PressAcademia

PressAcademia Procedia

ENGINEERING AND NATURAL SCIENCES

ISSN: 2459-0762

procedia@pressacademia.org

PressAcademia publishes journals, books and case studies and organizes international conferences.

http://www.pressacademia.org/procedia

PressAcademia Procedia

ABOUT PRESSACADEMIA PROCEDIA

Engineering and Natural Sciences

PressAcademia Procedia publishes conference proceedings. ***PressAcademia Procedia*** is a scientific, academic, refereed and open-access journal. The publication languages are English and Turkish. ***PressAcademia Procedia in Engineering and Natural Sciences*** invites all conferences in the area of engineering, engineering management and natural sciences. ***PressAcademia Procedia*** issues a DOI number for each manuscript published and a special volume number for each conference proceedings.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship,
May 12- 14, 2017, Istanbul, Turkey.

Year: 2017 Volume: 5

CONTENT

Title and Author/s **Page**

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

PERFORMANCE ANALYSIS OF A COOPERATIVE COMMUNICATION SYSTEM USING MASSIVE MIMO

DOI: 10.17261/Pressacademia.2017.563
PAP-WCTIE-V.5-2017(1)-p.1-9

Sidi Ethmane Amar[1], Tansal Gucluoglu[2]
[1]Yildiz Technical University. sidijedou@gmail.com
[2]Yildiz Technical University. tansal@yildiz.edu.tr

ABSTRACT

Massive MIMO will be an enabling technology for the deployment of the 5th generation of wireless systems. In the past three years, this technology has been put under the spotlight. Although there is a lot of research work on this topic, a number of challenges still need to be surpassed before reducing Massive MIMO to practice. In this paper, I studied and explored a cooperative communication scenario using massive MIMO at the base station. Based on the work previously done by the project WINNER and the Vienna LTE Simulator, we generated a propagation channel model and simulated the throughput of the system model using different types of schedulers and transmission modes.

Keywords: Massive MIMO, cooperation, scheduler, CoMP.

1. INTRODUCTION

In the last decade, data traffic has increased in a very exponential way due to the easy access to the internet driven by a dramatic growth and evolution of both wireless data consuming devices (like smartphones, tablets, and laptops) and broadband networks' infrastructure. In other words, the demand for wireless data traffic will increase even more dramatically in the foreseeable future. Thus, new disruptive technologies are required to meet this huge demand. The key factor to take into consideration in wireless data traffic is the wireless throughput (bits/s) which is defined as: Throughput = Bandwidth (Hz) × Spectral efficiency (bits/s/Hz). Evidently, to enhance this throughput, some novel technologies which can increase the bandwidth or the spectral efficiency or both should be explored.

Transmission with multiple-input multiple-output (MIMO) antennas is a well-known diversity technique to enhance the reliability of the communication. The more antennas the base station BS is equipped with, the more degrees of freedom are offered and hence, more users can simultaneously communicate in the same time-frequency resource. As a result, a huge sum throughput can be obtained.

With large antenna arrays, conventional signal processing techniques (e.g. maximum likelihood detection) become prohibitively complex due to the high signal dimensions. The main question is whether we can obtain the huge multiplexing gain with low-complexity signal processing and low-cost hardware implementation. Research has shown that the use of an excessive number of BS antennas compared with the number of active users makes simple linear processing nearly optimal in [1].

Due to the complexity and deployment consideration in practical scenarios at individual BSs, each BS site cannot be deployed with a large number of antennas. That means with a limited number of antennas, the inter-cell and intra-cell interference still exist if simple non-cooperative linear precoding is used individually in each base station site [2]. Cooperative massive MIMO [CM-MIMO] where multiple BSs cooperate together and form a distributed antenna array to serve multiple users simultaneously is an attractive alternative. In CM-MIMO, user data as well as CSI (channel state information) is shared among BSs that will provide more degrees of freedom for communication. Also, precoding can take into account inter-cell interference and thus mitigate inter-cell interference, which is especially critical for cell edge users that typically suffer more inter-cell interference.

Table 1: The Cisco VNI Forecast —Historical Internet Context

Year	Global Internet Traffic
1992	100 GB per day
1997	100 GB per hour
2002	100 GBps
2007	2,000 GBps
2015	20,235 GBps
2020	61,386 GBps

Source: Cisco VNI, 2016

Furthermore, CM-MIMO, where multiple BSs coordinate through the backhaul network, the bandwidth of the backhaul link and delay may create additional impairments on the system performance [3], [4]. In this paper, system level simulation performance of cooperative massive MIMO scenario (specifically with coordinated multipoint CoMP technique) is compared to other non-cooperative massive MIMO scenarios under the uniform framework of the LTE-A system. Here, the analysis is based on comparing the results of different simulations using the same setup, varying only the scheduler type and the transmission mode to see how it will affect the throughput and the spectral efficiency of the UEs. This analysis provides insight on the potential system performance that can be achieved by using cooperative massive MIMO. The rest of the paper is organized as follows. In Section 2, the 3GPP 3D channel model used for simulation is described. Section 3 presents the system simulation setup and simulation results. Finally, our conclusion is presented in Section 4.

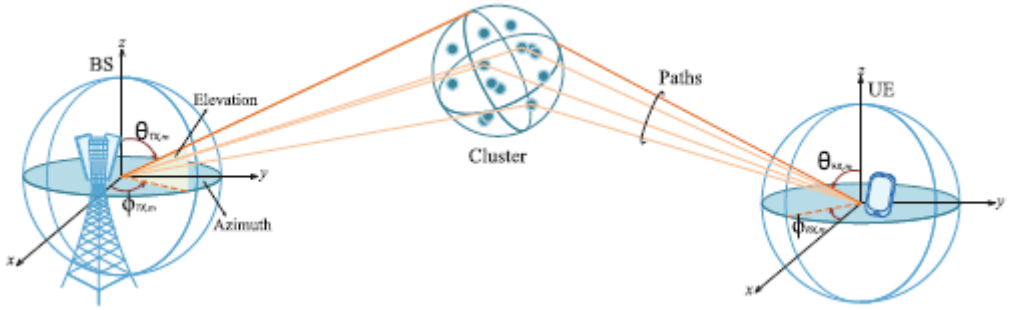
2. The 3GPP 3D Channel Model

The 3GPP 3D channel model, used in this paper, characterizes wireless communication channels of typical European cities. It is a 3D geometric stochastic model, describing the scattering environment between BS and UE in both azimuth and elevation dimensions. The scatterers are represented by statistical parameters without having a real physical location. In 3GPP TR 36.873 [5], three scenarios, Urban Macro cell (UMa), Urban Micro cell (UMi) and UMa-high rise (UMa-H) are specified. They represent typical urban macro-cell and micro-cell environments. Both UMa and UMa-H scenarios, consider a BS height of 25m, thus surpassing the surrounding buildings. UMa-H also specifies such environments with one high-rise building per Evolved Base Station (eNodeB) sector. UMi, considers a BS height of 10m, lying below the rooftop level. All three environments are assumed to be densely populated with buildings and take into account both indoor- and outdoor UEs.

The 3GPP 3D channel model specifies three propagation conditions, Line-Of- Sight (LOS), Non-Line of Sight (NLOS) and Outdoor-to-Indoor (O-to-I). For each of these conditions it defines different parameters for mean propagation path loss, macroscopic fading and microscopic fading. All three scenarios in [5], UMa, UMi and UMa-H, consider 80% of the UEs to be located indoors. The probability of being in LOS is determined separately for indoor and outdoor UEs and depends on the height of the UE as well as the break point distance. The break point distance characterizes the gap between transmitter and receiver at which the Fresnel zone is barely broken for the first time [6]. For an indoor UE, LOS refers to the signal propagation outside the building in which the UE is located.

For each UE location, large scale parameters are generated according to its geographic position as well as the propagation conditions at this location. The large-scale parameters incorporate shadow fading, the Ricean K-factor (only in the LOS case), delay spread, azimuth angle spread of departure- and arrival, as well as zenith angle spread of departure and arrival. The small-scale parameters incorporate delays, cluster powers as well as angles of departure and -arrival in azimuth and elevation direction, respectively. The model considers N clusters of scatterers, where each cluster is resolvable to M paths. A simplified sketch of the model is given in Fig. 1.

Figure 1: Scattering concept in the 3D model. Figure demonstrates a link that is resolvable to M paths. Elevation- and azimuth angles at BS and UE are denoted as θ and ϕ, respectively [8]

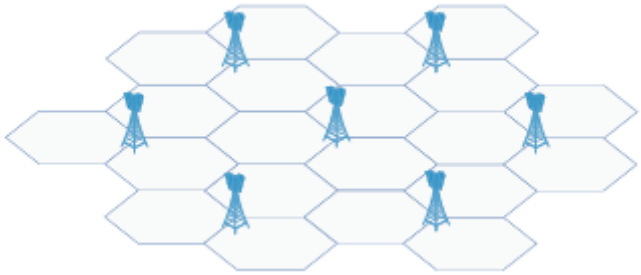


3. SYSTEM LEVEL SIMULATION

3.1. Simulation Setup

The macro-sites are arranged according to a hexagonal grid with an inter-site distance of 500m. Each site is equipped with three sector eNodeBs; each eNodeB employs *4*x2 antenna configuration on the macro-site. The minimum distance between a UE and its associated eNodeB is 35m.

Figure 2: Hexagonal-grid tri-sector macro-cell scenario. [8]



In the table 2. below, there is a detailed list of the employed simulation settings. We simulate at the center frequency of 2 GHz and an LTE bandwidth of 10 MHz; the total transmit power per eNodeB is 40W. The eNodeBs on the macro-site employ directional antennas. The signal propagation conditions follow the model of the 3GPP 3D channel, as mentioned previously. Each UE has two receive antennas and employs a Zero-forcing receiver ZF within each eNodeB sector. We assume 10 active users which are uniformly distributed and move at the speed of 5 Km per hour. We employ closed loop spatial multiplexing CLSM corresponding to transmission mode 4 of LTE. The feedback comprises channel quality indicator CQI, precoding metrics indicator PMI, and rank indicator RI. The feedback is delayed by three transmission time intervals TTI and computed with perfect channel knowledge. Resources are assigned according to CoMP scheduler. We assume a simple full buffer traffic model and simulate over a total length of 20TTI.

Table 2: Simulation Parameters

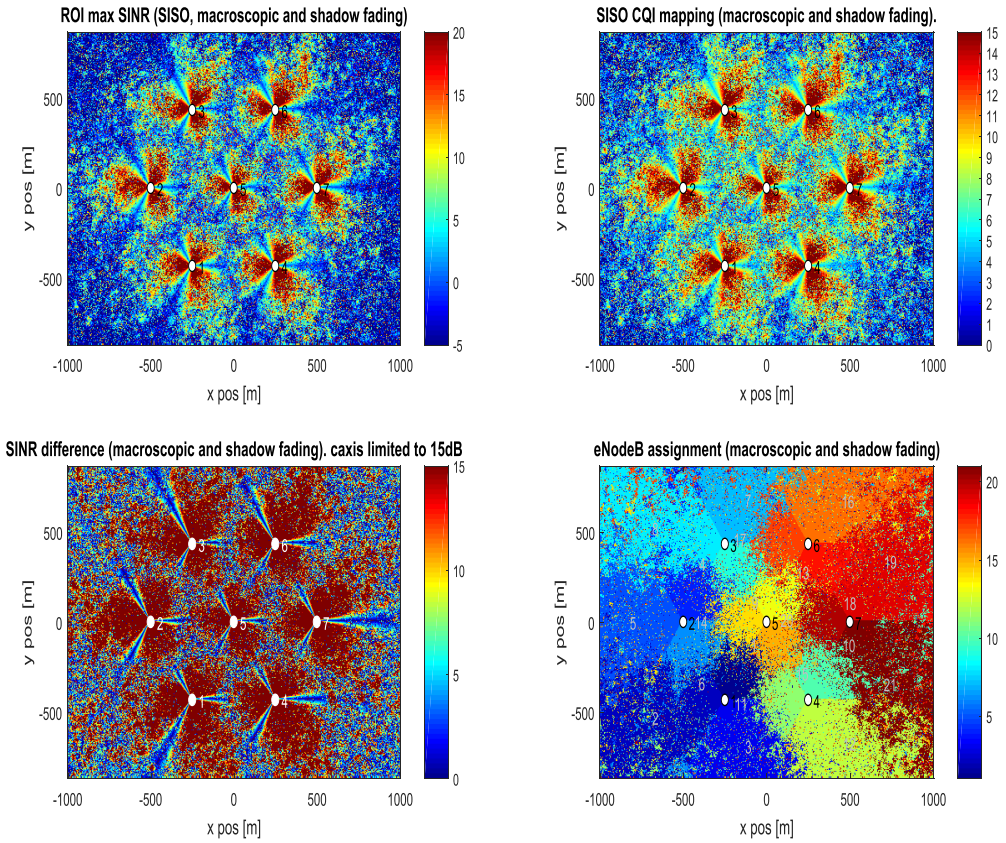
PARAMETER	VALUE
CARRIER FREQUENCY	2 GHz
LTE BANDWIDTH	10 MHz
MACRO-SITE DEPLOYMENT	Hexagonal grid
SCENARIOS	3D-UMa
ENODEB ANTENNA HEIGHT (UMA)	25 m
ENODEB TRANSMIT ANTENNA CONFIGURATION	N_{Tx} = 4, each has 8 ports

ENODEB RECEIVE ANTENNA CONFIGURATION	N_{Rx} = 2, each has 8 ports
ENODEB ANTENNA ELEMENTS PER PORT	M = 10
POLARIZED ANTENNA MODELING	Model 2
ENODEB ANTENNA POLARIZATION	X-pol *(+-45°)*
UE ANTENNA POLARIZATION	X-pol *(0/ + 90°)*
VERTICAL ANTENNA ELEMENT SPACING	0.5λ
HORIZONTAL ANTENNA ELEMENT SPACING	0.5λ
MAXIMUM ANTENNA ELEMENT GAIN	8dBi
UE ANTENNA PATTERN	Isotropic antenna gain
ELECTRICAL DOWNTILT	12°
UE DISTRIBUTION	Uniform in cell [5, Table 6-1]
NUMBER OF ENODEBS	21
ENODEB TRANSMIT POWER	40 W
LTE TRANSMISSION MODE (MIMO MODE)	Closed Loop Spatial Multiplexing CLSM, Mode 4
ANTENNAS PER UE	2
RECEIVER TYPE	Zero-forcing ZF
ACTIVE UES PER ENODEB	10
UE SPEED	5 Km/h
TRAFFIC MODEL	Full buffer
CHANNEL KNOWLEDGE	Perfect
FEEDBACK	AMC: CQI, MIMO: PMI and RI
FEEDBACK DELAY	3 TTI
SCHEDULER	CoMP,
SIMULATION LENGTH	20 TTI

3.2. Simulation Results

The simulator initially generates macro-site and the associated eNodeBs, then it computes the site specific pathloss and shadow fading maps. Based on the pathloss and the shadow fading maps, the simulator computes a maximum SINR map as shown in the upper left corner of Fig. 3 below. This map is then applied to calculate the UEs associations regions as shown in the lower right corner of Fig. 3. The spectrum patterns stem from the shadow fading. In each association region, the simulator uniformly distributes the specified number of UEs, in our case 10; then it initializes the schedulers and the resource blocks for each eNodeB. The next step is to calculate the fast fading channels.

Figure 3: Macroscopic and Shadow Fading



Our simulation employs mutual information based SNR mapping that maps the vector of post equalization SINR values from the link quality model to an effective SNR value in terms of mutual information; then using a lookup table and selecting the curve for the applied coding and modulation scheme allows to predict a block error ratio of the transport block.

Figure 4: SNR-to-CQI Mapping - CQI BLER Curves

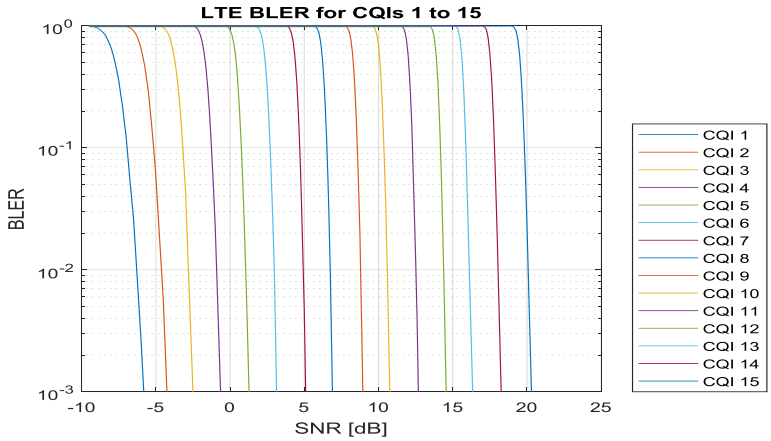
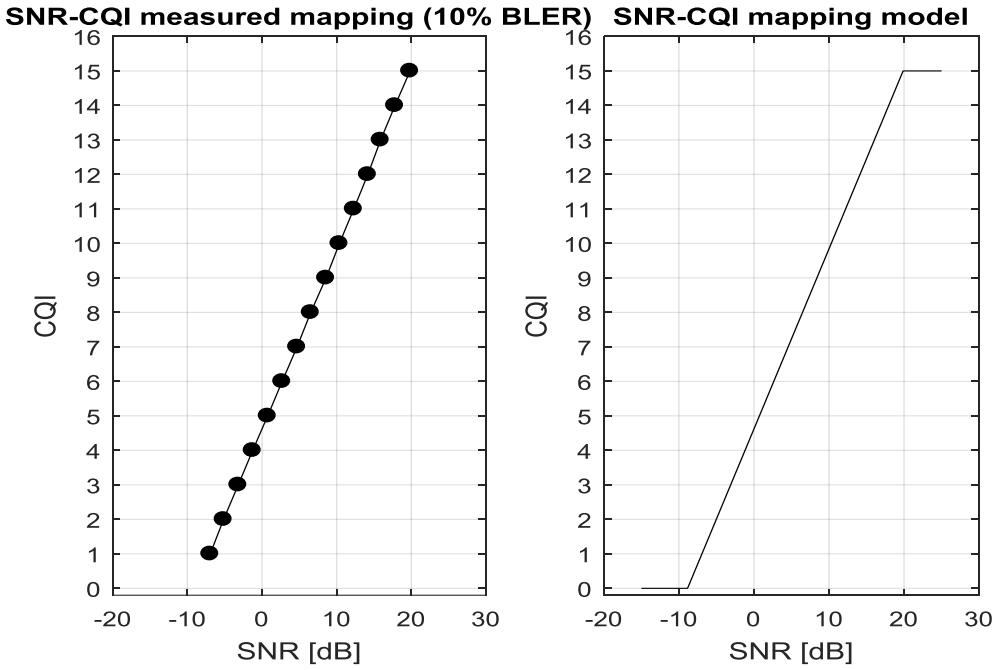
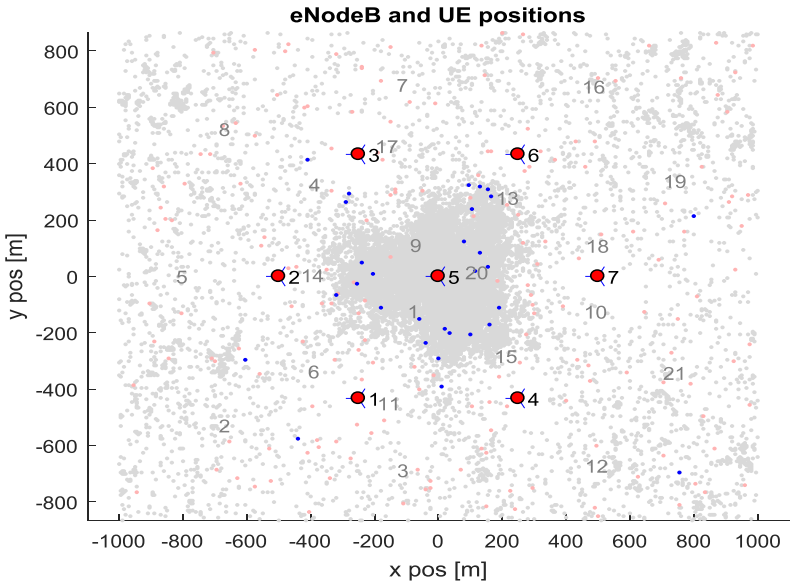


Figure 5: CQI Mapping Obtained from the 10% BLER Points



The presented curves were obtained with the Vienna LTE Advanced Link Level Simulator and form the only computationally costly physical layer evaluation. Eventually, combining the block error ratio with the size of transport block in bits yields the effective throughput.

Figure 6: eNodeBs and UEs Positions



The figure Fig. 6 shows the simulated network with the bold red dots denoting the macro-site, and the blue and grey dots correspond to the UEs.

Figure 7 contains empirical cumulative distribution function for the average UE throughput, the average UE spectral efficiency and the UE wideband SINR. It also provides plots for the average UE throughput and the UE spectral efficiency Vs. the wideband SINR.

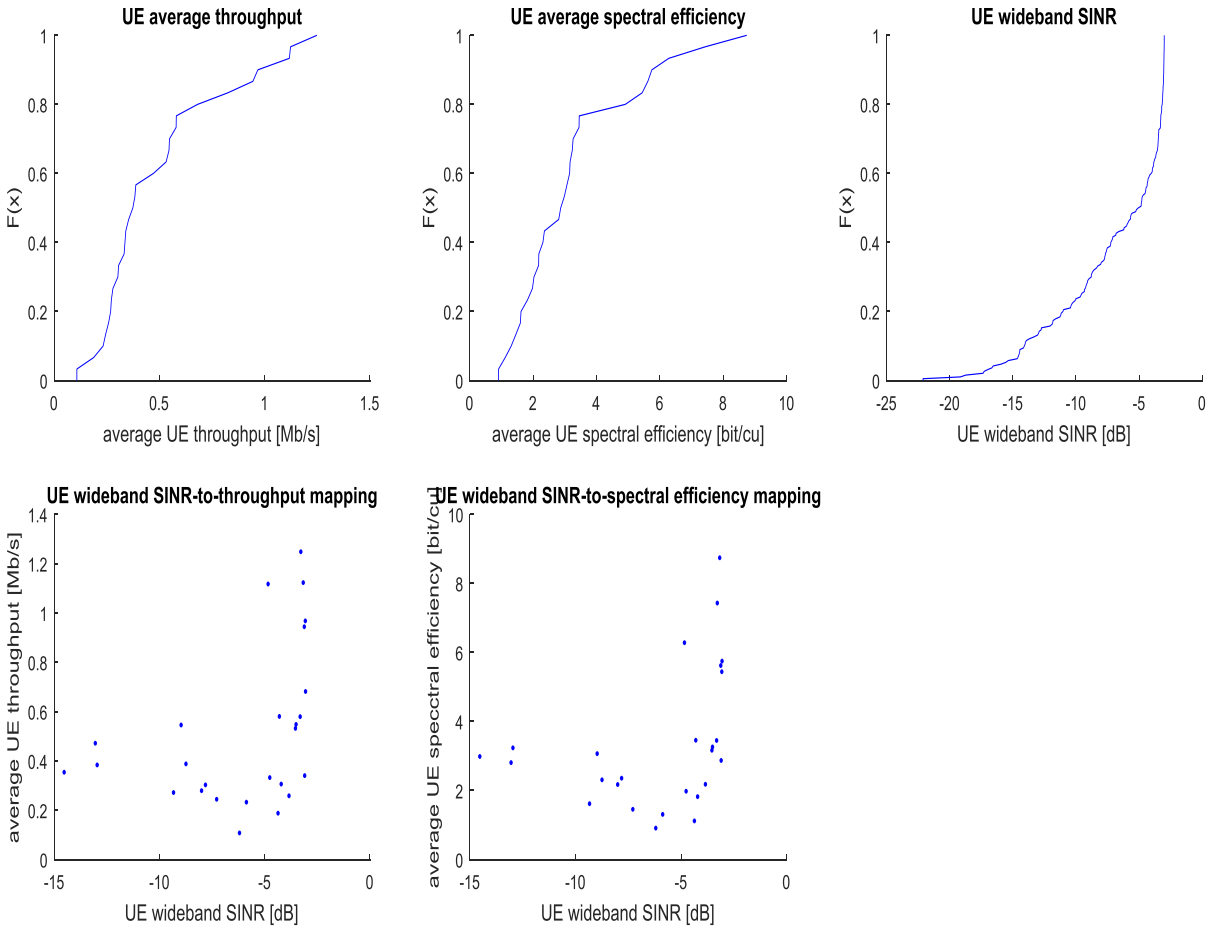


Figure 7: Plots show the throughput ECDF for the average UE throughput (upper-left), spectral efficiency (upper-middle), wideband SINR (upper-right); the mapping between the wideband SINR and the average throughput for each UE (lower-left) and the mapping between the wideband SINR and the spectal efficiency (lower-right).

3.3. Comparison and Evaluation of Results

In this part, we evaluate the results of different simulations under the same previous configuration. Using the transmission mode closed loop spatial multiplexing CLSM with three types of schedulers round robin, proportional fair, and CoMP in contrast to the transmission mode MU-MIMO with round robin scheduler, we obtained the aggregate results in the table below. Table. 3 contains the UE throughput values, the 95% (peak), average, and 5% (edge); the average UE spectral efficiency and the average number of resource blocks per TTI and UE; as well as the statistics of mean RB occupancy and Fairness index.

Table 3: Results of Different Simulations

Transmission Mode	Scheduler	Peak/Avg./edge UE throughput (Mb/s)	Avg. UE spectral eff. (bit/cu)	Avg. RBs/TTI/UE (RBs)	Average cell throughput (Mb/s)	Mean RB occupancy	Fairness index
CLSM	round robin	3.51/1.49/0.15	1.77	5	14.88	100%	0.490905
CLSM	Prop Fair Sun	3.81/1.21/0.12	1.71	5	12.15	100%	0.423385
CLSM	CoMP Intra-site	2.07/0.67/0.11	2.38	1.67	6.73	36.67%	0.568229
CLSM	CoMP Global	1.12/0.50/0.19	3.28	0.92	5.04	25.61%	0.737139
MU-MIMO	round robin MU	3.79/1.40/0.14	1.52	5.32	13.96	25.62%	0.621665

According to the comparative results shown in Table 3. Above, our cooperative scenario yields the least average UE and cell throughputs but it has slightly better edge UE throughput and in terms of resources allocation, it has the best average UE spectral efficiency and fairness index.

Figure 8: UE Average Throughput

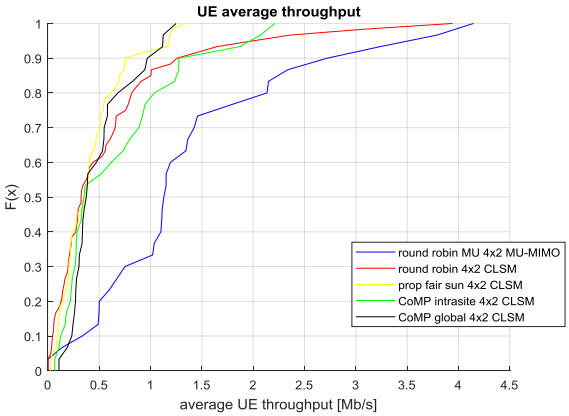


Figure 9: UE Average Spectral Efficiency

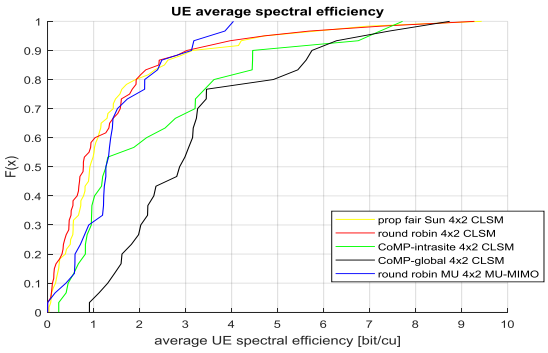
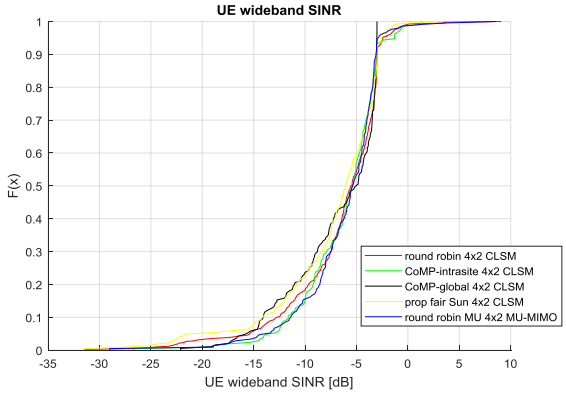


Figure 10: UE Wideband SINR



4. CONCLUSION

In this paper, downlink system level simulation performances of noncooperative and cooperative massive MIMO scenarios were evaluated and compared based upon current LTE-A system, considering different transmission schemes and scheduling types. The analysis of the results shows that the deployed cooperative scenario has better performance for the

cell edge users and it significantly improves the spectral efficiency of the system, whereas the cell average throughput is degraded owing to the power imbalance of the cell center users. These results go in accordance with the main reason for which the DL/UL Coordinated Multi-Point transmission/reception (CoMP) technique was introduced; which is to improve the system capacity and cell edge user throughput.

REFERENCES

[1] F. Rusek, D. Persson, Buon Kiong Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[2] C. He and R. D. Gitlin, "Limiting Performance of Massive MIMO Downlink Cellular Systems," in *Information Theory and Applications Workshop (ITA)*, Feb. 2016.

[3] R. W. Heath, T. Wu, Y. H. Kwon, and A. C. K. Soong, "Multiuser MIMO in Distributed Antenna Systems with Out-of-Cell Interference," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4885–4899, Oct. 2011.

[4] R. Zakhour and S. V. Hanly, "Base Station Cooperation on the Downlink: Large System Analysis," *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2079–2106, Apr. 2012.

[5] 3rd Generation Partnership Project (3GPP). Study on 3D channel model for LTE, 3rdGeneration Partnership Project (3GPP), TR 36.873 (2014)

[6] H. Masui, T. Kobayashi, M. Akaike, Microwave path-loss modeling in urban line-of-sight environments. IEEE J. Sel. Areas Commun. **20**(6), 1151–1155 (2002). doi :10.1109/JSAC. 2002.801215

[7] Chao He and Richard D. Gitlin "System Performance of Cooperative Massive MIMO Downlink 5G Cellular Systems" Department of Electrical Engineering University of South Florida Tampa, Florida 33620, USAIEEE WAMICON 2016 April 11-13, 2016

[8] Markus Rupp, Stefan Schwarz and Martin Taranetz "The Vienna LTE-Advanced Simulators Up and Downlink, Link and System Level Simulation" ISSN 1860-4862 ISSN 1860-4870 (electronic) Signals and Communication Technology ISBN 978-981-10-0616-6 ISBN 978-981-10-0617-3 (eBook) DOI 10.1007/978-981-10-0617-3

[9] M. Taranetz, T. Blazek, T. Kropfreiter, M. Mller, S. Schwarz, M. Rupp: "Runtime Precoding: Enabling Multipoint Transmission in LTE-Advanced System-Level Simulations"; IEEE Access, 3 (2015), 725 - 736

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

EVALUATION OF CARBON FOOTPRINT FOR WOOD BASED PANEL INDUSTRY IN TURKEY

DOI: 10.17261/Pressacademia.2017.564
PAP-WCTIE-V.5-2017(2)-p.10-18

Mehtap Erdil[1], Nural Yilgor[2], Yigit Gungor[3]
[1]Istanbul University, erdilmehtap@gmail.com
[2] Istanbul University, yilgorn@istanbul.edu.tr
[3] Gebze Technical University, yigitgungor@hotmail.com

ABSTRACT

In wood-based panel industry, the importance of carbon footprint has increased due to global warming and climate change all around the world. So this topic has been studied extensively in the world, especially in developed countries. Eventhough wood based panel industry is one of the most important industry in Turkey there is not any comprehensive study in this field so this study was planned to meet the requirement in industry which release a lot of emissions to the environment. In recent years the use of wood-based panels such as fiber board, particleboard, oriented strand board, plywood etc., have been increased due to the fast growth in bulding trade in Turkey. The world's largest board producers are China, USA, Germany and Turkey respectively. This charges the industry a great responsibility for the environment because sustainable development is a development that meets the needs of the present without making a concession the ability of future generations to meet their own needs. From this perspective, it is essential to know the current status of the sector concerning CO_2 footprint, energy footprint and water footprint. This study is aimed to investigate one of the largest board producer industry of Turkey whether clean production (less emission) is carried out or not. Besides, it is aimed to present some improvements to decrease the emission. Carbon footprint values are calculated as statistically with Tier method during the particle board production, and Pareto anaysis method is used for determine the footprint' problem. Consequently, a liveable environment will be provided and environmental oriented production will be supported to contribute to these properties in the industry for this issue in the scope of this research. Also, this study shall provide a general view and perception for the importance of the carbon footprint in the industrial sector.

Keywords: Carbon footprint, tier method, pareto analysis, wood based panel industry, particle board production.

1. INTRODUCTION

Carbon footprint originated from the terminology of ecological footprint which was proposed by Wackernagel and Rees, 1996 (Wackernagel and Rees, 1996). It is the calculation of the total value of direct or indirect CO_2 emissions of the activities occur during the life cycle. According to another definition, it is a technique to identify and measure of the greenhouse gas emissions (GHG) release from each operation or activity full life cycle (Carbon Trust, 2007; IPCC, 2006; IPCC, 2007). Grubb and Ellis, 2007, defined carbon footprint as the total amount of carbon dioxide release through the combustion of fossil fuels. The major greenhouse gases emitted into the atmosphere are CO_2, CH_4, N_2O and some different fluorine containing halogenated compounds (Muthu, 2014). However, the most significant greenhouse gas is CO_2 and it was produced from burning of fossil fuels to generate energy which is essential for the manufacture and transport of the goods. It is expressed as grams or kilograms of CO_2 equivalent per kilowatt hour of generation ($g/kgCO_{2eq}/kWh$) which being in charge of the different global warming effects of other greenhouse gases (Paliamentary Office of Science and Technology (POST), 2006). Hammond, 2007 suggested that ..." The property that is often referred to as a carbon footprint is actually a carbon weight of kilograms or tonnes per person or activity." Higher GHG concentrations in the earth`s atmosphere cause global warming and it causes climate change in the world. This reality forces the nations to get some precautions. The most important agreement of concerning global warming and climate change is known as Kyoto Protocol and six greenhouse gases are defined as CO, CO_2, CH_4, N_2O, perfluorocarbons, and hydrofluorocarbons which cause strongly global warming and it was thought those gases are released by human activities. (IPCC, 2007; ECCM, 2008; UN,1998). A carbon footprint measurement is a complicated and boring method. GHG produced by human activities either directly or indirectly. Direct GHG emissions are suggested as burning fossil fuels for generating electricity to heat and transport. But indirect emissions are less clear that

emitted through life-cycles of goods and services (ECCM, 2008; Wiedmann and Minx, 2007). In this study, it is aimed to calculate CO_2 emissions for each process in a particle board industry in a plant scale one of the largest producers in Marmara region, in Turkey. The plant named as X plant afterwards in this study. The amount of annual production of particle board is 423096 m^3/year in 2015. The carbon footprint is calculated according to Tier 1 method (IPCC, 2007).

2. LITERATURE REVIEW

2.1. Forest Products Industry

The forest products industry has been a strong power of economy for human since olden times (Mahapatra and Mitchell, 1997; Ok, 2005, İlter and Ok, 2007). So the use and supply of wood for the industry have a great importance from past through today (Tezcan et al., 2014). In recent years the use of wood-based panels such as fiberboard, particleboard, oriented strandboard, plywood etc., have been increasing due to the fast growth in building trade in Turkey. The world's largest board producers are China, USA, Germany and Turkey respectively. This charges the industry a great responsibility for the environment because sustainable development is a development that meets the needs of the present without making a concession the ability of future generations to meet their own needs (Tezcan et al., 2014; Yıldırım et al., 2014).

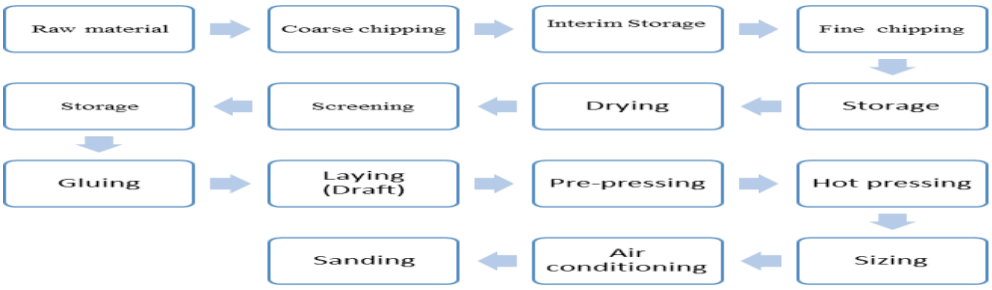
2.1.1. Researches About Carbon Footprint and Pareto Analysis

There are many studies about carbon footprint in wood industries throughout the world, but, there is not any research in our country even it is one of the most important developed industry in Turkey. In a study (Garcia et al., 2013) researched the carbon footprint of particleboard produced in Portugal. It was aimed the effect of different methodologies in the particleboard carbon footprint (CF) calculation in using four different CF measurement methods. Those methods are ISO-TS14067, the GHG Protocol Product Standard, PAS 2050, and Climate Declaration. In another study (Görener and Toker, 2013), applied in pareto analysis to a firm engaged in forest products industry which is specialized on MDF board (Medium Density Fiber) manufacturing. They aimed to determine and classify failure modes and then make suggestions due to their significance degree by Pareto analysis. They also investigated the occurrence of waste process by using Pareto analysis. In a study researched different uses of wood to see their efficiency by means of carbon and energy impacts to displace fossil energy (Lippke et al., 2012). They suggested when waste wood is consumed as biofuels instead of fossil fuels, the emissions were decreased. In another study (Bergman et al., 2014), researched in the carbon effects of wood products. This study defines to carbon emissions savings when wood products are used in constructing buildings in place of nonwood sources. In another study (Cetin et al., 2014), applied in pareto analysis on the scope and extent of extra work caused by management and workers' issues in the Turkish furniture industry. The aim of this study is to prove and detect factors that decrease of efficiency through issues of management, production processes, supervision of workers and aspects of the products themselves, thereby assisting enterprises obtain essential measures. This study was analyzed with using pareto charts and cause effect diyagram. In another study (Moroşanu et al., 2011), researched on identifying and analyzing defects on oak veneer for four regions. So, the pareto analysis was applied for developing the quality of the studied products.

3. DATA AND METHODOLOGY

3.1. Particleboard Manufacturing Process

Figure 1: Work Flow with regard to Particle Board Production in X Plant



X is a plant operating in forest products industry and it produces particle board, medium density fiberboard (MDF), and parquet in Marmara region. Work flow in connection with particle board manufacturing is shown in Figure 1. Particle board is described as a panel product manufactured from lignocellulosic materials, in the form of chipped particles in the combination of synthetic resin and suitable binders by the help of heat and pressure. The study is carried out in the particle board plant (X) considering the improvements of all of the energy flow processes comprises of following steps:

1-Occuring a study plan

2-Calculation of carbon footprints for each of process

3-Applying the Pareto analysis procedures

4-Drawing the Pareto diagram

5-Determine major emission problem(s) according to 80/20 law

6-Give suggestions for major emission problem(s)

3.2. Methodology

3.2.1. Tier Method

According to IPCC 2007 there are three tiers methods to evaluate emissions. Since the simplicity in application and suitability of the data that we have, Tier 1 method used in this study. Carbon footprints (CF) were calculated for each process with the inputs' emission factors and then Tier 1 method was applied. Process and machinery based data related to energy and fuel consumption used for emission calculations through the formula given below (Pekin, 2006; Atabey, 2013; Turanlı, 2015). Before carbon footprint calculation, it must be known fuel consumption and emission factor. Emission factor can be found in literature (Defra, 2010; Lelyveld and Woods, 2010; Url-1; Url-2). Emission is calculated according to the formula 1 is given in below and CF is calculated according to formula 2.

Emission = Energy consumption x Emission factor x Oxidation factor (1)
(Oxidation factor is taken as 1)

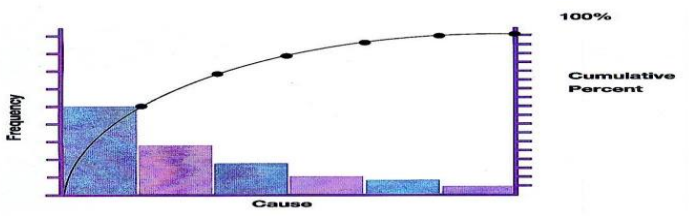
CF = Emissions ($kgCO_2$)/ Amounts of annual production (m^3) (2)

3.2.2. Pareto Analysis

Vilfredo Pareto was a 19th century Italian economist and analyzed economic problems using mathematics and this method was mentioned as his name and it helps to identify and classify the fault according to percentage significant. He observed that 80 percent of the land in Italy was owned by 20 percent of the population. Then. It is a way of evaluating reasons of problems to provide an effective solution. This method applies due to 80/20 law in general. Cause and effect diagram is useful tool in identifying the major causes. This diagram helps to build a relationship, brainstorming is done with utilizing these quality tools to provide an effective solution (Gitlow et al., 2005). Quality expert J.M. Juran used the rule to quality control and shown that 80 percent of problems stem from 20 percent of possible reasons. Pareto diagrams are the graphical tool applied in Pareto analysis (Cravaner et al.,1993; Leavengood and Reeb, 2002). Pareto analysis is a method which is applied to distinguish reasons from less important ones (Akın, 1996; Ozcan, 2001). Pareto analysis takes the procedures in below:

1. Problems should be identified and then classified
2. Data are classified according to the problem types. Total values which are in different categories and their percentages are defined.
3. A bar chart was drawn. In this bar graph, while the y-axis demonstrates the totals and percentages, the x-axis shows the classified groups.
4. Pareto diagram is realized to state qualitative totals therefore starting from the upper right-hand corner of the first bar (Akın and Oztürk, 2005) (Figure 2).

Figure 2: Pareto Curve (Url-3)



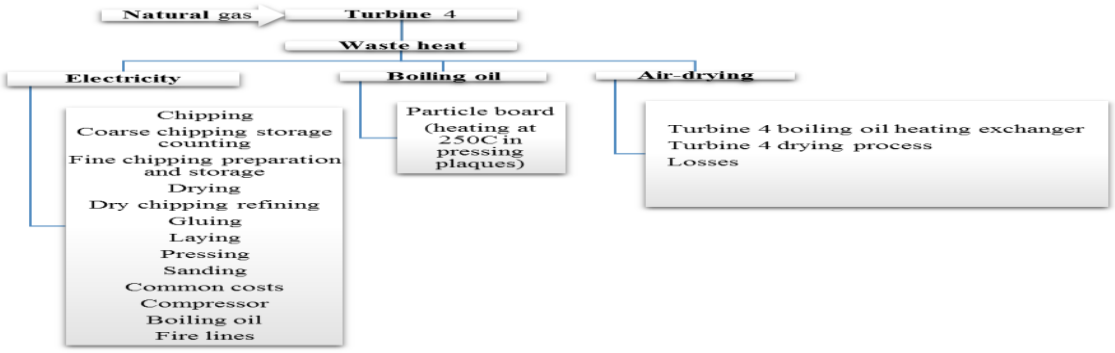
While failure modes are shown on x-axis, the frequency or cost are generally shown on y-axis on Pareto diagram.

4. FINDINGS AND DISCUSSIONS

4.1.Particle Board Manufacturing Process

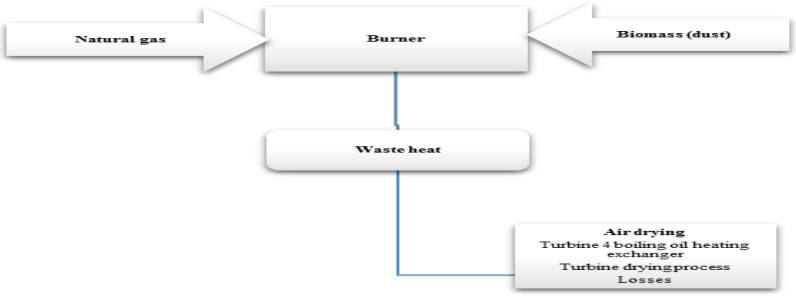
In this plant, natural gas, biomass (dust, wood chips, fiber, broken plate, etc.) and diesel fuel are using as the main inputs for obtaining energy. While natural gas and biomass are using in particle board production process, diesel fuel is consumed by transportation equipment (volvo, escalator and forklift) which are using in the field. Those inputs are primary and direct energy sources for particle board production in process. Since diesel value is very low, it is neglected so the main energy inputs of the process are natural gas and biomass. The total amount of natural gas and biomass as kWh is 232661527. While the ratio of natural gas consumed is about 75 %, biomass ratio is about 25%. Energy flow diagrams were shown in below.

Figure 3: Natural Gas Energy Flow Diagram



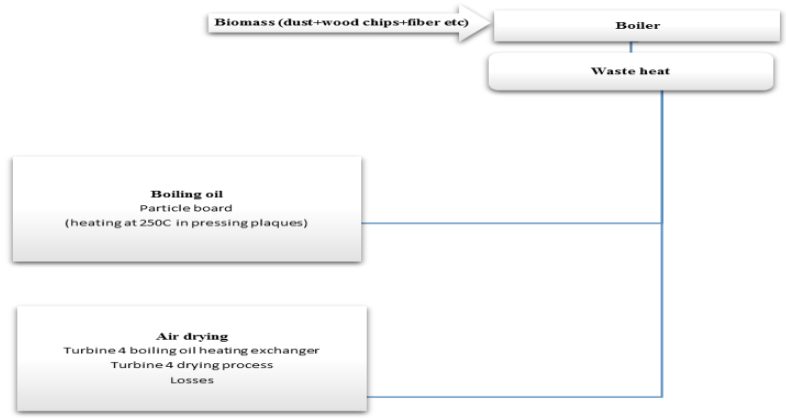
Natural gas is consuming in Turbine 4. As a result of the use of natural gas, waste heat releases. The waste heat is recovering in electricity, boiling oil, and air drying energy sources. Those sources are named as indirectly energy sources or seconder energy sources. So the company produces its own electricity in the plant.

Figure 4: Natural Gas and Biomass Energy Flow Diagram



At the same time natural gas and biomass are burned together in the burner to produce energy. As a result of the process, waste heat is also emitting. The waste heat is recovering in air drying energy sources. Those sources are called as indirectly energy sources or seconder energy sources.

Figure 5: Biomass Energy Flow Diagram



Biomass sources vary from wood dust, wood chips, bark, emery powder, etc. It is generally consisting of process waste. As a result of the process, waste heat is also emitted. The waste heat is recovered in boiling oil and air drying energy sources. Those sources are called as indirectly energy sources or seconder energy sources.

4.1. Application of Pareto Analysis

In this study, CF was calculated for each process sources in particleboard production. Pareto chart was drawn to identify the problems which were revealed by the help of 80/20 law. For this purpose, firstly, CF values of every process' sources were enumerated as shown in Table 1.

Table 1: CF Data According to Source

Serial No	Source	CF (kg CO2/m^3 particleboard)
1	Turbine 4	72.65
2	Burner	5.42
3	*Chipping*	0.62
4	*Coarse chipping storage*	0.73
5	*Fine chipping prepration and storage*	2.89
6	*Drying*	4.43
7	*Dry chipping refining*	2.26
8	*Glueing*	1.06
9	*Laying*	1,01
10	*Pressing*	1.85
11	*Sanding*	0.98
12	*Common costs*	0.36
13	*Compressor*	0.71
14	Boiling oil	0.21
15	Fire lines	0.01
16	Turbine 4 boiling exchanger	11.3
17	Turbine 4drying process	39.4
18	Losses1	9.3

19	Boiler	0,93
20	*Boiling oil boiler*	0.60
21	Losses 2	0.54
22	Transportation equipment	0.2

Then enumerated values were sorted from high to low and the total amount of CF was found as seen in Table 2.

Table 2: CF datas in sorted from high to low

Serial No	Source	CF (kg CO_2/m^3 particle board)
1	Turbine 4	72.65
17	Turbine 4 drying process	39.4
16	Turbine 4 boiling exchanger	11.3
18	Losses1	9.3
2	Burner	5.42
6	*Drying*	4.43
5	*Fine chipping preparation and storage*	2.89
7	*Dry chipping refining*	2.26
10	*Pressing*	1.85
8	*Glueing*	1.06
9	*Laying*	1.01
11	*Sanding*	0.98
19	Boiler	0.93
4	*Coarse chipping storage*	0.73
13	*Compressor*	0.71
3	*Chipping*	0.62
20	*Boiling oil boiler*	0.60
21	Losses 2	0.54
12	*Common costs*	0.36
14	Boiling oil	0.21
22	Transportati,on equipment	0.2
15	Fire lines	0.01
TOTAL		157.5

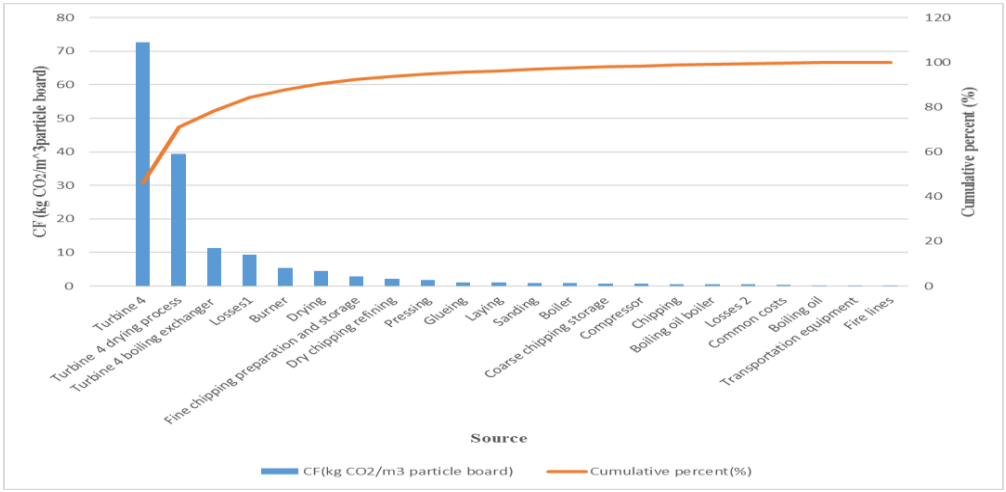
In the third step, as seen in Table 3, percent and cumulative percent were calculated for every sources' of values were took place in Table 2.

Pareto chart was drawn with 3 axes. While y axes in the left side shows CF values, y axes in the right side shows cumulative percent and x axes indicates the sources (Figure 6).

Table 3: Percent and Cumulative Percent of CF

Serial No	Source	CF (kg CO2/m3 particle board)	Percent (%)	Cumulative percent(%)
1	Turbine 4	72.65	46.131	46
17	Turbine 4 drying process	39.4	25.045	71
16	Turbine 4 boiling exchanger	11.3	7.155	78
18	Losses1	9.3	5.927	84
2	Burner	5.42	3.441	88
6	Drying	4.43	2.813	91
5	Fine chipping preparation and storage	2.89	1.836	92
7	Dry chipping refining	2.26	1.433	94
10	Pressing	1.85	1.173	95
8	Glueing	1.06	0.673	96
9	Laying	1.01	0.640	96
11	Sanding	0.98	0.622	97
19	Boiler	0.93	0.591	97
4	Coarse chipping storage	0.73	0.464	98
13	Compressor	0.71	0.448	98
3	Chipping	0.62	0,391	99
20	Boiling oil boiler	0.60	0.381	99
21	Losses 2	0.54	0.341	100
12	Common costs	0.36	0.227	100
14	Boiling oil	0.21	0.135	100
22	Transportation equipment	0.2	0.127	100
15	Fire lines	0.01	6E-03	100
TOTAL		157.5		

Figure 6: Application of Pareto Analysis according to Relationships between Source and Carbon Footprint



According to the Pareto chart, it was clearly seen that Turbine 4, Turbine 4 drying process and Turbine 4 boiling exchanger are the first three sources constitute 78% of the total sources (Figure 6).

According to the 80/20 law and the results that we obtained in the Pareto chart, the problems are originated from Turbine 4, Turbine 4 drying process and Turbine 4 boiling exchanger. While these causes of three sources are 22% of 22 sources, the first three sources occur 78% of total sources.

5. CONCLUSION

In this study, it was demonstrated that the total amount of 3 sources which take place in sequences of 22 sources in the process correspond to 78% of total amount of the sources with Pareto chart by the help of 80/20 law. So it can be suggested some improvements primarily for these 3 sources which are called Turbine 4, Turbine 4 drying process and Turbine 4 boiling exchanger. It can be offered that these sources may consume biomass energy instead of natural gas as an energy source. Besides, other renewables such as sun panels can be used as an energy source. Some best available techniques (BAT) can also be recommended. These techniques are explained below (Federal Environment Agency, 2011; BAT, 2014):

-Staff must be trained to improve environmental awareness periodically

-Environmental management system must be applied for control of procedures and carry out responsibilities by personnel

- Equipments' maintenance must be provided regularly

REFERENCES

Akın, B.1996, "ISO 9000 Uygulamasında İşletmelerde İstatistik Proses Kontrol Teknikleri", Istanbul: Bilim Teknik Yayınevi, pp 49-50.

Akın, B., Öztürk, E. 2005, "Istatistiksel Proses Kontrol Tekniklerinin Bilgisayar Ortamında Uygulanması", VII.Ulusal Ekonometri ve İstatistik Sempozyumu Bildiriler CD, Istanbul, pp 1-15.

Atabey, T. 2013, "Karbon Ayak İzinin Hesaplanması: Diyarbakır Örneği", Fırat Üniversitesi Yüksek Lisans Tezi, Elazığ.

Bergman, R., Taylor, A., Skog, K.E., Puettmann, M.E. 2014, "The Carbon Impacts of Wood Products", Forests Products Journal Vol 64 No. 7, 14 October.

Best Available Techniques (BAT) 2014, "Reference Document for the Production of Wood-based Panels"

Calculating our carbon footprint convert kWh electricity to kg CO2; draft, Url-2 http://www.leicsgov.ukcalculating_our_carbon_footprint.pdf

Carbon Trust, 2007, "Carbon Footprint Measurement Methodolgy" (Draft) www.carbontrust.co.uk

Cetin, T., Colak, M., Yılmaz, S., Yılmaz, T., Korkmaz, Y. 2014, "The Scope and Extent of Extra Work Caused by Management and Workers' Issues In The Turkish Furniture Industry", Journal of Asian Scientific Research 4 (12): pp 729-740, Muğla.

Cravener, T.L., Roush, W.B., Jordan, H., 1993, "Pareto Assessment of Quality Control in Poultry Processing Plants", Applied Poultry Science, 2: pp 297-302.

DEFRA,2010, biomass emission factors.

Federal Environment Agency, Brigitte Zletlow. lichtl Ethıcs& Brands, May 2011, "Environmental Standards In The Textile And Shoe Sector-A Guideline On The Basis Of The Brefs-Best Available Techniques Reference Documents Of The EU-Umwelt Bundes Amt"

Garcia, R., Freire, F., 2013, Carbon footprint of particleboard: a comparison between ISO/TS 14067, GHG Protocol, PAS 2050 and Climate Declaration.

Gitlow, H.S., Openheim, A.J., Oppenheim, R., Levine, D.M. 2005, "Quality Management", Third Edition, McGrawHill, USA, pp 366-367.

Görener, A., Toker, K. 2013, "Quality Improvement in Manufacturing Process to Defective Products using Pareto Analysis and FMEA", Beykent University Journal of Social Sciences-BUJSS Vol.6 No.2, ISSN:1307-5063, Istanbul.

Grubb and Ellis, 2007, "Meeting the Carbon Challenge: The Role of Commercial Real Estate Owners", Users & Managers, Chicago.

Guidelines for Measuring and Managing CO2 emission from Freight Transport Operations, March 2011(draft) Url-1 http://www.cefic.org/Documents/IndustrySupport/Transport-and-Logistics/Best%20Practice%20Guidelines&20-%20General%20 Guidelines/CeficECTA%20Guidelines%20for%20measuring%20and%20managing%20CO2%20emissions%20from%20transport%20operations%20Final%20,30.03.2011.pdf

Hammond, G. 2007, "Time to give due weight to the 'carbon footprint' issue". Nature 445(7125): 256. http://dx.doi.org/10.1038/445256b.

IPCC Climate Change, 2007: "Synthesis Report.Contribution of Working Groups I, II, and III to the Fourth Assessment Report of the Intergovermental Panel on Climate Change (IPCC)", Geneva

IPCC, 2006, "Guidelines for National Greenhouse Gas Inventories, Chapter 10: Emissions from livestock and manure management", Retrieved from http://www.ipcc nggip.iges.or.jp/public/2006gl/pdf/4_10_Ch10_Livestock.pdf.

IPCC 2007 Revised 2007" IPCC Guideliness for National Greenhouse Gas Inventories: Revence Manual Intergovermental Panel on Climate Change", Cambridge University Press, Cambridge

İlter, E.; Ok, K. 2007, "Ormancılık ve Orman Endüstrisinde Pazarlama İlkeleri ve Yönetimi" (Forestry and Forest Industry Marketing Principles and Management), Ankara, Turkey.

Leavengood, S. and Reeb, J.,2002, "Part 3: Pareto Analysis and Check Sheets, Performance excellence in the wood products industry Statistical Process Control", EM 8771, January 2002.

Lelyveld, T., Woods, P., AECOM 2010, "Carbon emission factors for fuels -Methodology and values for 2013-2016", 6 October 2010.

Lippke, B., Gustafson, R., Venditti, R., Steele, P., Volk, T.A., Oneil, E., Johnson, L.,Puettmann,M.E., Skog,K. 2012, "Comparing Life-Cycle Carbon and Energy Impacts for Biofuel, Wood Product, and Forest Management Alternatives" , Forest Products Journal Volt.62 (4): pp 247-257.

Mahapatra, A.; Mitchell, C.P. 1997, "Sustainable development of non-timber forest products: implication for forest Management in India", Forest Ecology and Management 94: pp 15–29.

Managers net, Pareto analysis(the 80:20 rule),(draft) Url-3 http://www.managers-net.com/paretoanalysis.html

Morşanu, C., Dumitraşcu, A., E., Ciobanu, V. 2011, "The Influence of Oak Raw Timber Defects on Decorative Veneer Cutting" , Bulletin of the Transilvania University of Braşov Series II Foresty Wood Industry Agricultural Food Engineering Vol 4(53) No. 2

Muthu, S.S. 2014, "Assessment of Carbon Footprint in Different Industrial Sectors", Volume 1, Springer, Environemntal Issues in Logistics and Manufacturing, ISBN 978-981-4560-41-2, London.

Ok, K. 2005, "Idea Marketing in Forestry: Some Implications from the Turkish Forestry Experience". *Forest Policy and Economics* 7(4): pp 493-500.

Özcan, S. 2001, "İstatistiksel Proses Kontrol Tekniklerinden Pareto Analizi ve Çimento Sanayiinde Bir Uygulama" , C.Ü. İktisadi ve İdari Bilimler Fakültesi Dergisi, 2 (2): pp 151-174.

Pekin, M., 2006, "Ulaştırma Sektöründen Kaynaklanan Sera Gazı Emisyonları Yüksek Lisans Tezi", Haziran 2006, İTÜ

POST, 2006, "Carbon footprint of electricity generation". POSTnote 268, October 2006, Parliamentary Office of Science and Technology, London, UK. http://www.parliament.uk/documents/upload/postp n268.pdf.

Turanlı, A. 2015, "Estimation of carbon footprint: A case study for Middle East Technical University Mastar Thesis" , May 2015, Middle East Technical University.

UN, 1998, "Kyoto protocol to the United Nations framework convention on climate change", United Nations.

Wackernagel, M., and Rees, W. 1996, "Urban ecological footprints Why cities cannot be sustainable-And why they are a key to sustainability", Environmental Impact Assessment Review Volume 16, Issues 4-6, July-November 1996, pp223-248.

What is a Carbon Footprint? Report presented to The Swedish Forest Industries Federation and Timcon by The Edinburg Centre for Carbon Management ECCM-EM-483-2007, February 2008 (Draft) http://www.timcon.orgCarbonCalculatorCarbon%20Footprint.pdf

Wiedmann, T., and Minx, J. 2007," A Definition of 'Carbon Footprint' ", ISA UK Research Report 07-01, June 2007.

Yildirim, T., Candan, Z. and Korkut.S. 2014, "Wood-Based Panels Industry In Turkey: Future Raw Material Challenges And Suggestions. Maderas. Ciencia y tecnología 16 (2): pp 175-186.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

EFFECT OF HIGH CURING TEMPERATURE ON MECHANICAL PROPERTIES OF CONCRETE

DOI: 10.17261/Pressacademia.2017.565
PAP-WCTIE-V.5-2017(3)-p.19-23

Olanrewaju D.O[1], A.Y Akinsanya[2], O.A Olowu[3]
[1], [2]- Department of Building Technology, Lagos State Polytechnic, Nigeria. kemiolan@yahoo.com
[3] - Department of Building Technology, Yaba college of Technology, Nigeria.

ABSTRACT

It has been shown that Concrete increases in strength with age after setting. The strength at a particular age can be further increased by suitable curing of the concrete while it is maturing. Such curing comprises the application of heat and/or the preservation of moisture within the concrete. In this research work, an attempt has been made to study the influence of elevated curing temperature on mechanical properties of concrete. Mechanical properties such as comprehensive strength and pullout strength were investigated. Concrete Specimens were cured under three different temperatures and curing age 30^0C, 60^0C and 100^0C and 7, 14 and 28 days respectively. Compression test and pull out test were conducted on concrete cubes. The results showed strong positive relationship with correlation coefficient 0.779 between number of days and strength of concrete and the P–value 0.004 implies that the relationship is significant at 5% level while the coefficient of Determination (COD) indicates that number of days have 60.75% effect on strength of concrete. Correlation coefficient 0.932 with a P-value 0.013 and COD of 86.79% indicate that temperature has a very strong effect on strength of concrete. The regression analysis showed a unit increase in temperature will lead to 0.062 N/mm^2 increase in the strength of concrete. However the rate of concrete strength at the increase in temperature is lower than the increase in number of days, this implies a tendency for concrete to be weak in future with increase in temperature especially above boiling point.

Keywords: Concrete, curing, compressive strength, pullout strength, elevated temperature.

1. INTRODUCTION

Mixture of coarse aggregate, fine aggeregate, cemententious material and water in suitable proportions, placed and compacted wherever required, solidifies after a lapse of time into what is known as concrete. Concrete is one of the major materials often used for construction work. Concrete is expected to function for its expected life span without loading, fatigue, weathering, abrasion and chemical attack. Many researchers have proved that most of the qualities desired of concrete benefit by increased compressive crushing strength, such as strength in tension, shear, resistance to weathering, abrasion and wear and impermeability, these are classified as hardened properties of concrete. Exceptions to this rule are lightness and thermal insulation. The properties of fresh concrete though important but cannot be compared to the properties of hardened concrete which is retained throughout the entire life span of the concrete. These properties are however requiring of concrete structure for different purpose and occasions: Among the important strength test to be carried out in this study are the compressive strength test and the pull out test.

Compressive strength is the capacity of a material or structure to withstand loads tending to reduce size. In other words, compressive strength resists compression (being pushed together) by definition, the ultimate compressive strength of a material is that value of uniaxial compressive stress reached when the material fails completely. The compressive strength is usually obtained by means of a compressive test. Compressive strength is one of the most important engineering property of concrete which designers are concerned of. The test requisite differ country to country. The compressive strength of concrete are used to determine the concrete mixture and meet the strength requirement of a specific job. It is calculated from the failure load divided by the cross-sectioned area resisting the load.

$\sigma = F/A$ where F = Load applied (N), A = Area (m^2)

Generally speaking concrete used for residential purposes is around 17 MPa (MegaPascals), around 28 MPa for commercial uses, and as high as 70 MPa for other specified applications. The tests are required to determine the strength of concrete and therefore its suitability for the job. Pullout strength of hardened concrete is by measure the force required to pull an embedded metal insert and the attached concrete fragment from a concrete test specimen or structure (ASTM 900-01). The insert is either cast into the fresh concrete or installed in hardened concrete. Pullout tests are used to determine whether the in-place strength of concrete has reached a specified level. In addition, post- installed pullout tests may also be used to estimate the strength of concrete in existing constructions.

There are factors which have the greatest effect upon this strength of concrete, among them are the cement to aggregate ratio, the compaction, the water to cement ratio of the mix and the method of curing. Curing according to Chithra and Dhinakaran (2014) is the process in which the concrete is protected from loss of moisture and kept within a reasonable temperature range. It can also be defined as providing adequate moisture, temperature, and time to allow the Concrete to achieve the desired properties for its intended use. Curing plays an important role on strength development and durability of Concrete, as it serves as key player in mitigating cracks which brutally influence durability. Curing conditions according to Chithra and Dhinakarah (2014), as strength gain is significant in the curing phase. Inappropriate curing may result into carbonation at early stage and reduces PH levels. In a bid to obtain good and high quality concrete certain measures have to be introduced to ensure good quality of curing process. The objectives of this research is to investigate the effect of high curing temperature on compressive and pull out strength of concrete cubes. The concrete behaviour at elevated curing temperature is of concern in predicting the safety of building and construction in response to current global warning conditions.

Many researchers in the field based on the field based on the available Literatures, focused on curing method, and performance of concrete after curing. However, few studies examine the effect of elevated curing temperature on strength properties of concrete. Mundle (2014) observed that the compressive strength increase after 72 hours of exposure to an elevated temperature up to 150^{o}c and after that the compressive strength of a concrete decreases with increasing temperature. Investigation by Eliverly and Evans (1964) confirmed that concrete specimens mixed in normal temperature and cured under high temperature (40^{0}c) have higher compressive strength than those mixed and cured under normal temperature. Cebeci (1987) states that concrete cured in water within (37^{0}c), is of higher strength up to 90 curing days and lower strength (360 days) compared with concrete cured within (17^{0}c). Konstantin et al. (2000) determined in their research that the compressive strength at 30^{0}c increases with time much faster compared to 20^{0} curing.

Cecconello and Iutikian (2012) investigate the effect of high curing temperature on concrete up to 14 days. The results reveal that concrete casted and cured at lower temperature develop strength slowly but later improves after 14 days: and the strength of concrete casted and cured at higher temperature readily gain strength but later reverse after 14 days. According to Selman (2001) compressive strength of concrete mixed and cast at temperature not exceeding (29^{0}c) and moist cured under hot weather for 7days, increases as the curing temperature is increased up to 90 days. The increase ranges between (4-22%) with respect to mixes cured at normal weather condition. Lo et al. (2009) Investigates effect of curing at higher temperature on compressive strength and carbonation depth of Pozzolaure structural lightweight concretes and pulverized Fuel Ash- and Silica Fumes incorporated lightweight. Concrete with PFA and SF as a cement substitute up to a percentage of 70 were compared. Results show PFA and SF incorporated concrete had greater strength under accelerated curing condition than under normal curing condition, but OPC mixes showed a different result that they gain higher strength under normal curing and lower strength under accelerated curing when compared with the former.

2. MATERIALS AND EXPERIMENTAL METHODS

2.1.Cement

Dangote ordinary Portland cement (OPC) was used throughout the experimental study. The cement complied with BS 12 and ASTM C150- standard specification for Portland cement.

2.2 .Aggregates

The importance of using the right type and quality aggregates cannot be overemphasized. The fine and coarse aggregates cannot be over emphasized. The fine and coarse aggregates generally occupy 60-75% of the concrete volume (70-85% by mass) and simply influence the concrete's freshly mixed and hardened properties, mixture proportion and economy. The coarse aggregate used was crushed granite chippings obtained from RatCon quarry site along Lagos/Ibadan expressway with size not greater than 20mm. The fine aggregate used was natural sand and not greater than 4.75mm.

2.3.Specimen's Preparation

The batching of the specimen was by weighing the constituent materials according to the adopted mix ration of 1:2:4. This ratio was adopted throughout with 0.55 water cement ratio. The mixing of the concrete specimen was done mechanically after which the slump and compacting sector test were carried out before casting into cube molds of 150mm x 150mm x 150mm in approximately 50mm layers with each layer given 35 strokes of the tamping rod. Specimen were held in the laboratory at an atmosphere of more than 95% relative humidity and ambient temperature, i.e. 25^{o}c before the cubes were demolded after 24 hours and the specimens oven cured at 30^{o}c, 60^{o}c and 100^{o}c for 7days, 14days and 28days.

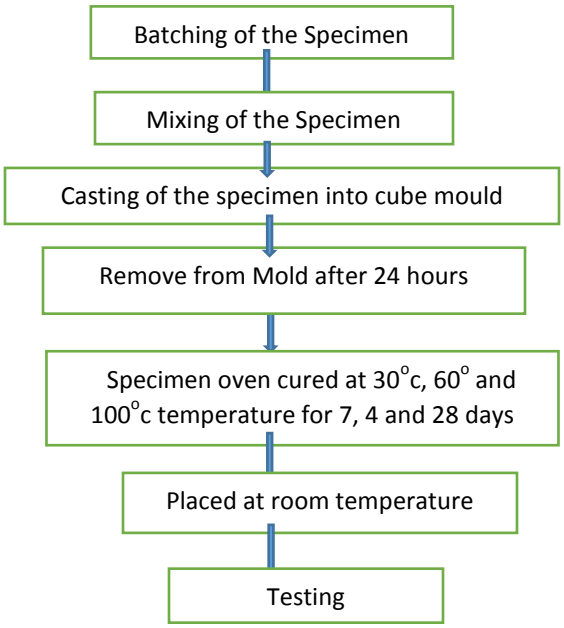
2.3.1.Test on the Concrete Samples

Tests on fresh concrete samples obtained as follows

Slump Test	75.0mm
Water cement ratio	0.55%
Mix proportion	M2O
Mode of mixing	Mechanical
Method of curing	Immersion in water

2.3.2.Flow Chart for Lab Work

Figure 1: Flow Chart for Lab Work



2.4.Testing Procedures

2.4.1.Compressive Strength Test

For the compressive strength test, a set of three standard cubes of 150mm x 150mm x 150mm size were texted using Electrically operated hydraulic compression machine with capacity of about 300KN. These were dance at different temperature 30^{o}c, 60^{o}c and 100^{o}c and at 7 days, 14days and 28days of curing. Compressive strength of any material is defined as the resistances to failure under the action of compressive force especially for concrete, compressive strength is an important parameter to determine the performance of the material during severe conditions. It is calculated by dividing the failure load with the area of application load.

$$\text{Compressive strength } (\upsilon) = \frac{\text{Failure load (KN)}}{\text{Area of application of load (mm}^2\text{)}}$$

2.4.2.Pullout Test

Pullout tests are used to determine whether the in-place strength of concrete has reached a specified level so that, other activities can proceed e.g. post-tensioning, termination of winter protection and curing. In addition, the strength of concrete is existing constructions may also be estimated using post-installed pullout tests.

The materials and tools used are:

(a) Pullout testing machine
(b) Concrete cone embedded with specially shaped steel rod.

The specimen is placed in the machine with the threaded shape from the embedded head attached to the machine. This is tie down by means of bearing plate& bearing ring after which the tension force is applied through the threaded shaft on the specimen. The reading on the calibration on the machine at the point threaded shaft detach itself from the specimen divided by the idealized area of frustum or area of specimen gives the pullout strength of that specimen.

3. RESULTS AND DISCUSSION

It is clear that the compressive strength pullout strength increases with curing days and at elevated curing temperature up to 100^0c. The result is better shown in figure 1. The line graph shows upward strength in different days at different temperature.

Table 1: Average compressive strength and pullout test result.

Days	Temperature			Pullout
	30^oc	60^oc	100^oc	
7 days	12.99	11.87	15.15	31.7
14 days	13.22	15.3	18.66	41.82
28 days	16.64	20.09	24	41.97

Figure 1: Line Graph Of Relationship Between Strength Of Concrete And Number Of Days At Different Temperatur

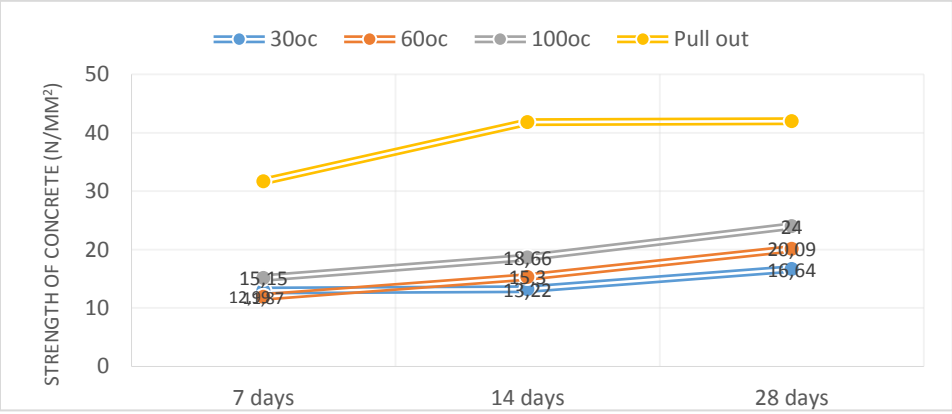


Table 2: Two-Way Analysis of Variance (ANOVA)

Source of Variation	SS	df	MS	F	P-value	F crit
Temperature	30.34	2	15.17	6.30	0.058	6.944
Days	71.45	2	35.72	14.83	0.014	6.944
Error	9.63	4	2.41			
Total	111.42	8				

Table 2 shows that Temperature F- value is 6.30 with the degree of freedom 2 and 4, and the p-value 0.058 which is higher than 0.05 (level of significant); hence, strength of concrete across different temperature are not significantly different. For Number of days, the F- value is 14.83 with the degree of freedom 2 and 4, and the p-value 0.014 which is less than 0.05 (level of significant); hence, strength of concrete across different Number of days are significantly different.

Table 3: Correlation Table between Number of Days and Strength of Concrete (Pull out)

Variable	Mean	Std Dev	Nr	COD	(%)	P-value
Days	16.33	9.56	6	**0.779**	**60.75**	**0.004**
Strength (N/mm^2)	1.647	0.256				

The correlation coefficient 0.779 shows that there is a fairly strong positive relationship between Number of Days and Strength of Concrete and the p- value 0.004 implies that the relationship is significant at 5% level while the coefficient of Determination (COD) indicated that Number of Days has 60.75% effect on strength of concrete.

Table 4: Regression Analysis of the Relationship between Strength of Concrete and Temperature

Coefficients	Standard Error	T Stat	P- value
Intercept	13.10	0.313	0.015
Temp	0.06	0.005	0.048

Table 4 above shows the regression analysis result which depicts that the **Strength of Concrete = 13.10 + 0.062 Temperature.** The regression equation shows that the strength of concrete will be **13.10 N/mm^2 at 0^{o}C** and a unit increase in temperature will lead to **0.062 N/mm^2** increase (impact) in the strength of concrete.

4. CONCLUSIONS

Based on the results and discussion presented above the following conclusions can be drawn

- At 7 days, 14 days and 28 days the strength of concretes cured at 100^{o}c are more greater than cured at 30^{o}c and 60^{o}c. Also the concrete cured at 60^{o}c is greater in strength than concrete at 30^{o}c.
- The increased average strength of concrete with increase in temperature is lower than the increase average strength of concrete with increased in number of days (see Table 1). Therefore, it shows tendency that there will be weak concrete in future with increased temperature.
- From Table 2, the p- value show the strength of concrete across different temperatures are not significant to that of strength of concrete across different number of days
- The regression analysis slowed a unit increases in temperature will lead to 0.062 N/mm^2 increases in the strength of concrete.

REFERENCES

Chithra S. and Dhinakaran G. 2014, "Effect of hot water Curing and hot air oven curing on admixed concrete", International Journal of ChemTech Research CODEN (USA); Vol.6, NO.2, PP 1516-1523.

Cebeci, O.Z., 1987, "Strength of concrete in Warm and Dry Environment", Materials and Structures, Vol.20, pp. 270-222.

Eliverly, R.H. and Evans, E.P. 1964, "The effect of curing condition on physical properties of concrete", Magazine of Concrete Research, Vol.16, NO.46, pp. 11-20.

Konstantin, K, Isaak, S, Shin, I, and Aron, B, 2000, "Influence of Mixture Proportions and curing conditions on compressive strength of High-performance concrete". ACI Materials Journal, Vol.97, NO.1, pp.25.

Selman, M.H., 2001, "Effect of Hot Weather on the Mechanical Properties of Concrete Produced using local Furnace slag". M.Sc. Thesis, Al-Mustansiriyah University, Baghdad, Iraq, pp.29,43.

Klieger, P.H., 1958, "Effect of Mixing and Curing on Concrete Strength", ACI Journal, Proceeding, Vol.54, NO.12, pp.1063-1081.

Lo. TY, Nadeem A, Tang WCP, Yu PC, 2009, "The effect of high temperature curing on the strength and carbonation of pozzolanic structural lightweight concretes, Construction and Building Materials", 23: 1306-1310.

Cecconello, V and Tutikkian, B 2010 The Influence of low Temperature on the Evolution of Concrete Strength Journal of Rev. IBRACON Estrut. Mater Vol. 5 NO.1, pp. 1-15

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

INNOVATIONS IN THE BIODIESEL PRODUCTION

DOI: 10.17261/Pressacademia.2017.566
PAP-WCTIE-V.5-2017(4)--p.24-28

Emine Emel Cakirca[1], Ayse Nilgün Akin[1]
[1] Kocaeli University, Engineering Faculty. emel.cakirca@kocaeli.edu.tr

ABSTRACT

Demand on energy from alternative fuels are growing rapidly due to depletion of fossil fuels and global warming crises. Biodiesel instead of conventional petroleum diesel, is a non toxic, biodegradable and renewable fuels. Therefore, biodiesel can be considered as a promising liquid fuel for transport sector. Among the routes of biodiesel synthesis, the most commonly used are transesterification of oil feedstocks and esterification of free fatty acids, in which these are carried forward in the presence of a catalyst. In transesterification process, a large majority of feedstocks come from vegetable/edible oil, as the first generation biodiesel feedstocks, in many regions of the world. The use of edible oils is not convenient in production of biodiesel due to insufficient amount of edible oil feedstocks and demand for food materials in the world. When it is thought that over 70% of price of biodiesel includes cost of feedstocks, use of cheaper and more sustainable resources have become inevitable. Waste cooking oil and algae oil, which have not been used as food sources, are termed as second generation feedstocks and seem more attractive and promising for biodiesel production. In addition to raw material source, the type of alcohol and catalyst also affects to price and quailty of biodiesel. Since biodiesel is a process includes sustainable technology, all inputs are investigated in terms of usability, sustainability and eco-friendly. The aim of this study to give an overwive on the biodiesel production in that select of the most effective and available feedstocks and summarize the recent development and innovation in the production process.

Keywords: Biodiesel, raw materials, oil, catalyst, innovation

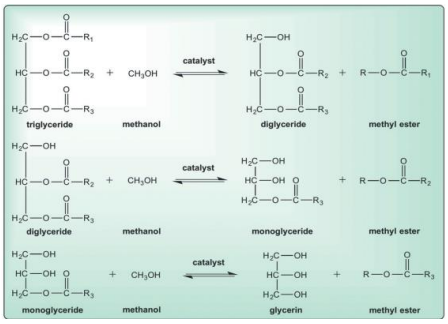
1. INTRODUCTION

In recent decades, increasing human population and industrialization around the world led to depletion of fossil fuels and emergence of energy crisis. Furthermore scarcity of fossil fuels reserves and occurrence greenhouse gases through the use of fossil fuels have made alternative fuel resources more attractive for the future. Among the renewable energy resources, biodiesel is the most common alternative energy instead of conventional diesel. Biodiesel has gained attention recently due to its biodegradability, lack of sulphur, non toxisity, renewability and positive properties compared to petroleum diesel (Avhad and Marchtti, 2016; Tabatabaei et al., 2015).

It is derived from transesterification and esterification of edible or non edible oil and animal fats with short chain alcohols such as methanol and ethanol in presence of a catalyst. As shown in Figure 1, transesterification is a multi-step reaction; Triglycerides are converted to diglycerides, diglycerides are converted to monoglycerides and then monoglycerides are converted into biodiesel and glycerine as by-product. There are three important inputs in this reaction. The oil source that accounts for majority of biodiesel cost; the second is the type of catalyst that accelerates the reaction and improves biodiesel efficiency, and finally the alcohol sources.

The most widely used type of catalyst in the commercial production of biodiesel is homogeneous basic catalysts, KOH and NaOH. Methanol is used as a reactant because of its cheap and easy availability. A wide variety of edible cooking oils are used as an oil source such as sunflower oil and canola oil for commercial production of biodiesel. In this study, recent studies and innovations in biodiesel production have been reviewed and summarized through the raw material sources.

Figure 1: General Equation of Transesterification (Lima et al., 2016)



2. RAW MATERIALS FOR BIODIESEL PRODUCTION

Transesterification process is one of the most widely used method in the commercial production of biodiesel. This reaction takes place between triglycerides with alcohols having 1-8 carbon atoms in the presence of a catalyst. The raw materials for biodiesel production are triglycerides and short chain alcohols. In addition to them, catalysts also play an important role in transesterification reaction. It can be concluded that there are three major materials which affect on biodiesel production in that its quality and cost. Among these materials, oil feedstock is the most important one since it plays decisive role in terms of cost of biodiesel.

2.1. Oil Feedstocks

Currently, majority of the biodiesel production is from edible oil (more than 95%) (Ahmia et al., 2014). Table 1 shows classification of sources of biodiesel. In generally, oil sources are categorized in three groups: edible oil as 1st generation, non edible oil as second generation and other sources.

Table 1: Classification of the Raw Materials for Biodiesel Production (Verma et al., 2016)

Edible oil (1st generation)	Non edible oil (2nd generation)	Other sources (3rd generation)
Soybean oil	Jatropha oil	Waste cooking oil
Rapeseed oil	Jojoba oil	Fish oil
Sunflower oil	Rubber seed oil	Microalgae
Palm oil	Milk bush oil	Animal fats
Canola oil	Linseed oil	Chicken fat oil
Corn oil	Neem oil	
Castor oil	Tall oil	

Leading countries around the world for biodiesel production are USA, Brazil, Germany, France and Argentina. In the US, soybean oil is still the largest biodiesel raw material. In Europe, canola oil is primary raw material in 2015. Rapeseed in Canada and palm oil in Southeast Asia are also used in biodisel production (Falch et al., 2016; Ahmia et al., 2014). The choise of raw materials for biodiesel depends on some parameters: (1) avail-ability of land, (2) cultivation practices, (3) energy supply and balance, (4) greenhouse gases, (5) logistics costs, (6) soil erosion and fertility, (7) direct economic value of the feedstocks (Bhuiya et al., 2016). Since rapeseed produced so much as to export in Europe, is commonly used in Europe in biodiesel production. Likewise, US and the other countries also use edible oils in the production of biodiesel, which are available for production on their own soil (Verma et al., 2016). However, usage of edible oil for production biodiesel competes with the use of land for food production and lead to increasing price of oil and causes food scarcity all over the world (Mardhiah et al., 2017; Chehetri et al, 2008, Verma et al., 2016). Non edible oil sources needs to be provided in biodiesel production for economical and environmental benefits. New generation biodiesel also known 2nd generation, produced from non edible oil such as Jatropha oil, microalgae oil or waste cooking oil. Utilization of renewable feedstocks provide sustainability for biodiesel and reduce the cost of biodiesel production by 60 – 90% (Baskar et al., 2016). Although microalgae and waste cooking oil are seen other sources as 3rd generation from Table 1, these sources are also regarded as the 2nd generation biodiesel feedstocks in literature (Bhuiya et al., 2016).

Microalgae as new generation raw materials for biodiesel is an important alternative due to be able to grow in any environmental conditions and its high oil yields. Microalgae can play an important role in solving the problem between the production of food and biofuels in the near future. The advantages of microalgal biodiesel are : (1) contains no sulpur, (2) non toxic, (3) highly bio-degradable, (4) reduce carbon emissions (Demirbaş et al., 2011). Petvoy et al. (2012) reported that despite of the fact that algal biomass converted into their esters, algal biomass is much more expensive to produce

biodiesel. So, researchs on the algal biodiesel focused on the economic viability of the algal biomass. From literature, it is seen that algal biodiesel will be one of the most important alternative fuels as long as it is economically healing.

One of the other promising raw material is waste cooking oil which has high potential for biodiesel production. Research shows that about 380,000 tonnes of waste cooking oil are produced in the US and 700,000 - 1,000,000 tonnes / year in Europe. Approximately 1.7 million tons of vegetable oil is consumed annually in Turkey, and a serious waste oil is generated (http://www.cygm.gov.tr). By using these oils as biodiesel resources, undesirable contaminant wastes are removed in an environmentally friendly way, while at the same time the dependence on edible resources in biofuel production will be reduced (Maneerung et al., 2016; Mohammadshirazi et al., 2014). Al-Hamamre et al. (2014) reported that the use of edible oils in biodiesel production affects approximately 70-95% of the cost of biodiesel. Therefore, minimizing the cost of raw material is a very important advantage for cheap and effective fuel production.

2.2. Alcohol Feedstocks

Transesterification is a reversible reaction and contains stoichiometrically 3: 1 alcohol to oil. However, excess alcohol is used in the reaction since using of excess alcohol pushes the reaction towards the products. Methanol as a cheap alcohol source is widely used because of its physical and chemical properties suitable for transesterification. But methanol has a low boiling point and the risk of explosion is very high. Methanol and methoxide are very dangerous substances that must be handled with care (Stamenkovic et al., 2011). Because of this, ethanol can be seen as an alternative to methanol due to its easy production and availability (Anastopoulos et al., 2013). The use of ethanol instead of methanol in the transesterification reaction is a promising step because it can be produced from agricultural renewable sources and has less toxicity than methanol.

2.3. Catalysts For Biodiesel Production

Transesterification is carried out by catalytically in order to increase reaction rate. Homogeneous catalysts such as NaOH, KOH, H2SO4 are applied in conventional commercial methods. These type of catalysts have high yield in a short time in transesterification reaction. However, the steps such as washing and purification in homogeneous process increase the production cost. In addition, by using homogeneous catalysts it is seen that presence of K / Na traces in the biodiesel and reduce product quality (Semwal et al., 2011). The main disadvantage of homogeneous catalysts is generation of large scale waste water from the washing process. For these reasons, research have been based on new technologies in terms of enviromentally, economically and reducing problems just mentioned.

According to free fatty acid contents of oil, asidic or basic catalysts are used. The amount of free fatty acids content in the oil is the determining parameter for the use of acidic or basic catalyst. When the free fatty acid content is too high, the use of basic catalyst causes the formation of soap, which is undesirable (Endalew et al., 2011). On the contrary, it is suitable to use an acidic catalyst in this case. But the reaction rate is quite slow.

One of the recent developments in the production of biodiesel is the design of heterogeneous catalysts that are more active, produced from alternative cheaper feedstocks and to be reusable.

The most important advantage of heterogeneous catalysts is lifetime of the catalysts which is longer than homogeneous catalysts in biodiesel production. This feature is related to stability of the crystal structure which of the catalytic surface. The poisoning and leaching of the catalyst during the reaction led to change the physical properties as well as deactivation of the catalyst. For this reason the catalyst must be properly formulated (Yan et al., 2010).

Table 2. shows that effect of the different type of oil, catalyst and alcohol on the biodiesel production. As it seen, heterogeneous catalysts exhibit competitive activity with homogeneous catalysts. CaO-based catalysts are widely used because of their long lifetime, exhibit high activity, and operate under moderate reaction conditions. CaO has also attracted attention due to obtained from waste, eggshell and mollusk shells in natural and inexpensive ways (Borgez et al., 2012).

Table 2: Summary of Raw Materials Used in Biodiesel Production in the Literature

Process	Catalysts	Type of oil	Type of alcohol	Yield, %	References
Homogeneous	KOH	Freshwater algae	Ethanol	-	Vijayaraghavan et al., 2009
Homogeneous	H2SO4	Microalgae	Methanol	-	Miao et al., 2006
Homogeneous	NaOH, KOH	Frying oil	Ethanol	74	Encinar et al., 2007
Heterogeneous	MgFe2O4@CaO	Soybean oil	Methanol	89	Liu et al., 2016
Heterogeneous	CaMgO	Rapeseed oil	Methanol	85	Zakaria et al.,

					2012
Heterogeneous	Mg–Al hydrotalcite	poultry fat	Methanol	98	Liu et et al., 2007
Heterogeneous	Lipaz	Palm oil	Ethanol	82	Raita et al., 2010
Heterogeneous	Li-CaO	Jatropha curcas oil	Methanol	100	Endalew et al., 2011
Heterogeneous	Dolomite	Canola oil	Methanol	92	İlgen, 2011
Heterogeneous	Ca/Al/Fe3O4	Rapeseed oil	Methanol	98	Tang et al, 2012
Heterogeneous	rice husk	Many kinds of oils	Methanol	94	Roschat et al., 2016

3. CONCLUSION

Biodiesel is produced under different conditions by using a wide variety of oil feedstocks. But physical properties, environmental and economic efficiency of the oil are important effects on biodiesel production. The importance of biodiesel raw materials is inevitable especially considering the environmental and economic consequences. For this reason microalgae oil and waste oil are promising as future generation raw materials for biodiesel production. The current studies are about increasing the oil yield of microalgae and making them more commercially viable. Using waste oils in the production of biodiesel draws attention due to the fact that wastes are disposed as an environmentally friendly way and dependence on edible oils is reduced considerably as well. Alcohol, another raw material in biodiesel production, is based on methanol and ethanol in the literature. Ethanol has advantages over methanol because it can be biologically produced. But the ethanolysis reaction has not yet been fully investigated. Heterogeneous catalysts used in transesterification attract attention because of their being environment friendly, easy to find, reusable and cheap.

REFERENCES

Ahmia, A.C., Danane, F., Bessah, R., Boumesbah, I. 2014, "Raw material for biodiesel production", vol. 17, no 2, pp. 335-343.

Al-Hamamre, Z., Yamin, J. 2014, "Parametric study of the alkali catalyzed transesterification of wast frying oil for Biodiesel production", Energy Conversion and Management, vol. 79, pp. 246-254.

Anastopoulos, G., Dodos, G.S., Kalligeros, S., Zannikos, F. 2013, "Biodiesel Production by Ethanolysis of Various Vegetable Oils Using Calcium Ethoxide as a Solid Base Catalyst", International Journal of Green Energy, vol. 10, pp. 468–481.

Avhad, M.R. and Marchtti, J.M. 2016, "Innovation in solid heterogeneous catalysis for the generation of economically viable and ecofriendly biodiesel: A review", vol. 58, no 2, pp. 157-208.

Baskar, G., Aiswarya, R. (2016), "Trends in catalytic production of biodiesel from various feedstocks", Renewable and Sustainable Energy Reviews, vol. 57, pp. 496–504.

Bhuiya, M.M.K., Rasul, M.G., Khan, M.M.K., Ashwath N., Azad, A.K. (2016), "Prospects of 2nd generation biodiesel as a sustainable fuel—Part: 1 selection of feedstocks, oil extraction techniques and conversion Technologies", Renewable and Sustainable Energy Reviews, vol. 55, pp. 1109–1128.

Chhetri, A.B., Watts, K.C., Islam, M.R. 2008, "Waste Cooking Oil as an Alternate Feedstock for Biodiesel Production", Energies, vol. 1, pp. 3-18, DOI: 10.3390/en1010003.

Demirbaş, A., Demirbaş, M. F. 2011, "Importance of algae oil as a source of biodiesel", Energy Conversion and Management, vol. 52, pp. 163–170.

Encinar, J.M., González, J.F., Rodríguez-Reinares, A. 2007, "Ethanolysis of used frying oil. Biodiesel preparation and characterization", Fuel Processing Technology, vol. 88, pp. 513–522.

Endalew, A.K., Kiros, Y., Zanzi, R. 2011, "Heterogeneous catalysis for biodiesel production from Jatropha curcas oil (JCO)" Energy , vol. 36, pp. 2693-2700.

Flach, B., Lieberz, S., Rondon, M., Williams, B., Wilson, C. 2016, "EU Biofuels Annual 2016 ", pp. 22-29.

İlgen, O. 2011, "Dolomite as a heterogeneous catalyst for transesterification of canola oil", Fuel Processing Technology, vol. 92, pp. 452–455.

Lima, A.L.; Ronconi, C.M., Mota, C.J.A. 2016, "Heterogeneous basic catalysts for biodiesel production", Catalysis Science and Technology, vol. 6, pp. 2877–2891.

Liu, Y., Zhang, P., Fan, M., Jiang, P. 2016, "Biodiesel production from soybean oil catalyzed by magnetic nanoparticle MgFe2O4@CaO", Fuel, vol. 164, pp. 314–321.

Maneerung, T., Kawi, S., Dai Y., Wang, C. 2016, "Sustainable biodiesel production via transesterification of waste cooking oil by using CaO catalysts prepared from chicken manure", Energy Conversion and Management, vol. 123, pp. 487–497.

Mardhiah, H. H., Ong, H.C., Masjuki, H.H., Lim, S., Lee, H.V. 2017, "A review on latest developments and future prospects of heterogeneous catalyst in biodiesel production from non-edible oils", Renewable and Sustainable Energy Reviews, vol.67, pp. 1225–1236.

Miao, M., Wu, Q. 2006, "Biodiesel production from heterotrophic microalgal oil", Bioresource Technology, vol. 97, pp. 841–846.

Mohammadshirazi, A., Akram, A., Rafiee, S., Kalhor, E.B. 2014, "Energy and cost analyses of biodiesel production from waste cooking oil", Renewable and Sustainable Energy Reviews, vol. 33, pp. 44–49.

Petkov, G., Ivanova, A., Iliev, I., Vaseva, İ. 2012, "A critical look at the microalgae biodiesel", Eur. J. Lipid Sci. Technol.,vol. 114, pp. 103–111.

Raita, M., Champreda, V., Laosiripojana, L. 2010, "Biocatalytic ethanolysis of palm oil for biodiesel production using microcrystalline lipase in tert-butanol system", Process Biochemistry, vol. 45, pp. 829–834.

Roschat, W., Siritanon, T., Yoosuk, B., Promarak, V. 2016, "Rice husk-derived sodium silicate as a highly efficient and low-cost basic heterogeneous catalyst for biodiesel production", Energy Conversion and Management, vol. 119, pp. 453–462.

Semwal, S., Arora, A.K., Badoni, R.P., Tuli, D.K. 2011,"Biodiesel production using heterogeneous catalysts", Bioresource Technology, vol. 102, pp. 2151–2161.

Stamenkovic, O. S., Velickovic, A.V., Veljkovic, V.B. 2011, "The Production Of Biodisel From Vegatable Oils by Ethanolysis: Current State and Perspectives", Fuel, vol. 90, pp. 3141-3155.

Tang, S., Wang, L., Zhang, Y., Li, S., Tian, S., Wang, B. 2012, "Study on preparation of Ca/Al/Fe 3 O 4 magnetic composite solid catalyst and its application in biodiesel transesterification", Fuel Processing Technology, vol. 95, pp. 84–89.

Tabatabaei, M., Karimi, K., Horváth, I. S., Kumar, R. 2015, "Recent trends in biodiesel production", Biofuel Research Journal, vol. 7, pp. 258-267.

Verma, P., Sharma, M.P., Dwivedi, G., 2016, "Impact of alcohol on biodiesel production and properties", vol.56, pp. 319-333.

Vijayaraghavan, K., Hemanathan, K. 2009, "Biodiesel Production from Freshwater Algae", Energy Fuels, vol. 23, pp. 5448–5453.

Yan, S., DiMaggio, C., Mohan, S., Kim, M., Salley, S.O., Simon Ng, K.Y. 2010, "Advancements in Heterogeneous Catalysis for Biodiesel Synthesis", Top Catal, vol. 53, pp. 721–736.

Zakaria, R., Harvey, A.P. 2012, "Direct production of biodiesel from rapeseed by reactive extraction/in situ transesterification", Fuel Processing Technology, vol. 102, pp. 53–60.

http://www.cygm.gov.tr/CYGM/Files/yayinlar/kitap/bitkisel_atik_yaglarin_yonetimi_kitapcigi.pdf

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

INNOVATION APPLICATIONS AT IGDAS, A NATURAL GAS DISTRIBUTION COMPANY OPERATING IN THE ENERGY SECTOR

DOI: 10.17261/Pressacademia.2017.567
PAP-WCTIE-V.5-2017(5)-p.29-33

Mehmet Akif Demirtas
Manager of Stategic Planning & Management Systems, İGDAS, İstanbul, Turkey

ABSTRACT

In this study, the past, present and the future of ınnovation process in IGDAS (Istanbul Gas Distribution Industry and Trade Incorporated Company), which has been operating for 30 years as the biggest company of the intercity natural gas distribution sector of Turkey, are mentioned. In this regard, it can be expressed that it is also mentioned in a few published studies that include experiences with regards to collecting suggestions from internal and external resources on job development in the service sector related to electric, water and gas distribution (utility sector), and solving some problems, evaluating them, giving feedback, their implementation and awarding stages. In the notice, it was aimed at publicizing the company with its main field of activity, development, and some awards received, and afterwards, its innovation journey starting from the 2000s was explained via a case study approach, step by step, detailed under each title.

Keywords: Innovation, energy, natural gas, IGDAS

1. INTRODUCTION

Istanbul Gas Distribution Industry and Trade Incorporated Company (İGDAŞ) was founded on December 25, 1986, with the clear intent of saving Istanbul from air pollution, to conduct the intercity natural gas distribution works. Infrastructure works oriented towards this goal were conducted by the SAE consortium formed by French Sofregaz and Turkish Alarko companies, and natural gas started to be used in 1992. The ratio of sulfur dioxide in the air at that period, 219 micrograms/m^3, was threatening the human health, however, it caught a downwards trend as natural gas usage became widespread, and the ratio reached 20 micrograms/m^3 in the beginning of 2000s, the limit level determined by the World Health Organization (WHO). As a result of the increased population in the city, energy consumption has increased in relation with the need for heating since the 1990s. Even though this was followed by the increasing number of vehicles and exhaust gas emissions, the above-mentioned ratio of sulfur dioxide has been kept at the single-digit level. This is a praiseworthy success. In a research conducted by Price Waterhouse Coopers (PwC) and an NGO called Partnership for New York City, in which the world's leading metropolises were evaluated under business life, daily life and innovation categories, Istanbul was successful to become one of the three metropolises with its air quality.

İGDAŞ, within the framework of the "sustainability" concept that it emphasizes both on its mission-vision and among its values, has been proceeding to become the flagship of the sector with its infrastructural investments and increasing number of subscribers, and has become a role model company in Turkey its sector with its maturing corporate model. As transparently shared from our website for years (Activity Report - 2015), more than 6 million BBS (number of independent units) subscribers are provided with 5.5 billion m3 of natural gas annually, via a natural gas line that is approximately 17.000 kilometers long, in a high quality, uninterrupted, and safe way. As a result of this, it has been among the top 20-30 companies of the biggest 500 companies of our country for the past 5 years. The concept of quality adopted by İGDAŞ has not only allowed the institution to ensure customer satisfaction, but also contributed to its becoming a structure that foresees continuous improvement and encourages employer motivation, and together with the understanding of quality it internalized with the participation of all its employees, the company adopted as a principle to evaluate the concepts of environment and work safety in a holistic manner.

It took the first step in its journey towards excellency by signing the goodwill declaration with KalDer (Turkey Quality Association) in 2007, with the aims of proceeding in accordance with its vision and strategic targets, focusing on the

fundamental requirements and needs of its shareholders, forming future-oriented scenarios by determining their short- and long-term priorities, and managing the risks that may occur accordingly. İGDAŞ revealed its success with regards to the applied management model especially in the evaluations conducted by KalDer based on the EFQM (European Foundation for Quality Management) Excellence Model, by receiving 4 stars in 2009 on Competence in Excellence, Success Award in 2010, and National Quality Grand Award in 2011 in the Large-Scale Businesses category. The process was crowned with the Sustainability Excellence Award received in 2015.

So as to add value for the shareholders, works are monitored with an approach that is suitable with the logic of strategic planning, in relation with strategies, key performance indicators (KPI), activities, and process performance indicators concepts, studies to comply with customer satisfaction and corporate social responsibility principles are supported with survey results. Thanks to the different practices in relation with the process monitoring methodology and innovation works, these successes were crowned with the recently received 2011 Best Call Center (IMI Conferences), Corporate Social Responsibility Sustainability, and 2013 Company Adopting Customer Satisfaction As A Principle (Republic of Turkey, Ministry of Customs and Trade) awards. Moreover, the corporate maturity of İGDAŞ was openly proven with the Non-Public Company with the Highest Score Award received from Turkish Corporate Management Association in 2013, based on the rating about corporate management policies.

At this point, as mentioned at BUYEM (2014), intrapreneurship is an entrepreneurship approach maintained by all human resources in an established system and a functioning structure, and it is fundamentally an in-house open innovation approach. In this regard, the studies whose details are explained below and which were conducted within the framework of the strategical aim of strengthening the knowledge and technology leadership (technology, know-how, design, innovation) of İGDAŞ by improving its Innovation and Design capacity, which is among its strategical aims, provided added value in achieving the above-mentioned successes.

2. IGDAS INNOVATION JOURNEY

2.1. Everyone's Opinion is Valuable

As expressed by Kırım (2012), when innovation is considered, what is meant is not always the radical, Silicon Valley-style innovations. Any company in any sector should have a command on various types of innovation that they can put into effect in different situations and succeed, and especially the conditions under which these types give the best solutions. Starting from a similar point of view, suggestions began to be collected in 2001 so as to solve some problems faced during the business processes. Main logic here is the assumption that the person who conducts his/her work in the best way is the one who knows about it the most, and that he/she is the first one to encounter the problems and search for solutions. The environment in which the verbal ideas would, so to speak, evaporate in the beginning turned into a process in which written suggestions were collected, evaluated by a commission, and improvements, although gradually, started to be made. In this 3-year period, coordination was ensured by an administrative unit, and as the main area of activity of the company is natural gas distribution and as the received suggestions dominantly have a technical nature due to the technical background of most of the employees, the business was required to be conducted under the responsibility of a technical unit.

2.2. Suggestion Evaluation

From Adair's (2008) point of view; while innovation is organizationally a natural human activity, it should be as voluntary and planned as possible. If you fail to plan, you plan to fail. Based on the internationally accepted reality, one can suggest that an organization that avoids to face change or does not need innovation is doomed to stagnate, decay and, finally, die. Just as the trees start to die from the top, this saddening process in the companies starts from the top management and spreads towards those in the sub-levels. Managerial leaders are needed for a beneficial change program. Change swiftly increases the need for leaders. Leaders are those that will ensure change. Even though innovation is a natural process, it becomes more effective when led appropriately. This means that change shall be welcomed, planned, controlled, monitored, and most importantly, led in accordance with the aims of the organization. As stated in section 2.1., some technical problems could be solved thanks to the approach of a technical department that speaks the same language, and, when a possible resistance towards change is considered, the fact that the top management spread the change towards the employees, in the meaning of mentality, starting from itself, and made it systematic. In the period between 2004 and 2006, the number of the suggestion evaluation meetings increased, and the foundations of an archive, although manually, were laid.

2.3. Suggestion Follow-up System

Drucker (2003), who has various studies on innovation, defines the concept of innovation management as "the activities companies perform so as to conduct innovation under control within a process by activating both the unorganized and complex structure and their managerial activities in order to comply with the internal and external changes". When used by

itself, the concept of innovation may express the process, however, the concept of innovation management expresses the controlling of innovation by means of managerial activities during implementation. Since 2006, advancement was contributed when a unit that is in close relation with the company plans, has both technical and administrative features when it comes to know-how, and that reports to the company's top management. The support of the top management allowed the problem of encountering some negativity during the implementation stage of the accepted suggestions to be solved faster.

2.3.1. Solution at Work Practice

Even though quality was increased thanks to the monthly suggestion-specific evaluations conducted by specialist assessors, if the desired number of suggestions were not reached, the following method may be applied: All employees may be asked via a letter written directly to the personnel by the top management (general manager of CEO) for a certain period of time (30-45 days, for instance) whether they have suggestions with regards to their own job. Thomas et al. (2000) expresses in a study that the objectives of the innovation process in companies may be listed as follows:

- Ensuring structural attitudes for systematic and continuous innovations,
- Defining environmental and legal possibilities and determining the ideas that may lead innovation,
- Determining financial and strategic objectives in order for the innovation effectiveness of the company.

With the Solution at Work project implemented in 2008, the results expressed in articles 1 and 2, namely the annual number of collected suggestions, were achieved in a short period of time - 1 to 2 months. At this point, the personnel desired not only to express the suggestion in black and white, but also verbally, in front of the team that evaluates it, and this was implemented.

2.3.2. Digital Infrastructure Installation

When Özdaşlı's (2006) approach is considered, innovation process in organizations should not be left to coincidences, but be planned and foreseen. In order for innovation to be managed, it is required to determine and plan objectives that are really related to innovation, and the related plans to be implemented. So as to manage innovation in a planned way, it shall be realized that innovation is a process, and the stages of creating and obtaining new ideas in this process, as well as the stage of implementing these ideas shall be organized well. An innovative movement systematically implemented for companies may be effective. In this regard, especially within the scope of the EFQM excellency model works mentioned in the introduction part, companies experienced on innovation were visited on-site so as to learn and compare, and as a result, it was decided that a system shall be established, which is able to follow-up on the system from end to end, whose evaluation criteria were determined, feedback logic is settled, and which will also support the corporate memory. Therefore, a more professional process management is aimed at. Thanks to an easily usable-widespread module that is compatible with the digital system used within the company, the suggestions entered into the system are, if necessary, are pre-screened based on the announced criteria (compatibility with the main field of activity, added value to the company activities, originality, etc.). As a result of the evaluation meeting, the correspondence with the relevant units, award process, etc. information is recorded. Therefore, such a software support both supports the formation of innovative ideas and enables the managing of ideas with a portfolio logic.

2.3.3. Innovation Volunteers' Team

In the Uzkurt's (2008) work, Morgan argues that an innovative organizational culture can only be developed by means of a cultural change and transformation. O'Reilly, on the other hand, lists the necessary activities for such a change and transformation as follows:

- Ensuring employee participation in the process,
- Developing a management understanding that supports and motives change,
- Evaluating the internal and external information, and
- Establishing a comprehensive award system for the change-related activities.

Especially with regards to the issues mentioned in the first two articles, an announcement was made to all units and all stages, and an "Innovation Volunteers' Team" was formed with representatives from all units - from operators to managers. In order for the team to speak the same language, training was received about the subject, and suggestion evaluation works started to be conducted more systematically by means of this team.

On top of that, as the information and culture level on the subject increased, the team organized works with special groups formed within themselves as with a project logic, on specific issues determined by the top management, and reported the results to them. Therefore, steps were taken not only towards gradual, but radical innovation.

2.4. Innovatıon Management System

Starting from the phenomenon of "change is the only constant", the established systems shall be improved in time, and be made more answerable to developing needs. In accordance with the demands of the employees that supported the innovation system with their suggestions and the ideas developed by the innovation volunteers, a new digital infrastructure was established in 2012, which is much more developed when compared to its 2008 version, through which the owner is able to follow up on his/her suggestion at any stage, by which repetitive/similar suggestions are detected and grouped more easily, and which has a pool that can be re-evaluated when the conditions change, even when the given suggestions were not accepted due to various reasons. In this period, the team of assessors that change based on the idea continued to work in the meetings led by the related assistant general manager, in which the owner is personally present to offer his/her suggestion. Moreover, the increase in the number of suggestions, namely quantity, was important in the previous periods, while today, the value added to the company with the suggestion, namely quality, has gained importance.

2.4.1 Open Innovation

Except for a standardized website allowing suggestions to be received externally, designing a solely innovation-oriented website is a practice that was again supported with the idea created by the innovation volunteers, and received acceptance of the top management. Therefore, an innovation communication channel was established, which is open to suppliers, subscribers, industrial users, and the world. Furthermore, a visual design called "inovig" was made for publicity purposes, and used in the website and other activities. The company's innovation website was awarded the Golden Ant Award in the Communication Management Projects category in 2012, given by Marmara Municipalities Union.

2.4.2. Innovation Day Event

As expressed in the EFQM Excellence Model, employees are appreciated, known and supervised. Excellent institutions encourage their employees to participate in the improvement and entrepreneurship process, and recognize their endeavors and successes. In this regard, the suggestion-owners who deserved to be awarded are done so according to the procedures and principles of the company on "giving awards and gifts to the personnel", by means of the periodical ratings conducted from the beginning. The issue is announced to the other personnel via appropriate communication channels. While these events can be conducted in-house, the award ceremony can be organized in an environment such as a university, where innovative products are displayed, narrated to the visitors by means of posters and LCD TV screens, and in which applications are performed via which the participators can have interactive roles, by learning especially from the events organized by international companies on this issue. This situation is also expressed as "people's contributions to the company are realized and valued" at the IiP (Investors in People) - the award our company deserved to receive in 2015 - standards, which are followed by few companies in our country.

2.4.3. Spreading with Education in Innovation

In 2014, together with the changes in the innovation volunteers' team, a comprehensive and professional training oriented at the "Conjoining with Innovation Coaching Bases and Company Innovation Process" for a more systematical spread of innovation culture within the company was received in collaboration with a rooted university, which enabled the team to renew itself and maturity level to increase with regards to the style in approaching the issue.

2.4.4. Implementing and Monitoring the Suggestions

As expressed in the previous sections, the suggestions which were decided to be implemented via the Innovation Management System Module is implemented by the related unit, its results are recorded in various dimensions, and monitored accordingly. In addition, the implementations can also be monitored within the framework of Strategic Planning, which has been efficiently implemented in the company for long years. The critical success factors of the strategic objectives included in the 5-year Strategic Plans are detected, and they are measured by means of key performance indicators (KPI). The activities that are annually updated so as to ensure these are monitored by means of performance target indicator and value during the review meeting of the management organized at each quarter of every year, and reported to the top management. Therefore, the appropriate suggestions originating from the Innovation Management System can naturally be evaluated within the scope of the above-mentioned strategic activity.

3. FINDINGS

Evaluating the results of a conducted study/activity, receiving feedback via methods such as survey, and making improvements on the process, if necessary, are very good opportunities. Within the scope of the periodical Employee Satisfaction Surveys made with this purpose, questions on innovation activities were prepared, and how the issue is perceived by internal customers, namely the employees, began to be monitored. The results were entered into the process improvement module, and the related units were ensured to conduct necessary process improvements. Beneficial findings can be obtained via methods such as audit, evaluation or investigation. A professional off-site eye can evaluate you on different dimensions and compare you with your counterparts, and you may benefit from it. With this aim, the İGDAŞ Innovation Management System that was tried to be explained above in general terms was desired to be evaluated with a non-partial and systematic point of view. In 2014, İnovaLİG was participated in, which is an innovation development program conducted by Turkish Exporters' Assembly (TİM) with the support of A.T. Kearney. During the first implementation of the competition, which is practiced in 17 European countries and based on the IMP^{3}rove methodology, in our country, the standards that were developed in Europe in years were complied with in a sensitive way. Among the 460 firms registered for the competition, our company ranked among the first 250 companies thanks to its data quality and the attention paid. When these are evaluated according to the A.T.Kearney "House of Innovation" methodology, the average points received by these companies were surpassed in the fields of innovation strategy, innovation organization and culture, and enabling factors. In this group, in which there are companies that keep innovation activities at the center of their main working field, have years of experience, and are well-known in their sectors, our company especially took justified pride of being among the top 5% group in the "Enabling Factors" category.

4. CONCLUSION

The more a work, either in the social life or business life, is valued, the more success or beneficial result is obtained from it. The fact that the structures established in the companies are well-founded, systematic and integrated with other practices is among the essential factors of sustainability. The findings reported by an external observer, which are especially expressed in the third section, provide us with a good idea on the functionality of the established system. Therefore, it can be asserted that an innovation role model that can be implemented in the companies which operate in the utility sector is put forth. In this notice that was summarized above, an "Innovation Management System Model" implemented in a large-scaled company operating in the natural gas distribution sector of the energy industry was shared with case studies.

REFERENCES

Adair, J. (2008). Yenilikçi Liderlik, (1st Edition), Istanbul

Büyem, 2014, "Boğaziçi University Lifelong Learning Center: Conjoining with Innovation Coaching Bases and Company Innovation Process Training Notes"

Drucker, P. F. , "The Discipline of Innovation", Trans. Ahmet Kardam, Mess Yayınları, Istanbul, 2003, p. 120

İgdaş , 2015, "Annual Activity Report – 2015 "

Kırım, A. (2012). Mor İneğin Akıllısı, (44th Edition), Istanbul

Kuczmarski , T., Middlebrooks, A., Swaddling , J., Innovating the Corporation: Creating Value for Customers and Shareholders, NTC Business Books/Contemporary Publishing Group, 2000, p.141

Özdaşlı, K., "Toplam Kalite Yönetimi ve Yenilik İlişkisi: Bir Örnek Olay", AKADEMİK BAKIŞ Uluslararası Hakemli Sosyal Bilimler E-Dergisi, Issue: 10, September 2006, p.14

Uzkurt, C.; Yenilik Yönetimi ve Yenilikçi Örgüt Kültürü, Beta Yayınları, Istanbul, 2008

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

DETERMINING SHIPBOARD INTEGRATION REQUIREMENTS of MAINTENANCE 4.0 CONCEPT in MARINE ENGINEERING

DOI: 10.17261/Pressacademia.2017.568
PAP-WCTIE-V.5-2017(6)-p.34-40

Cagatay Kandemir[1], Metin Celik[2]
[1]Department of Marine Engineering, Istanbul Technical University. ckandemir@itu.edu.tr
[2]Department of Marine Engineering, Istanbul Technical University. celikmet@itu.edu.tr

ABSTRACT

Ship machinery maintenance is a core technical aspect to achieve expected reliability, availability and efficiency in system level. Since maintenance practiceon-board ships is so critical, the integrity of planning, coordination and execution stages are expected to be well functioned. At this insight, this study considers the implementation potential of maintenance 4.0, relatively a new concept defined within industrial 4.0 initiative, in marine engineering field. In detail, the generic design and implementation requirements of maintenance 4.0 including key opportunities and challenges are identified. Moreover, a comprehensive requirement analysis is conducted to adopt maintenance 4.0 concept into ship machinery systems of existing and newbuilding merchant ships.

Keywords: Marine engineering, ship maintenance, maintenance 4.0, industry 4.0

1. INTRODUCTION

Maintenance is a necessary action for industrial organizations in order to retain and conserve organizational assets such as machinery, equipments, replaceable parts, hardwares and other devices. Any failure caused by poor maintenance policy is a threat to operational running of the planned work flow and schedule. On occasion, one of the system failures could lead extremely costful consequences depending on the business type and sectoral vulnerabilities. Besides, newly developed expensive processing systems are used in a widespread manner and companies increase their workforce on maintenance issues. For instance, in chemical industries 30% of the total workforce is assigned to the maintenance related departments (Jonge at al. 2017). Because of such reasons, maintenance strategy has an important role for organizations to achieve business goals by smooth and effective way. In the literature, there are various maintenance strategies which are often categorized in two groups as corrective maintenance (CM) and preventive maintenance (PM). It must be noted that the categorizing process is still under scrutiny by researchers in this field. CM is a maintenance activity which focuses on repairing or mending the failed part or equipment of a system. This process could be involved in more than one recovery approaches such as malfunction isolation, decomposition, replace, re-install, adjustment, verification and fix the relevant failed elements (Fang and Zhaodong, 2013). Therefore, malfunction time, detection and recovery time of a malfunction and the cost of malfunction could be considered for future operations. In run to failure (RTF) maintenance approach, the relevant part, equipment or machinery is simply allowed to breakdown. If the occurance of the problem begins to interfere system functionality, repairing & replacement process is conducted (Piatrowski, 2001). This approach is recommended when operational shutdowns have no serious effect on productibility and equipment costs are not important.

Preventive Maintenance (PM) is a systematic inspection which focused on detecting a malfunction incipient failure. In general, time based maintenance schedule is planned to avoid unexpected shutdowns; therefore system availability is aimed to be maximized (Chalabia at al, 2016; Restrepo, Hennequin & Aguezzoul, 2016; Hadidi, Al-Turki & Rahim, 2012). Depending to the system characteristics, various modified PM applications have been studied, formulated or combined with other maintenance techniques in the literature. Despite these advantages of PM implementations, catastrophic failures are likely to occur and PM activities are usually criticized as labor intensive actions due to involvement of unneeded maintenance operations. In order to eliminate these disadvantages and enhancing the effectiveness of the processes, predictive maintenance (PDM) strategies have been applied in most of the industrial facilities. The PDM differs from PM by

focusing on the actual condition of specified parts of a system notwithstanding a planned schedule. On that sense, regarding technical aspects and physical parameters such as vibration, pressure, flow, or voltage etc. must be monitored by responsible experts who involved in PDM implementations (Raza and Ulansky, 2016). Thus service life of equipments is aimed to be maximized while minimizing the unintended breakdowns and failure risk of system operation (Baidya and Gosh, 2015). In this context, condition based maintenance (CBM) approach come into prominence by many scientists as an effective PDM technique. At the present time, vibration analysis, acoustic emission, ultrasonic testing implementations, oil analysis, strain measurement, electrical effects, shock pulse method, radiographic inspection and thermographic monitoring technologies are utilized to have actual data interpretation in CBM (Marquez at al., 2012). According to obtained data, if the condition gets below of a specified system level, then repair or replacement process initiates after a controlled shutdown.

However, as "maintenance" concept addresses a vast of different fields in global industry; even the latest maintenance techniques could have some ambiguities for a specific system character. Hence, integration of all previously discussed maintenance approaches is utilized in order to have more effective operational results. For instance, reliability centered maintenance (RCM) proposes a combined maintenance approaches philosophy in concert with root cause analysis. Accordingly, RCM gives cause for a mathematical alghoritm or software program which focuses on specific failure modes most probably to occur (Yssaad, Khiat & Chaker, 2012). If more than 20-25%of maintenance workloadis based on breakdowns in an organization, RCM method could be beneficial to the overall process(Smith and Hitchcliffe, 2004).

With the emerging of new technologies such as remote diagnostic systems, e-maintenance concept has come to the forefront since 2000. E-maintenance provides an opening and gives wide coverage to information and communication technologies (ICT) for global companies who involved in competitive industrial activities in the world. Successful applications of e-maintenance strategy will bring benefits to the system reliability while integrating the costumers and suppliers with the aim of zero downtime performance (Han and Young, 2006). Even so, challenges of e-maintenance have been widely discussed regarding to aspects of human resource and training requirements, decision making process, difficulties when designing new business models and complexity of industrial adoption (Chowdhury at al., 2012).

Ever-growing technologic advancements on automation systems and data exchange became an opportunity to open a new door for even more futuristic maintenance concepts. Most of the industrial environments consider maintenance 4.0 as one of the most prudential subjects originated from "industry 4.0" theme. In the industry 4.0, the main idea is to achieve higher availability performance, reduced downtimes, optimized energy consumption and cut general maintenance costs by interoperability of machines supported with internet of things (IoT), internet of services (IoS) and internet of people (IoP) (Zezulka at al., 2016). Moreover, in addition of the cyber physical systems (CPS); equipments, sensors, machines, products, supply chain elements and customers get linked with each other thus a system's essential objects will share information in order to conduct control operations autonomously (Qin, Liu & Grosvenor, 2016). So, the term of industry 4.0 have become the most popular subject among industry and academia, as result it has already dubbed as the fourth industrial revoulition with the expected marginal effect (Kagermann at al., 2013). In this industrial transformation, the maintenance 4.0 concept has an important role to contribute future's maintenance standard and it lies in the core of industry 4.0 due to requirements for new developed expensive systems, smart machines, autonomous devices and tools.

2. LITERATURE REVIEW

As industry 4.0 is still in the early stages of implementation, academia is discussed this phenomena and its key components in attempt to develop smart facilities of the future. CPS is the most used term in the industry 4.0 literature. According to Kagermann (2014), CPS is a combination of virtual environment and real life. Indeed, considering the underpinning aspects of industry 4.0, CPS has crucial importance due to potential benefits such as decentralizing system data. In contrast, case studies could be more explanatory to have more explicit understanding for both current challenges and potential handicaps. For instance, Stock and Seliger (2016) interrogate the industry 4.0 in the context of micro and macro perspectives in a case study for manufacturing a desktop tool. In this study, expected developments and opportunities elaboratively examined and findings demonstrated widely. They have created human-machine interaction through integrating CPS into their project and proposed a smart production factory model. Lee, Kao and Yang (2014) highlighted the importance of big data environment with the purpose of effective CPS implementations and in addition, a case study over Komatsu smart bulldozer has been demonstrated to have wide range of perspective on remote controlled self-maintenance availability in the scope of industry 4.0. Another case study on ongoing projects have been conducted by Carlos Toro et al. (2015) with the intent of establishing an architecture framework on knowledge based intelligent systems. That is because, delays of communication systems on given equipments cause an uncertainty, so reliable middlewares are needed if unpredicted situations desired to be eliminated. In addition to the previous studies, Bagheri at al. (2015) have designed an adaptive method between differently settled units and integrated this approach into their case study. Ivanov, Sokolov and Ivanova (2016) bring benefits to the CPS utilization with a mathematical alghoritm in order to establish smooth dynamic coordination of gathered data. In reference, any problem is aimed to be solved simultaneously by means of effective

information system. Different angles of CPS integration into industry 4.0 is widely discussed in the literature; see also Albers at al. (2016), Landherr, Schneider & Bauernhansl (2016), Kohlberg and Zuhlke (2015), Lee (2008), Wittenberg (2016), Dutra and Silva (2016), Monostori at al (2016), Garcia-walls at al (2017). As a general idea, there are various challenges to achieve for proper practice of CPS such as complexity, static and dynamic composition of sub-systems, processing big data, governance, general uncertainties and required employee consciousness. As a whole, challenges for the industry 4.0 concept are also particularly studied. To measure maturity level and readiness of manufacturing enterprises, Schumacher, Erol & Sin (2016) submitted 9 dimensions and 62 assigned items and tested them in several companies for validation. The 9 dimensions such as strategy, leadership, customers, products, operations, culture, people, governance and technologymake possible a self assessment model of the organizations who attempt to be a part of industry 4.0 applications. Similarly, Thames and Schaefer (2016) noted that agility, flexibility and adaptability should be enhanced for actual development so, in this regard; they have suggested software architecture (software defined cloud manufacturing). As another scale, human adaptation is also under discussion by some scientists. Because, the general opinion; necessity to the human workforce is expected to diminish with the emerging of this concept. Nevertheless, new technologic developments will be in need of human contribution but from different aspects. In this sense, Erol at al (2016) put emphasys on the complexity and abstractness of the actual execution and discussed required competences for workers as another challenge to achieve. To eliminate this, they proposed a scenario based learning approach for different type of employee roles as manager, engineer and worker. Hecklau at al (2016) also highlighted the same challenge and they proposed required competences in their paper as comprehensive list of essential competences to support human resource management issue. In contrast, Blöch and Schneider (2016) also underlined training of workers and discussed a learning factory which conducts training practices by simulations. In order to train human workers properly; Quint, Sebastian and Gorecky (2015) proposed a system architecture for a mixed environment of CPS through augmented reality visulation techniques.

Eventually, in these studies; opportunies, challenges, benefits and potential impacts of industry 4.0 are thoroughly discussed. The concept of industry 4.0 and the term of maintenance 4.0 are very concentric with each other and as a consequence, in most of relevant studies; the term of “maintenance 4.0”is not preferred by authors even if they contribute to the maintenance literature. However, examined papers in the literature review have broad information on both industry 4.0 and maintenance 4.0. In order to contribute to the literature; this paper discusses the shipboard integration requirements and challenges via “maintenance 4.0” terminology.

3. SHIP MAINTENANCE PROCESS

A ship performing at sea has many complex aspects comparing with other industrial transportation units. Any breakdown could lead very costly consequences due to limited repair and recovery options while operating far from the land based service facilities such as shipyards, ports etc. In addition, unexpected downtimes and delays mean high amount of money loses for shipping companies so they could be verydifficult to compansate. At times, defects are not detected onboard ship until it is too late. Despite the current efforts on ship maintenance, machinery damage is still most common cause of ship incidents by 36% according to the Allianz Safety & Shipping Review (2016). As ship maintenance process executed by few number of engineers, detailed checklists and everlasting machinery tasks must be followed anyhow in accord with planned maintenance system (PMS). PMS is widely used system which allows conducting maintenance operations in sync with the requirements of classification societies and other including parties such as shipyards. Because of the preventive maintenance (PM) nature of PMS, ship crew is responsible to fulfill PMS tasks properly at the right time. In this system, time based schedules and running hour based plannings allow ship operators to maintain equipments before exposing risks which may lead breakdowns. In other words, any machinery part already involved in PMS checklist has to be overhauled even if there is no predicted problem. Carried out tasks, PMS checklists and related documents are supervised by periodic inspections at the intervals in the scope of ISM Code (International Safety Management Code, Chapter 5, Section 10).

In addition to the PMS, corrective maintenance (CM) and run to failure (RTF) is also conducted in marine vessels at operational level. As mentioned previously, in RTF maintenance process is conducted only if a breakdown occures. Depending to the equipment type, existing spare parts or usable lifespan of an item; CM activities could be advantageous in some specific cases. However, CM and RTF most likely to cause major problems in the long term and as a matter of fact, it may lead more costly results such as ship disruption, expensive repair actions, port or class detentions, etc. For this reason, more reliable alternative ship maintenance methods have been analysed by researchers and followingly; condition based maintenance (CBM) has come forward thanks to technologic advancements in condition monitoring appliances recently.

CBM is still not a commonly using maintenance policy due to its additional requirements comparing with classical maintenances techniques but when it is done properly, operational efficiencies are expected to improve significantly. Maintenance costs and machinery failures could be reduced by means of instant monitoring of the critical system components. For existing ships, CBM is not preferred due to ship owners’ unwillingness towards its actual integration thus newbuilding ships are considered as most likely candidates for this method. However, it is simpler to implement on new ships during their building period because required sensors, data networks and other technologic infrastructures can be

easily integrated before entering service life. Although CBM pledges cost effective maintenance operations onboard, has some challenges as: breakdowns on technologic systems, missing or false monitoring due to sensor deterioration, competence requirements for ship operators and software based errors on data processing experiences.Apart from these maintenance approaches, reliability centered maintenance (RCM), e-maintenance and maintenance 4.0 applications have not been preferred notably in current shipping activities.

4. REQUIREMENT OVERVIEWS

Comparing with other systems, ship machinery is often larger, complex and difficult to operate so repairing process can be time-consuming, costly and can be even hazardous activities. Hence, PDM concepts such as maintenance 4.0 are crucial for ships in the sense of reducing downtime, financial loses and potential risks to human health. Since industy 4.0 is labelled as the fourth industrial revolution which is upcoming with highly expected major contributions into many industries, challenges are shown differences based on the types of industrial fields. In this point of view, requirements of maintenance 4.0 for maritime industry and potential challenges caused by industrial characteristics must be identified.

A ship's main engine system's critical elements generally consist of: pistons, injections, exhausts, valves, pipings and as the fundamental energy source; diesel oil, naturally. In order to obtain data from these parts of main engine for data process through IoT and cloud based software systems, new generation condition measurement units are required. Newly developed smart sensors are very capable for such duties as their functions are not limited with representing physical quantites, they are also able to receive, transmit, elaborate data and commands via digital channels (Fumagalli at al., 2010). For local data exchange of main engine system, radio-frequency identification (RFID) technology is suggested. RFID devices can support data exchange and also data storage automatically in a given local area, effectively. (Fumagalli at al., 2010). If RFID is not preferred, an IoT based autonomous alternative data transferrable and storageable items should be used. In addition to the previous technology use, PLC (programmable logic controller) or infrared technologies could be chosen based on the failure modes.

Failure modes can be various for each system component in main engine. Selected failure modes for the given main engine parts are could be: Noise, vibration, temperature, oil leakages, unexpected suspendings for no reason, fail to start, structural deficiency, abnormal fluid viscosity and indicator malfunction. Obtained data from given equipments and failure mode types must be transmitted to integrated data acquisition and signal processing tool: Most widely used system is Supervisory Control & Data Acquisition Systems (SCADA). By means of SCADA, elaborated information could be provided for the operating engineer (Fumagalli at al., 2010). In addition, system condition is monitored by engineering staff even if they are not on their working area by means of digital devices and synchronized network between human and machine (Campos at al., 2009). However, all obtained information must be gathered on a database to analyse, manage, redesign or reschedule the current maintenance operations. Cloud system which is supported with a big data and a decision making model is adviced at this stage. Even so, newly developed software based alghoritms and architectures in the scope of maintenance 4.0 may be proper alternatives in the future, as if they would be more effective. Wireless connection type (See also: Bluetooth, GSM, UMTS... etc) could be various thus it should be determined properly considering the whole preferred maintenance 4.0 items. Besides, connection with the shore, company head office, flag state administrations, port state controls or regional memorandums would be very beneficial to the merchant ship's overall conditions including maintenance and repairing issues. In the context of maintenance 4.0, augmented reality should also be implemented for use of ship crew.

In addition to the ship main engine and its components, a comprehensive maintenance 4.0 design should be implemented in ship's hull, steering gear systems, ballast water tank and fuel tank equipments, boiler and related systems etc. By synchronization of such local systems to the main database network, maintenance 4.0 requirements will be fulfilled. However, these arrangements are likely to face with many challenges for the near future.

Integrating maintenance 4.0 concept into newbuilding merchant ships are relatively easy comparing with the existing ships. Because, from the very beginning of their life cycle, ships could be designed conveniently with maintenance 4.0 concept. Besides, there could be a plenty of time for analyses and tests which are executed for probable failures. Existing ships have serious limitations for this application. Ship owners are usually reluctant to integrate new systems into their ships due to financial concerns. A maintenance 4.0 system's implementation process could take many days depending on ship characteristics. In addition, their complex and outworn structure will not allow for an easy modification processes. Moreover, there are various types of ships and also too many different sized ships. For this reason, software based failures are considered as one of the inevitable problems.

Additionally, for both newbuilding and existing ships; there are still several common challenges. Ships are operating under different environmental conditions so already recorded previous failure data may be not compatible for another ship. Another concern is national and international strict norms dictated by regulatory bodies such as international maritime organization (IMO), classification societies, flag state control, port state control etc. Any change that not recognized by

these societies could not be permitted until new regulation is entried into force. If somehow, maintenance 4.0 succesfully applied to a ship's body; this new concept imposed extra workload to the already overburdened ship operators. Furthermore, ship operators must be trained for maintenance 4.0 specificly. However, they would not be eager for this new training program as they have been already attended for numerous certificate programs in order to earn their current job positions. Even if they have trained well, crew of a ship is changing frequenly and familiarization for a new ship's complex system is going to take a long time.

Despite the challenges, if mentioned problems are solved; ship operations become significantly efficient and maintenance based downtimes, breakdowns and failures are eliminated on a large scale. Likewise, incidents and potential risks sourced by maintenance operations are likely to diminish substantially. In the same way, maintenance operations of a ship may be inspected remotely with the networks that accessed to the cloud systems thus rigid prescriptive requirements of recognized bodies could be ameliorated. As result, fuel consumption of merchant ships becomes even lower due to conveniently maintained machinery system and followingly, more environmental friendly worldwide shipping services could be achieved.

5. CONCLUSION

This study discusses the implementation potential of maintenance 4.0 into marine engineering field. Firstly, maintenance methods CM, RTF, PM, PDM, RCM, e-maintenance and maintenance 4.0 are examined. Hence, technological requirements of maintenance 4.0 and industry 4.0 activities thoroughly analysed and presented for ships. Requirements are very dependent on type and characteristic features of a ship but in general, the most importantrequirements are:

- New generation condition measurement units : Smart sensors and RFID items.
- Based on the failure modes, PLC and infrared monitoring technologies may be needed.
- For signal processing tool, SCADA is required.
- CLOUD based system supported with big data or proper decision making approaches.
- Wireless connections such as Bluetooth, GSM or UMTS.
- Specialized failure mode sensors and data transfer units for engine room, ship hull, steering gear, ballast water system, fuel tanks and boiler systems.

Almost every field of global industry has some typical challenges for actual integration of the maintenance 4.0 concept. Identically, marine engineering has distinctive challenges as well as has typical opportunities. So, challenges and opportunities are identified via conducting a requirement analysis for onboard maintenance operations. Summarily, old ships have very little potential for the maintenance technique of fourth industrial revolution. Whilst newbuilding shipsmore likely to be a part of maintenance 4.0 but there are still serious challenges and uncertainties such as;

- Ships are operating in ever changing conditions
- There are numerous strict norms, classification societies and other organizations in maritime sector
- Crew is unskilled for such innovations and they are already overloaded
- Comprehensive specialized education programmes are required for ship engineers
- There are uncertainties in softwaring processes due to various ship types
- Unwillingnessof ship owners to the major innovative integrations
- Different ship characteristics will retard the know-how process from experienced ships
- Different ship characteristics make it difficult to manage for newly assigned ship operators to another system design.

REFERENCES

Albert A., B, Gladysz B., T, Pinner, V., Butenko, T., Stürmlinger (2016).Procedure for Defining the System of Objectives in the Initial Phase of an Industry 4.0 Project Focusing on Intelligent Quality Control Systems, Procedia CIRP, Volume 52, 2016, Pages 262-267, ISSN 2212-8271, http://dx.doi.org/10.1016/j.procir.2016.07.067.

Allianz Shipping Review (2016) Allianz 2016, http://www.agcs.allianz.com/assets/PDFs/Reports/AGCS_Safety_Shipping_Review_2016.pdf, date retrieved: 01.03.2017.

Baidya, R., & Ghosh, S. K. (2015). Model for a predictive maintenance system effectiveness using the analytical hierarchy process as analytical tool. IFAC-PapersOnLine, 48(3), 1463-1468.

Bagheri, B., Yang, S., Kao, H. A., & Lee, J. (2015). Cyber-physical Systems Architecture for Self-Aware Machines in Industry 4.0 Environment. IFAC-PapersOnLine, 48(3), 1622-1627.

Blöchl, S. J., & Schneider, M. (2016). Simulation Game for Intelligent Production Logistics–The PuLL® Learning Factory. Procedia CIRP, 54, 130-135.

Campos, Jaime, Erkki Jantunen, and Om Prakash. "A weband mobile device architecture for mobile emaintenance." *The International Journal of Advanced Manufacturing Technology* , 2009.

Carlos Toro et al. (2015)*Procedia Computer Science 60* (2015) 362 – 370

Chalabia,N. , M. Dahaneb, M., Beldjilalia, B., Nekic, A. (2016). Optimisation of preventive maintenance grouping strategy for multi-component series systems: Particle swarm based approach. *Computers & Industrial Engineering*. Volume 102, December 2016, Pages 440–451

Chowdhury, Soumitra, Asif Akram, and Maria Åkesson. "E-maintenance as an emerging customer value generating it-enabled resource." *7th Mediterranean Conference on Information Systems*, Guimarães, Portugal, September 8-10, 2012. 2012.

Dutra D. S., Silva, J. R. (2016). Product-Service Architecture (PSA): toward a Service Engineering perspective in Industry 4.0, IFAC-PapersOnLine, Volume 49, Issue 31, 2016, Pages 91-96, ISSN 2405-8963, http://dx.doi.org/10.1016/j.ifacol.2016.12.167.

Erol, S., Jäger, A., Hold, P., Ott, K., & Sihn, W. (2016). Tangible Industry 4.0: a scenario-based approach to learning for the future of production. Procedia CIRP, 54, 13-18.

Fang, L.. Zhaodong, H. (2013). System Dynamics Based Simulation Approach on Corrective Maintenance Cost of Aviation Equipments. *Procedia Engineering*, Volume 99, 2015, Pages 150-155

Fumagalli, L., Di Leone, F., Jantunen, E., & MacChi, M. (2010). Economic value of technologies in an e-Maintenance platform. *IFAC Proceedings Volumes* (IFAC-PapersOnline), 1(PART 1), 18–23. https://doi.org/http://dx.doi.org/10.3182/20100701-2-PT-4012.00005

García-Valls, M., Calva-Urrego, C., Alonso, A., & de la Puente, J. A. (2016, April). Adjusting middleware knobs to suit CPS domains. In Proceedings of the 31st Annual ACM Symposium on Applied Computing (pp. 2027-2030). ACM.

Hadidi, A., Al-Turki, U, M., Rahim, U, A. (2012) Practical implications of managerial decisions to integrate production scheduling and maintenance, *International Journal of System Assurance Engineering and Management*, 224-230.

Han, T., Young, B. (2006). Development of an e-maintenance system integrating advancedtechniques. *Computers in Industry*, 57(6), pp. 569-580

Hecklau, F., Galeitzke, M., Flachs, S., & Kohl, H. (2016). Holistic Approach for Human Resource Management in Industry 4.0. Procedia CIRP, 54, 1-6.

ISM, ISM code, chapter 5. Retrieved 28 December 2013

Ivanov, D., Dolgui, A., Sokolov, B., Werner, F., & Ivanova, M. (2016). A dynamic model and an algorithm for short-term supply chain scheduling in the smart factory Industry 4.0. International Journal of Production Research, 54(2), 386-402.

Jonge, B., Teunter, R., Tinga, T. (2016) 30The influence of practical factors on the benefits of condition-based maintenance over time-based maintenance. *Reliability Engineering & System Safety*, 2017 vol: 158 pp: 21-

Kagermann, H. J. Helbig, A. Hellinger, W. Wahlster (2013). Recommendationsfor Implementing the Strategic Initiative INDUSTRIE 4.0: Securing theFuture of German Manufacturing Industry; *Final Report of the Industrie 4.0 Working Group*, Forschungsunion, 2013.

Kagermann, H. (2014). Chancen von Industrie 4.0 nutzen.Anwendung, *Technologien und Migration*, 603-614.

Kolberg, D., Zühlke, D. (2015). Lean Automation enabled by Industry 4.0 Technologies, IFAC-PapersOnLine, Volume 48, Issue 3, 2015, Pages 1870-1875, ISSN 2405-8963, http://dx.doi.org/10.1016/j.ifacol.2015.06.359.

Landherr. M, Schneider, U., Bauernhansl, T. (2016) The Application Center Industrie 4.0 - Industry-driven Manufacturing, Research and Development, Procedia CIRP, Volume 57, 2016, Pages 26-31, ISSN 2212-8271, http://dx.doi.org/10.1016/j.procir.2016.11.006.

Lee, E. A. (2008). Cyber Physical Systems: Design Challenges, 11th *IEEE International Symposium on Object Oriented Real-Time Distributed Computing*, pp 363-369.

Lee, J., Kao, H. A., & Yang, S. (2014). Service innovation and smart analytics for industry 4.0 and big data environment. Procedia Cirp, 16, 3-8.

Marquez F.P.G, Tobias, A.M. and Perez, J.M.P. (2012). Condition monitoring of wind turbines: Techniques and methods, Mayorkinas Papaelias Renewable Energy, 46, 169-178.

Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., … Ueda, K. (2016). Cyber-physical systems in manufacturing. CIRP Annals - Manufacturing Technology, 65(2), 621–641. https://doi.org/10.1016/j.cirp.2016.06.005 (http://www.sciencedirect.com/science/article/pii/S2212827114003497)

Piotrowski, J. April 2, 2001. Pro-Active Maintenance for Pumps, Archives, February 2001.

Qin, J., Liu, Y., Grosvenor, R. (2016) A Categorical Framework of Manufacturing for Industry 4.0 and Beyond, Procedia CIRP, Volume 52, 2016, Pages 173-178, ISSN 2212-8271, http://dx.doi.org/10.1016/j.procir.2016.08.005

Quint, F., Sebastian, K., & Gorecky, D. (2015). A mixed-reality learning environment. Procedia Computer Science, 75, 43-48.

Raza, A., Ulansky, V. (2016) Modelling of Predictive Maintenance for a Periodically Inspected System, DOI: 10.1016/j.procir.2016.09.032.

Restrepo L.M.R, Hennequin, S., Aguezzoul, A. (2016). Optimization of integrated preventive maintenance based on infinitesimal perturbation analysis. *Comput. Ind. Eng. 98, C (August 2016), 470-482*. DOI: http://dx.doi.org/10.1016/j.cie.2016.06.019

Schumacher, A. Erol, S. Sihn, W. (2016) A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises, Procedia CIRP, Volume 52, 2016, Pages 161-166, ISSN 2212-8271, http://dx.doi.org/10.1016/j.procir.2016.07.040.

Smith, Anthony M., and Glenn R. Hinchcliffe. 2004. RCM: gateway to world class maintenance. Amsterdam: Elsevier Butterworth-Heinemann. http://www.123library.org/book_details/?id=36146.

Stock, T., & Seliger, G. (2016). Opportunities of Sustainable Manufacturing in Industry 4.0. Procedia CIRP, 40, 536–541. https://doi.org/10.1016/j.procir.2016.01.129

Thames, L., & Schaefer, D. (2016). Software-defined Cloud Manufacturing for Industry 4.0. Procedia CIRP, 52, 12-17.

Wittenberg, C. (2016). Human-CPS Interaction - requirements and human-machine interaction methods for the Industry 4.0, IFAC-PapersOnLine, Volume 49, Issue 19, 2016, Pages 420-425, ISSN 2405-8963, http://dx.doi.org/10.1016/j.ifacol.2016.10.602

Yssaad, B., Khiat, M., & Chaker, A. (2012). Maintenance Optimization for Equipment of Power Distribution System Based on FMECA Method. Journal Acta Electrotehnica, 53, 218-223.

Zezulka, F., Marcon, P., Vesely, I., & Sajdl, O. (2016). Industry 4.0 – An Introduction in the phenomenon. *IFAC-PapersOnLine*, 49(25), 8–12. https://doi.org/10.1016/j.ifacol.2016.12.002

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

BUDESONIDE SELF-NANOEMULSION FORMULATION AS AN ORAL DRUG DELIVERY SYSTEM: PREPARATION, CHARACTERIZATION AND CYTOTOXICITY STUDIES

DOI: 10.17261/Pressacademia.2017.569
PAP-WCTIE-V.5-2017(7)-p.41-46

Burcu Mesut, Seyma Hande Tekarslan Sahin, Yildiz Ozsoy
Istanbul University Faculty of Pharmacy Pharmaceutical Technology Dept., bmesut@istanbul.edu.tr

ABSTRACT

Lipophilic drugs which are poorly water soluble need specific applications like lipid based delivery systems to increase their oral absorption across gastrointestinal tract. Selfnanoemulsifying drug delivery systems are Lipid Based Drug Delivery systems (SNEEDs) and are consists of an isotropic mixture of oil, surfactant(s) and co-surfactant. This drug delivery systems are prepared at room temperature with gentle stirring process without using any other component or heating process and become nanoemulsion form upon gentle agitation in aqueous phase. SNEDDs have many advantages such as they can be produced easily and they are affected less from other outside factors, its scale-up process is easy, unaffected by lipid digestion, carry and protect peptides which can be enzymatically hydrolyzed. The aim of this study was to increase solubility of Budesonide by using new Snedds formulation approach.

Keywords: Budesonide, self-nanoemulsion drug delivery, oral drug delivery, cytotoxicity.

1. INTRODUCTION

Budesonide (16,17-butylidenebis(oxy)-11,21-dihydroxypregna-1,4-diene-3,20-dione) is a synthetic glucocorticoid steroid used in the treatment of asthma and an inflammatory bowel disease (Hvizdos and Jarvis, 2000). Molecular weight of Budesonide is 430.5, log P values is 3.2, BCS class II and pratically insoluble in water (28 μg/mL) (Ali et al., 2010). It is a potent nonhalogenated corticosteroid and has a relative glucocorticoid receptor affinity of 935 compared with 100 for dexamethasone and a 200-fold higher glucocorticoid receptor affinity and a 1000-fold higher topical anti-inflammatory potency than cortisol.(Ek et al., 1999). Budesonide is completely absorbed from the gastrointestinal tract, however its bioavailability is around 10% due to hepatic first-pass effect (Meloche et al., 2002). Due to degradation of Budesonide in the liver easily, an alternative way should be developed to avoid the first-pass effect thereby the cellular uptake in the site of action is enhanced (Piao et al., 2009). Approximately 50% of the drugs in use all over the world are oral medications and the market share continues to increase despite the emergence of new dosage forms (http://www.slideshare.net/jaydeep.adhikari/evolution-of-controlled-release-Technology,http://www.marketsandmarkets.com/Market-eports/north-american-drug-delivery-technologies-market-1209.html). Oral drug delivery is the most accepted routes because of its patient compliance (Bernkop-Schnürch, 2013). Even though many pharmaceutical compounds have poor water solubility (Amidon et al., 1995), lipid based drug delivery systems (LBDDS) overcome the low bioavailability problem by increasing solubility of this compounds (Pouton, 2000). Self-nanoemulsifying drug delivery systems (SNEEDs) are LBDDS which consist of oil, surfactant and co-surfactant (Bu et al., 2016).

SNEDDs are thermodinamically stable systems. They form emulsion spontaneously when they encounter gastrointestinal fluids, therefore they have fast dispersion process and are less affected by other factors such as food effect during its formation process (Jankovic et al., 2016). After dilution, they create nanoscale size emulsion droplets. Decreasing droplet size increase surface area and improve the dissolution rates as well as bioavailability (Balakumar et al., 2013).

Cell viability assays are used for in vitro toxicity studies. MTT assay is the most used assay for cell viability. MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) is a tetrazolium salt and is converted to an insoluble purple formazan in viable cells. This conversion is done by succinate dehydrogenase within the mitochondria (Fotakis and Timbrell, 2005). Initial protocol of MTT assay was described by Mossmann (1983). Mitochondrial dehydrogenases in living cells

reduce MTT into formazon salts which form blue-magenta colour. When cytotoxic compounds damage the cells, MTT can not be reduced by cells and as a result, there is a decrease of the reduction of MTT to formazan (Mueller et al., 2004). The aim of this study was to develop new SNEDD formulation of Budesonide and thus overcoming its first-pass degradation, evaluating specification of this new formulation and performing cytotoxicity studies of budesonide-loaded SNEDD formulation.

2. MATERIALS AND METHODS

2.1. Materials

Budesonide was kindly donated from Referans Temsilcilik, Turkey. Labrafac Lipophile WL1349, Labrafac PG, Peceol, Pliurol oleique, Transcutol, Labrasol, Labrafil M2125 was gifted from Gatteffosse. RPMI 1640 medium, 0.05% trypsin–EDTA solution, DPBS without Calcium, Magnesium and phenol red, Fetal bovine serum (FBS), were purchased from PAN-Biotech (Germany). Penicillin-Streptomycin (10,000 U/mL) were obtained from Gibco (USA). 3-(4,5-dimethyl-thiazol-2-yl)-2,5-diphenyl-tetrazolium bromide (MTT) were obtained from Sigma. All other chemicals were of analytical grade.

2.2. Methods

2.2.1. Preparation of SNEDDs Formulations

Self-nanoemulsions were prepared by mixing of oil phase; Labrafac Lipophile WL1349, Labrafac PG and Peceol, surfactant; Labrasol® and Gelucire 48/16® and co-surfactant; Transcutol®. Ternary phase diagrams of surfactants, co-surfactants and oils were constructed to recognize the zone of nanoemulsion formation. Samples were prepared and for each sample oil and surfactant(s): co-surfactant ratio were mixed in ratios ranging from %10:90 - %90:10 like 1:9, 2:8, 3:7, 4:6, 5:5, 6:4, 7:3, 8:2 and 1:9. Once the self-emulsifying region was identified, the desired component ratios of SNEDDS were selected.

2.2.2. Analytical Method

Budesonide is an epimeric mixture of two isomers. USP Method was modified to quantitate budesonide in formulations. The HPLC system utilized consists of a LC-20AD liquid chromatography equipped with 20A DAD detector. Chromatographic separations were performed on C18 (4.6 × 250 mm, L1; 5 µm packing) which was attached with loop 100 µL.Mobile phase of acetonitril: Buffer solution (pH 3.2) (32:68 (V/V)) was used. Buffer solutions were filtered through filtration unit (Millipore, 0.45 µm pore size) and degassed before use. The flow rate was maintained at 1.5 mL/min and injection volume was 50 µL. Detection was performed at a wavelength of 254 nm and analysis was carried out at ambient temperature.

2.2.3. Solubility Studies in Excipients

Determination of drug solubility of Budesonide an excess amount active pharmaceutical ingredient (API) in various oils, surfactants and co-surfactants were carried out by adding an excess amount of API in to 1 ml of the vehicle. The mixture was shaken in an orbital shaker (25 ± 1 °C)(Thermo, USA) for 24 h. The equilibrated mixture was centrifuged at 15000 rpm for 15 min and excess insoluble budesonide was removed by filtration. The filtrate was analyzed for the amount of budesonide RP-HPLC (Shimadzu, Japan) at 254 nm.

2.2.4. Solubility Studies at Different pHs

Budesonide solubility at different pH were analyzed. The Buffer solutions were prepared at three deifferent pH (1.2, 4.5 ve 7.4) to determine solubility of budesonide at physiologyc pH values. An excess amount of budesonide were added in 1 mL buffer solutions. The mixture was shaken at 25 ± 1 °C in an orbital shaker (Thermo, USA) for 24 h. The equilibrated mixture was centrifuged at 15000 rpm for 15 min and excess insoluble budesonide was removed by filtration. The filtrate was analyzed.

2.2.5. Physicochemical Properties

The pH value of Snedds formulations were analysed at 25 ℃ temperature by using Inolab pHmeter. Droplet size and polidispersity index values were measured by Zetasizer after diluting in purified water and buffer solutions at different physiological pHs at 25℃ and 37℃.

2.2.6. Drug Loading Capacity of Chosen Formulations

The solubility of budesonide in formulations was determined. The same procedure was applied as excipients solubility studies.

2.2.7. Cell Culture

The Human Colorectal Adenocarcinoma Caco-2 was cultured in RPMI1640 supplemented with 10% FBS and 100 units/mL penicillin-Streptomycin. Cells were cultured at 37°C in a humidified atmosphere of 5% CO_2. Medium was removed third times a week. Confluent cells were removed from cell culture dishes with 0.25% sterile trypsin.

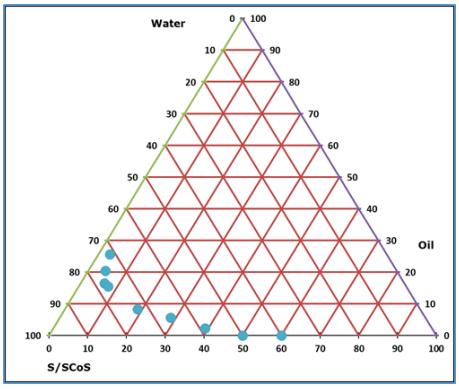
2.2.8. Cell Viability Assay

MTT (3-[4,5-dimethylthiazol-2-yl]-2,5 diphenyltetrazolium bromide) assay was used to determine cell viability. Caco-2 cells (2× 104 cells/well) were cultured in 96 well plate at 37°C in a humidified atmosphere of 5% CO2. After cells were grown for 24 hours, budesonide loaded SNEDDS were applied to wells at concentration of 0.05, 0.10, 0.25, and 0.50% (v/v) for 2 h and 24 h. Cells treated with medium only is used as control group. After treatment with SNEDDS, the supernatant was discarded and MTT solution (5 mg/ml in PBS) and medium were introduced to wells. After incubation with MTT solution, MTT solution was replaced with DMSO to dissolve the formazan crystals. The optical density (OD) of the wells was determined using a plate reader (Thermoscientific Multiskan Ex, USA) at a test wavelength of 570 nm and a reference wavelength of 630 nm. All experiments were performed in triplicate, and the relative cell viability (%) was expressed as the absorbance ratio between budesonide-loaded SNEDDS treated and untreated control cells.

3. RESULTS and DISCUSSIONS

Phase diagram studies help to understand phase behavior of the formed nanoemulsion (Balakumar et al., 2013). Diagrams were constructed to determine the self-nanoemulsifying area and to make a decision about an optimum concentration of oil, surfactant and co-surfactant for the development of liquid SNEDDS formulations. An example of phase diagram studies was given in Figure 1.

Figure 1: A Sample of Phase Diagram Study (oil: Peceol, surfactant/Co-surfactant: Labrasol-Transcutol 1:1)



Solubility studies in excipients play important role to determine SNEDD formulation stablities and to prevent in situ precipitaion previously (Parmar et al., 2011).

Figure 2: Drug Solubility Results in Excipients

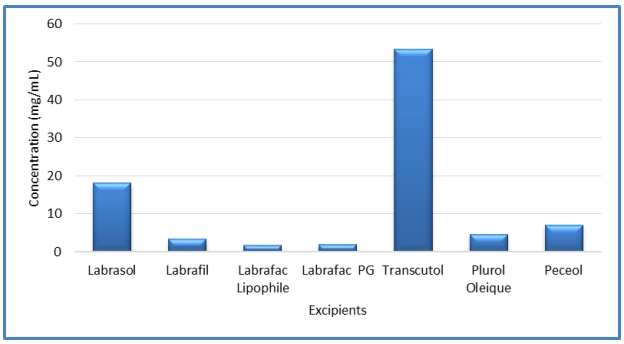


Figure 3: Drug Solubility Results at Different pH Values

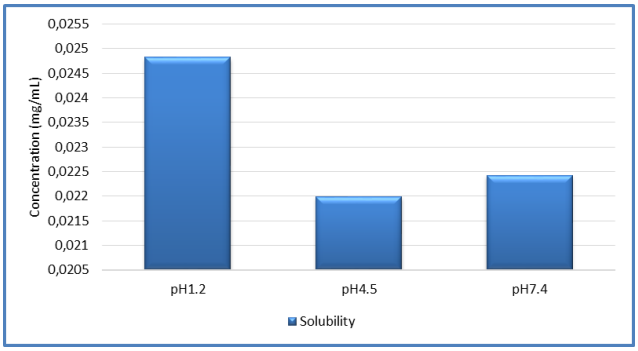
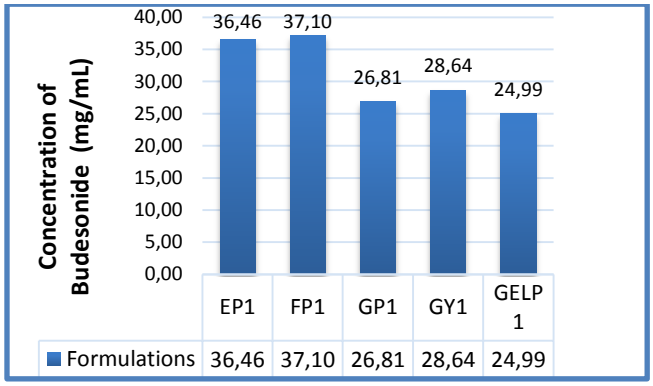


Figure 4: Drug Loading Capacity of Thechosen Formulations



The obtained mixture of the formulation should be clear, monobasic at the room temperature and isotropic (Qi et al., 2011). Nanoemulsion Droplet size is an important characteristics to evaluate formulation stability and a critical criterion of enhancing drug bioavailability (Xi et al., 2009). The droplet size of 1% solution of SNEDD in the different buffer solution pHs were approximatly 18 nm and PDI values are below then 0.3. This results show us how SNEDD formulation is stable even at different pH values and that means SNEDD globules are not affected at different physiological fluids (See Table 1).

Table 1: Droplet Size (nm) of 1% Solution in Different pH Buffer Solutions

	pH 1.2		pH 4.5		pH 6.8		pH 7.4	
	25℃	37℃	25℃	37℃	25℃	37℃	25℃	37℃
$Size_{average}$ (nm)	16,78	17,24	19,22	19,73	17,37	17,95	18,01	18,56
$PDI_{average}$	0,238	0,273	0,062	0,115	0,135	0,153	0,268	0,286

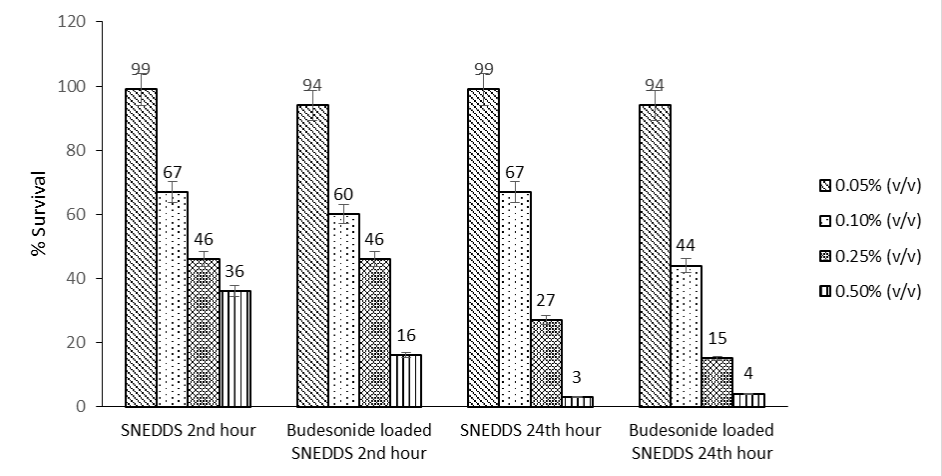
The change in droplet size after 24 hours in 1% pH 6.8 buffer solution was investigated. There was no change in droplet size after 24 hours. Test results were given at Table 2.

Table 2: Droplet Size (nm) of 1% Solution at pH 6.8 after 24 h.

	pH 6.8	
	25℃	37℃
$Size_{average}$ (nm)	17.17	17.86
$PDI_{average}$	0.128	0.167

MTT assay was carried out to determine suitable concentration for SNEDDS formulation. Caco-2 cells were exposed to budesonide-loaded SNEDDS for 2 and 24 h. MTT assay test of this formulation showed that toxicity of formulations were concentration-dependent. The empty SNEDDS formulation and budesonide-loaded formulation have significantly higher viability at a concentration of 0.05 (v/v) compared to the other concentrations (0.10, 0.25, and 0.50% (v/v)). Cytotoxicity is not observed in SNEDDS at concentration of 0.05 (v/v). Loss of cell viability was obvious at concentration of 0.10 (v/v). While there is no difference between 2h and 24h for survival ratios at concentration of 0.05 (v/v), increase of the concentration decreases survival ratios. As the concentration of SNEDD formulation gets higher (0.25 (v/v) and 0.50% (v/v)), administration of the SNEDD formulation for 24h becomes more toxic compared to 2 h. These results showed that the cell viability was decreased as a function of time at concentration of (0.25 (v/v) and 0.50% (v/v)). Budesonide loaded SNEDD and empty SNEDD at concentration of 0.05 (v/v) can be regarded as non-toxic for Caco-2 cells during the incubation time. SNEDDS at concentration of 0.25 (v/v) and 0.50% (v/v) result in a significant lower viability in Caco-2 cells due to containing higher concentration of Peceol which is a surfactant.

Figure 5: Cell Viability for Caco-2 Cells Following Exposure to SNEDDS as Derived from MTT Assay (n=3).



4. CONCLUSION

Each formulation was evaluated based on their physicochemical characteristic and formulations which contain Peceol as oil phase, Labrasol and Gelucire 48/16 as surfactants and Transcutol as co-surfactant were chosen. To have an idea about the fate of the formulation in GIT, in vitro release studies will be performed.

This is the first study, to our knowledge, budesonide loaded SNEDDS were produced and characterized. Budesonide-loaded formulation (0.05, 0.10% (v/v)) exhibited low toxicity. It demonstrated that this formulation at concentrations of 0.05%, 0.10% (v/v) are convenient for oral administration.

ACKNOWLEDGMENT

This study supported by Istanbul University BAP no: 47691.

REFERENCES

Ali, H. S., York, P., Blagden, N., Soltanpour, S., Acree Jr., W. E., Jouyban, A. 2010. "Solubility of budesonide, hydrocortisone, and prednisolone in ethanol + water mixtures at 298.2 K," , Journal of Chemical and Engineering Data, vol. 55(1), pp. 578–582.

Amidon, G.L., Lennernäs, H., Shah, V.P., Crison, J.R. 1995. "A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability", Pharmaceutical Research, vol. 12, pp. 413–420.

Balakumar, K., Raghavan, C.V., Selvan, N.T., Prasad, R.H., Abdu, S. 2013. "Self nanoemulsifying drug delivery system (SNEDDS) of Rosuvastatin calcium: Design, formulation, bioavailability and pharmacokinetic evaluation". Colloids and Surfaces B: Biointerfaces, vol. 112, pp. 337–343.

Bernkop-Schnürch, A. 2013, "Nanocarrier systems for oral drug delivery: Dowereally need them?". European Journal of Pharmaceutical Sciences, vol. 49, pp. 272–277.

Bua, P., Narayanana, S., Dalrymplec, D., Chenga, X., Serajuddin A.T.M. 2016, "Cytotoxicity assessment of lipid-based self-emulsifying drug delivery system with Caco-2 cell model: Cremophor EL as the surfactant", European Journal of Pharmaceutical Sciences, vol. 91, pp. 162–171.

Ek, A., Larsson, K., Siljerud, S., Palmberg, L. 1999. "Fluticasone and budesonide inhibit cytokine release in human lung epithelial cells and alveolar macrophages", Allergy, vol. 54, pp. 691–699.

Fotakis, G., Timbrell, J.A. 2006. "In vitro cytotoxicity assays: comparison of LDH, neutral red, MTT and protein assay in hepatoma cell lines following exposure to cadmium chloride", Toxicol Lett., vol. 160(2), pp. 171-177.

http://www.marketsandmarkets.com/Market-Reports/north-american-drug-delivery-technologies-market-1209.html, Accessed on 17.03.2017

http://www.marketsandmarkets.com/Market-Reports/north-american-drug-delivery-technologies-market-1209.html, Accessed on 17.03.2017

Hvizdos, K.M., Jarvis, B, 2000. "Budesonide inhalation suspension: a review of its inflammatory respiratory disorders", Drugs, vol. 60 (5), pp.1141–1178.

Janković, J., Djekic, L., Dobričićc, V., Primoracb. 2016. "Evaluation of critical formulation parameters in design and differentiation of self-microemulsifying drug delivery systems (SMEDDSs) for oral delivery of acyclovir", International Journal of Pharmacy, vol. 497(1–2), pp. 301–311.

Meloche, C.A., Sharma, V., Swedmark, S., Andersson, P., Falany, C.N. 2002. "Sulfation of budesonide by human cytosolic sulfotransferase, dehydroepiandrosteronesulfotransferase (DHEA-ST)", Drug Metabolism & Disposition, vol. 30, pp. 582–585.

Mosmann T. 1983. "Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays". Journal of Immunologyc Methods, vol. 65(1-2), pp. 55-63.

Mueller, H., Kassack, M.U., Wiese, M. 2004. "Comparison of the usefulness of the MTT, ATP, and calcein assays to predict the potency of cytotoxic agents in various human cancer cell lines". Journal of Biomolecular Screening, vol. 9(6), pp. 506-515.

Parmar, N., Singla, N., Amin, S., Kohli, K. 2011. "Study of cosurfactant effect on nanoemulsifying area and development of lercanidipine loaded (SNEDDS) self nanoemulsifying drug delivery system". Colloids Surface. B: Biointerfaces, vol. 86, pp. 327–338.

Piao, H.-M., Cho, H.-J., Oh, E.-C., Chung, S.-J., Shim, .C.-K., Kim, D.-D. 2009. "Budesonide Microemulsions for Enhancing Solubility and Dissolution Rate", Journal of Korea Pharmaceutical Sciences, vol. 39 (6), pp. 417-422.

Pouton C.W. 2000, "Lipid formulations for oral administration of drugs: non-emulsifying, self-emulsifying and 'self-microemulsifying' drug delivery systems", European Journal of Pharmaceutical Science, vol. 11, pp. 93–98.

Xi, J., Chang, Q., Chan, C.K., Meng, Z.Y., Wang, G.N., Sun, J.B., Wang, Y.T., Tong, H.H., Zheng, Y. 2009 "Formulation development and bioavailability evaluation of a self-nanoemulsified drug delivery system of oleanolic acid". AAPS PharmSciTech, vol. 10, pp. 172–182.

USP 31, http://www.usp.org/sites/default/files/usp_pdf/EN/USPNF/revisions/m10458budesonide.pdf, Accesed on 20.03.2017

Qi, X., Wang, L., Zhu, J., Hu, Z., Zhang, J. 2011 "Self-double-emulsifying drug delivery system (SDEDDS): a new way for oral delivery of drugs with high solubility and low permeability" International Journal of Pharmacy, vol. 409, pp. 245–251.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

DISPLAYING OF DATA'S OF DEVELOPED ENVIRONMENTALLY INTELLIGENT SYSTEMS USING WIRELESS SENSOR NETWORKS ON ANDROID PLATFORMS

DOI: 10.17261/Pressacademia.2016.570
PAP-WCTIE-V.5-2017(8)-p.47-51

Murat Dener[1]
[1]Gazi University, muratdener@gazi.edu.tr

ABSTRACT

Wireless Sensor Networks consist of large number of tiny sensor nodes. Each sensor node has communication, sensing and computing capabilities. These sensor nodes perform the environmental measurement tasks with the cooperation. Intelligent systems can be developed using these sensor nodes. Also, the number of mobile applications used in the display of intelligent systems has increased rapidly in recent years. As a result of this rapid increase, displaying and control of data over these mobile platforms have also begun. In this study, a wireless sensor network that detects the environmental data has been established using WiSeN sensor nodes. In addition, a mobile application has been developed to monitor the data's. This study brings innovation to the monitoring of environmental data's.

Keywords: Wireless sensor networks, android platforms, environmental data's

1. INTRODUCTION

It gets harder and harder every day to count the number of the technologies coming into our lives. The concept of mobile has also come into our lives as a result of these developments. Most people think of smart phones when they hear about this concept. Apart from smart phones, many other devices serving for far different areas are also being produced. All mobile devices fall into mobility. Mobile phones have made the concept of mobile get about. These devices used to be utilised for searching and messaging only. But nowadays, many transactions such as downloading from the Internet, making modifications on them and forwarding to someone else are also performed. It would be a big mistake to consider electronic devices, as the software inside is not visible, as a hardware only. As far as the mobile phones are concerned, each one of them has a unique operating system. As per the report of IDC (2017), the Android operating system dominates 84% of the market in the third quarter of 2014. Therefore, the Android operating system is the one with most users. DiMarzio (2008) states that an android operating system is an open source coded mobile operating system created using Linux 2.6 kernel. However, unless it has got applications on it, it resembles a car with no engine. It should be noted that when there is no applications installed, not even search can be made. Every service used on Android operating system is an application.

Wireless Sensor Networks are created by means of connecting many sensor nodes over a wireless connection. These nodes can make happen data communication from one region over the other in a network they create in which they communicate with each other. The data collected from the wireless sensor networks are sent to the database by means of the Internet Module. The Android application created takes the data that have been sent from the data base using web services. The data received are displayed on the Android interface. For instance, the values of temperature, humidity, light and such can be displayed on the interface of the Android instantly using a sensor network infrastructure.

The paper is organized as follows. In the second part of the study information on wireless sensor networks and the sensor node are presented. In the third part, making of the application is explained in detail. In the last part, the conclusion is presented.

2. WIRELESS SENSOR NETWORKS

Akyildiz et al. (2002) state that the progresses in the low-cost sensor architectures have made the Wireless Sensor Networks a new and known research area. Chong et al. (2003) state that Wireless Sensor Networks consist of small, limited-capacity sensor nodes. Lin et al. (2004) state that each sensor node has the ability to calculate, detect and communicate. When the energy of the sensor nodes is exhausted, the life of the node ends. Because of this, low power listening techniques are working within the sensor nodes. The sensor nodes send the information they have detected through cooperation to the gateway node. The information from the gateway node is displayed on the web or mobile platforms. Because of these features, Wang et al. (2006) state that they can be used in a wide variety of areas. Area of usage of the Wireless Sensor Networks is increasing with each passing day. Wireless Sensor Networks can be utilised in many fields including; guarding of battle fields, monitoring of enemy moves, reconnoitring of the fields, monitoring of military and personnel vehicles, monitoring of friendly forces, and determination of the speed and location of the targets in Military applications, in the environmental applications requiring determination of weather forecast, air pollution, monitoring of natural disasters such as floods, earthquakes forest fires, monitoring of agricultural activities and monitoring of animal farms, in medical applications consisting of determining the location of the doctors in the Hospital, monitoring of the conditions of the patients, guarding of the elderly and following some medical parameters, in commercial applications such as monitoring and determining of vehicles, monitoring of power lines, following of children by their parents, checking of lightings, checking of traffic lights and fire alarm systems and in home automation systems requiring smart home application and building security systems. The WiSeN sensor node used in study is shown in Figure 1.

Figure 1: WiSeN Sensor Node

General Characteristics of the WiSeN sensor node (2017): Zigbee IEEE 802.15.4, 2.4 GHz ISM band, CC2530 Radio Module, MSP430 Microcontroller, Low Power Consumption, Low Cost, To integrate the desired sensor, To acquire features to sensor node with C language, Programming via the USB Interface, Remote access to each sensor node, Self-organizing, The advantage of gateway node (Server, Sms, Call), Following datas over WEB, Mobile (Android, iOS) platforms, Ability to communicate with different wireless sensor nodes.

3. APPLICATION

A database is a space in which related data are brought together. With the use of these spaces the access to these data are lot easier. And the web services act as a door for these data to be accessed from outside. Any application requesting to get data from the database is connected to the web service first. Then whatever data the web service is supposed to receive is taken and sent to the application. In the web service created within the framework of this study PHP language has been used. The data are sent and received in JSON. JSON (JavaScript Object Notation) is a text format that facilitates structured data exchange between all programming languages. The application needed in order for the environmental data received from Wireless Sensor Networks to be displayed in Android Platforms comprises of four sections. These are: Application Splash Screen, Registered Users Entry Screen, Category Selection Screen, Screen for Scenarios.

3.1. Application Splash Screen

This is the section where the logo and slogan of the application is located. When the application is started, this section runs first. It stays on for a while. This duration can be determined by the developer. The application splash screen is shown in Figure 2.

Figure 2: Application Splash Screen

3.2. Registered Users Entry Screen

It is the second screen coming up. The users enter their user names and passwords. If an earlier session has been initiated, this screen is not displayed. The user name and the password is encrypted in JSON and is asked to the web service of the server defined in a POST query. Following the checking of user name and password by the web service, the "User ID", the "User Name" and the "User Name and Surname" are sent to the application. The application splits these data and forwards to the required configurations in which they are to be used. The registered users entry screen is shown in Figure 3.

Figure 3: Registered Users Entry Screen

3.3. Category Selection Screen

This section is divided into two as left and right. These sections are called "frames". In the right frame are the categories. The contents of the categories are in shown in the left frame. For instance; when the "Temperature" category on the right frame is selected, the temperature data is displayed in the left frame. When any of the categories in the right frame is selected, the application pulls the data from the web service and displays the required data on the left frame. The data coming to this section is transiently updated. For instance, the temperature values sensed by the sensor nodes will transiently be displayed in this section. As many categories as desired can be placed in the right frame. If the categories are represented in symbols, the visuality of the interface would be more attractive to the end user. More than one button can be added on the action bar in this section. These can be logging out and scenarios. When the log out button is hit, the session of the user is ended and the session opening screen is displayed. When the scenarios button is hit, the screen with scenarios is opened up. Category selection screen is shown in Figure 4.

Figure 4: Category Selection Screen

3.4. Screen for Scenarios

In order for the users to enter into the section for scenarios, they select scenarios in the action bar. As a result of the selection scenarios screen will be entered into. There is a spinner object in this screen (Figure 5). With this object, the scenarios can be viewed in a list. The data for this list extracted from the database. When w scenario is selected from the list, the list for selection is wiped off from the screen, as a result of the feature of the spinner, and the features of the scenario selected is brought up to the screen (Figure 6). For instance, with a scenario made up in the temperature category, the application can be instructed to send an SMS to the user if the temperature goes over a pre-defined value.

Figure 5: Screen for Scenarios 1

Figure 6: Screen for Scenarios 2

4. CONCLUSION

In this work, data from wireless sensor nodes was monitored on the Android platform. The Android application created in the methods mentioned can be run on all devices using Android operating system. As the application gets in connection with the web service, it does not get affected by the changes made in database. Therefore, when there is change in the database section of an application uploaded in the application market, the application can run without needing for an upgrade.

REFERENCES

Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E. 2002, "A survey on sensor networks", IEEE Communications Magazine, 40(8), 102-114.

Chong, C.Y., Kumar, S.P, 2003, "Sensor Networks : Evolution, opportunities, and challenges", Proc IEEE, Vol 91, No 8, 1247-1256.

DiMarzio, J.F. 2008, Android A Programmers Guide, McGraw Hill Professional.

Lin, R., Wang, Z., Sun, Y. 2004, "Energy Efficient Medium Access Control Protocols for Wireless Sensor Networks and Its State-of-Art", IEEE, pp 669-674.

New Age Technologies 2017, http://yenicagteknolojileri.com/portfolio-item/wsn-sensor-dugumu/.

Smartphone OS Market Share 2017, IDC [Online]. Available: http://www.idc.com/prodserv/smartphone-os-market-share.jsp.

Wang, y., Attebury, G., Ramamurthy, B. 2006, "A Survey Of Security Issues In Wireless Sensor Networks" , IEEE Communications Surveys & Tutorials, Volume 8, No. 2, 2nd Quarter.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

INTELLIGENT TRANSPORTATION SYSTEMS and TRAFFIC MANAGEMENT in WIRELESS SENSOR NETWORKS

DOI: 10.17261/Pressacademia.2017.571
PAP-WCTIE-V.5-2017(9)-p.52-57

Murat Dener[1]
[1]Gazi University, muratdener@gazi.edu.tr

ABSTRACT

Wireless Sensor Networks are the source of intelligent systems today. With the development of sensor nodes that form Wireless Sensor Networks, these systems can be easily used in our daily life. In this study, intelligent transportation systems and traffic management with Wireless Sensor Networks are described. Within the scope of intelligent transportation systems, vehicle sensors, road sensors, weather sensors and environmental sensors are mentioned. Within the scope of Traffic Management, dynamic signalization systems are mentioned. It is estimated that the work done will be beneficial to readers and practitioners.

Keywords: Wireless sensor networks, sensor node, zigbee, intelligent transportation systems, traffic management

1. INTRODUCTION

Sensors that detect the information such as temperature, humidity and light in the environment are called sensor nodes. Sensor nodes have the ability to calculate, detect, and communicate. These nodes perform the measurement task with cooperation. The networks that sensor nodes create are called Wireless Sensor Networks. Sensor nodes that form Wireless Sensor Networks are the source of intelligent systems. Two features of the sensor nodes are available. First, the desired sensor can be integrated into the sensor node. The second feature is that wirelessly transmitted data can be sent to remote stations. The system developed in this way can be monitored and controlled from time to place and from time to time. Because of all these features, sensor nodes constitute the infrastructure of the systems required for a smart city. In Libelium webpage (2017), authors state that systems that can be done with Wireless Sensor Networks are as follows: Smart Parking, Structural health, Noise Urban Maps, Smartphones Detection, Eletromagnetic Field Levels, Traffic Congestion, Smart Lighting, Waste management, Smart Roads, Forest Fire Detection, Air Pollution, Snow Level Monitoring, Earthquake Early Detection, Landslide and Avalanche Prevention, Potable water monitoring, Chemical leakage detection in rivers, Swimming pool remote measurement, Pollution levels in the sea, Smart Grid, Radiation Levels, Supply Chain Control, Smart Product Management, Quality of Shipment Conditions, M2M Applications, Smart Industrial Control, Smart Agriculture, Smart Animal Farming, Home Automation, Smart Health and more....

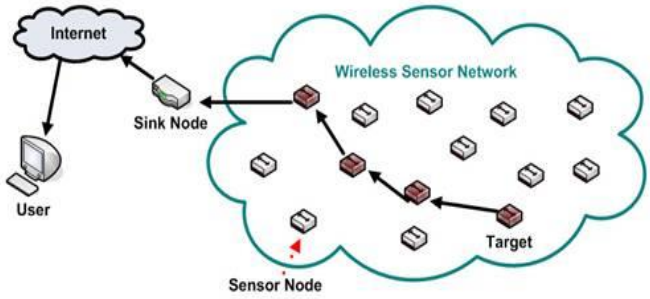
Systems developed with Wireless Sensor Networks can be prepared at half the cost of existing technologies. It is possible with these applications to increase the quality of life and raise the standard of living of people from every angle without distinguishing the lower, middle or upper part.

The rest of the article is as follows. In the second part, we talk about wireless sensor networks and ZigBee technology, in the third part, intelligent transportation systems that can be developed with Wireless Sensor Networks, and in the fourth chapter, traffic management that can be developed with Wireless Sensor Networks. The last part of the article contains the general conclusions of the study.

2. WIRELESS SENSOR NETWORKS

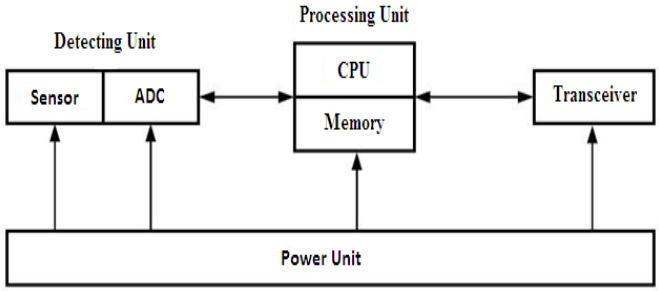
Ali et al. (2015) and Dener (2016) state that Wireless Sensor Networks consist of small-sized sensor nodes installed on the environment. These nodes carry, by collaborating in a physical ground, what they learn from the physical world to the virtual world platform. It is shown in figure 1.

Figure 1: Wireless Sensor Network



Challoo et al. (2012) state that wireless sensor nodes can use Wi-Fi, Bluetooth or ZigBee as wireless communication module. It is possible to choose according to the feature of the application to be made. While the distance of Bluetooth is low, ZigBee and Wi-Fi are more. The energy of the network established with Wi-Fi is exhausted after hours; the energy of the network established with Bluetooth after days and the energy of network established with ZigBee are exhausted after years. However, the network established with Wi-Fi and Bluetooth is more complex, while the network established with ZigBee is less complex. ZigBee's standard is IEEE 802.15.4, Bluetooth's standard is IEEE 802.15.1, and Wi-Fi's standard is IEE 802.11. ZigBee's data rate is 250 Kbits/s, Bluetooth's data rate is 723 Kbits/s, and Wi-Fi's data rate is 11 to 105 Mbits/s. Jun (2011) states that sensor nodes consist of, as shown in the figure 2, a detecting unit, a processing unit, a transceiver and a power unit. All constituents are explained below.

Figure 2: Constituents of Sensor Node



The main function of the detecting unit is to physically measure data in the detection and target zone. Khemapech et al. (2005) state that analogue voltage or signal is generated by the detector as a result of the event being monitored. Hill (2003) states that the continuous data is digitised by an ADC – Analogue to Digital Converter and then is sent to the processing unit for analysing. The processing unit plays a significant role in the management of the collaboration between the detectors in order to accomplish pre-described tasks. There are some families in this unit such as microcontrollers, microprocessors and field-programmable gate arrays – FPGAs. Interfaces such as Non-volatile Memory and ADCs can be integrated on to a single integrated circuit. Akyildiz et al. (2002) state that the processing unit needs a memory to run its tasks and minimizes the number of messages sent by means of local processing and data collection. There are three main communication channels in sensors; Optical Communication (Laser), Infrared and RF – Radio Frequency. Laser consumes less power compared to Radio and provides more security, however, requires line of sight and is sensitive to the atmospheric conditions. Feng et al. (2002) and Vieira et al. (2003) state that the infrared does not require an antenna as Laser does but broadcasting capacity is limited. RF is easy to use but an antenna is needed. The power consumption is a significant weakness in sensor networks. Any power saving arrangements could help to prolong the life of a sensor. Oliveira et al. (2011) state that the batteries used in sensors can be categorised in two groups; chargeable or non-chargeable. Generally in severe conditions, it is impossible to change or charge the battery.

3. INTELLIGENT TRANSPORTATION SYSTEMS

In extraordinary natural phenomena such as storms and earthquakes that occur suddenly, people must move from one point to another. Fuel consumption and vehicle emissions must be minimized and the loads must be delivered. The real-time vehicle and location information should provide the most advantageous journeys in terms of time and cost. Tufan (2014) states that the creation of such transport systems may seem a difficult target. In fact, accidents must be reduced to a minimum by warning drivers about any situation, such as weather conditions, traffic congestion, potential other hazards, etc. Establishing a transport system that allows maximum use of system capacity can be considered as a difficult goal to achieve. However, the use of technologies produced by today's scientific developments for a safe, efficient and sustainable transportation system in different fields is generally defined as Intelligent Transportation Systems. Strategy Development Presidency (2014) states that detection technologies used in intelligent transportation systems are mentioned below. These technologies include sensors that provide support for safe driving to the driver, and road and weather sensors.

3.1. Vehicle Sensors

Strategy Development Presidency (2014) states that sensors mounted on vehicles and providing safe driving support with functions such as lane detection, parking, blind spot warning, vehicle tracking distance warning.

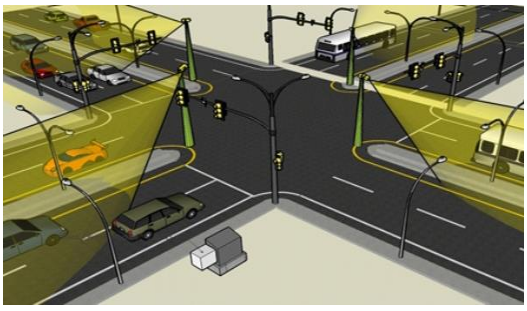
Figure 3: Vehicle Sensors (Mercedes-Benz 2017)

Thanks to these sensors, safe driving can be more. With the lane detection sensors, the vehicle will have increased its travel rate from its own lane. By creating an alert on lane changes, accidents will be reduced. Parking and blind spot warning sensors will prevent small accidents that may occur, and traffic congestion due to the park will be eliminated. Vehicle tracking distance sensors will also be used to maintain the officially designated tracking distance.

3.2. Road Sensors

Strategy Development Presidency (2014) states that it is the sensors that manage the intersections and adjust the signaling times, collect the data necessary to determine traffic flow information on the main arteries and peripheral roads throughout the city, and to direct traffic infrastructure for more efficient use.

Figure 4: Road Sensors (Memurlar 2017)



Nowadays, especially in big cities, the number of vehicles has increased considerably. Good coordination is necessary to ensure alignment in a region where there are thousands of vehicles. This coordination is primarily provided by traffic lights. However, the fact that the duration of the traffic lights at an intersection is fixed increases the traffic intensity. This means that the duration of the traffic lights must be dynamic for each intersection. In addition, this dynamism should be ensured at every hour of the day, and signaling times should be adjusted according to traffic intensity.

3.3. Weather Sensors

Strategy Development Presidency (2014) states that it is a powerful and versatile sensor that measures atmospheric conditions, live monitoring of ground temperatures, humidity and temperature.

Figure 5: Weather Sensors (PCE 2017)

Dynamically changing the maximum speed on the road by following atmospheric conditions such as rain and snow will both warn the drivers and reduce the rate of traffic accidents. In addition, when the temperatures of the ground are informed, it will be ensured that the drivers have already been informed and can take precautions in the event of icing or frost.

3.4. Environmental Sensing Systems

Strategy Development Presidency (2014) states that these are short and long range radar systems, imaging systems and software.

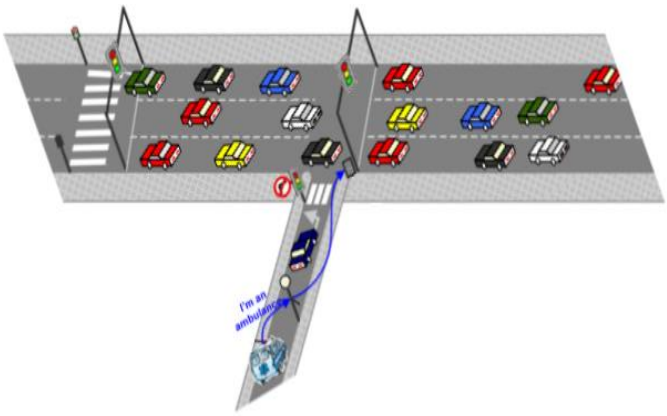
Figure 6: Environmental Sensing Systems (TETA 2017)

With the help of short and long range radar systems, vehicles passing the maximum speed within the region can be detected and automatic warning messages can be sent to these vehicles. Along with the imaging systems and software, however, the plates of vehicles which are in or out of a zone must be automatically detected and alerted to the necessary units in an emergency.

4. TRAFFIC MANAGEMENT

Traffic management, operation and supervision is the application of Intelligent Transportation System that provide efficiency in traffic, increase the quality of service and reduce traffic congestion. Strategy Development Presidency (2014) states that with these systems, it is aimed to inform drivers about dangerous situations, to take preventive measures in a timely manner, to give information about routes coming from dynamic traffic information systems and route going and to provide continuity in traffic flow.

Figure 7: Traffic Management (Selvarajah et al. 2012)



Effective traffic management ensures maximum comfort for people in traffic. Vehicle drivers do not expect much in traffic. In case of emergency such as ambulance, fire brigade, police car, gendarmerie in the traffic, it is also possible to set the transition priority of vehicles on the road.

5. CONCLUSION

Sensor nodes play an important role in an effective intelligent transport system and traffic management. In the vehicles, on the roads, with the help of the sensors in the environment, the information in the environment is gathered and the behaviors to be done together with the analysis software arise. The resulting information is sent back to the sensors, allowing the air route, speed and signaling times to be changed via the two-way communication. Thanks to effective management, traffic congestion and traffic accidents are reduced. This decrease also indirectly reduces air pollution. Increasing people's standard of living, facilitating people's lives with intelligent technologies can be considered as the advantages of intelligent transport systems.

REFERENCES

Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E. 2002, “Wireless Sensor Networks: A Survey,” Computer Networks, Vol 38, 393-422.

Ali, S., Qaisar, S.B., Saeed, H., Khan, M.F., Naeem, M., Anpalagan, A. 2015, "Network Challenges for CyberPhysical Systems with Tiny Wireless Devices: A Case Study on Reliable Pipeline Condition Monitoring", Sensors, Vol 15(4), 7172-7205.

Challoo, R., Oladeinde, A., Yilmazer, N., Ozcelik, S., Challoo, L. 2012, "An Overview and Assessment of Wireless Technologies and Coexistence of ZigBee, Bluetooth and Wi-Fi Devices", Procedia Computer Science, Vol 12, 386 – 391.

Dener, M. 2016, “A new gateway node for wireless sensor network applications”, Scientific Research and Essays, 11(20), 213-220.

Feng, J., Koushanfar, F., Potkonjak, M. 2002, “System-Architecture for Sensor Networks Issues, Alternatives and Directions,” ICCD’02.

Hill, J.L. 2003, “System Architecture for Wireless Sensor Network,” PhD Dissertation, University of California, Berkeley.

Jun, W.U. 2011, "Studies on Security and Protection for Wireless Sensor Networks", Doctoral Dissertation, Graduate School of Global Information and Telecommunication Studies, Waseda University.

Khemapech, I., Duncan, I., Miller, A. 2005, "A survey of wireless sensor networks technology," in In PGNET, Proceedings of the 6th Annual PostGraduate Symposium on the Convergence of Telecommunications Networking and Broadcasting.

Libelium 2017, 50 Sensor Applications for a Smarter World, http://www.libelium.com/resources/top_50_iot_sensor_applications_ranking/.

Memurlar 2017, Intelligent Transportation System, http://www.memurlar.net/haber/554242/.

Mercedes-Benz 2017, Vehicle Sensors http://t.mercedes-benz.com.tr/.

Oliveira, L.M.L., Rodrigues, J.P.C. 2011, "Wireless Sensor Networks: a Survey on Environmental Monitoring", Journal of Communications, Vol 6(2), 143-151.

PCE 2017, Weather Sensors, http://www.pce-cihazlari.com.tr/teknik-bilgiler/dokunmatik-ekranli-hava-istasyonu-pce-fws-20.htm.

Selvarajah, K., Arief, B., Tully, A., Blythe, P. 2012, "Deploying Wireless Sensor Devices in Intelligent Transportation System Applications", Intelligent Transportation Systems, Kent Academic Repository.

Strategy Development Presidency 2014, "National Intelligent Transportation Systems Strategy Document (2014-2023), Action Plan (2014-2016)", Republic of Turkey, Ministry of Transport, Maritime Affairs and Communications, Ankara.

TETA 2017, Recognition Systems, http://teta.com.tr/plaka-tanima-sistemleri.

Tufan, H. 2014, "Intelligent Transportation Systems Applications and an AUS Architecture Recommendation for Turkey", Republic of Turkey, Ministry of Transport, Maritime Affairs and Communications, Master thesis, Ankara.

Vieira, M.A.M, Coelho, C.N., Silva, D.C., Mata, J.M. 2003, "Survey on Wireless Sensor Network Devices," IEEE, Vol 1, 537-544.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

ASSESSMENT OF QUALITY REQUIREMENT AND IMPORTANCE FOR TEXTILE INDUSTRY IN TURKEY

DOI: 10.17261/Pressacademia.2017.572
PAP-WCTIE-V.5-2017(10)-p.58-66

Mehtap Erdil[1], Aysenur Erdil[2]
[1]Istanbul University, erdilmehtap@gmail.com
[2]Marmara-Yalova University, erdil.aysenur@gmail.com.tr

ABSTRACT

Today's competitory disposition of business obligates companies to implement an extensive business quality control system that highlights sustainable improvement of quality and productivity. In every sector, it is significant to support industries, meet the quality requirement of current's market by providing and improving the quality contol policy. Management development of work structures and frameworks should be involved in quality oriented production and quality control strategy. This topic should be supported by top management dedication to increase the quality level of production as a business policy and strategy within supportive resources. It is now generally accepted that textile is a fabric which made from fibers but the fibers may either be converted into yarn firstly and then the yarns put unitedly in one of a diversity of methods to make fabrics or the fibers can be altered directly into a fabric. The textile sector is selected as a research area in this study, because this is one of the most requirement and fundamental industries for households and populations. In this study, the integration steps of Quality Oriented manufacturing is hightlighted in the textile sector. An industrial perception is presented in the scope of this research. In addition of this, an overview of textile ındustry in Turkey will be presented and an approach for the importance of quality oriented production in garment industry is emphasized and is declared its strengths, opportunities, threats and weaknesses aspects according to the business quality policy.

Keywords: Consumption, industry, quality, production, swot, textile.

1. INTRODUCTION

Textile is declared that it includes clothing material such as fur, leather, suede and unsupported plastic sheeting. The first three items are natural materials and made form fibers although the fibers as fur can be detached from the skin and used as textile fibers. Leather and suede are fibrous in structure but the fibers have no separate identify in a textile sense and plastic sheeting has no fiber content at all. Fur, leather and suede can be simulated in textile structures and in combinations of textile and plastic materials (Potter and Corbman, 1967; Miller, 1995).

1.1. Textile Industry in the World in Turkey Particularly

From the middle of the century onwards, the textile industry has played a key role in initial industrialisation process in most countries. This is not surprising, given the combination of a large domestic demand (clothing being a basic necessity) and the early success at mechanising textile production. Textile industry of Turkey has the biggest share in Turkish foreign trade.

Turkey is the sixth largest producer of cotton yarn in the world and the producer of the world's finest cotton, Turkey was Europe's number one supplier of ready-wear clothes in 1995. This has a huge internal market with a population of 65 million having a potential purchasing power of five thousand and five hundred dollars per capita. The apparel market in Turkey is valued at 25 billion dollars, which accounts for 38 per cent of the country's total foreign trade and eight per cent of the gross national product (GNP), a highly attractive prospect for the global textile industry. Since the industry has the largest export/production ratio in the manufacturing industry, much importance is given to this industry in the seventh five-year development plan (Turkey in Statistics, 1997).

1.2. The Stages of Study

This research will be handled in three main stages. The first stage is literature survey. The sources are mainly gathered via technical reports and books. The second stage is an overview of textile industry in Turkey. The third stage is the analysis of the information gathered with Industrial overview. The base of literature survey begins with an overview of the concept of textile. The main steps of overview of textile industry in Turkey are as follows:

1-Current situation of textile and ready-made sectors in Turkey.

2-Textile and ready made sectors' trade relations with certain countries.

The existing and future importance of the clothing industry in Turkey's potential exports increases the essence of this research. Such a detailed study of the sector is a good example in the field.

2. OVERVIEW OF TEXTILE INDUSTRY IN TURKEY

Turkish textile sector whose past goes to 1920s, is developed in the leadership of government until 1950s. In the 1938 regenerated cellulose, after 1960s synthetic fiber prodcution is begun. Private sector entered cotton thread production in 1960s and synthetic thread production at 1970s. In the first years of Republic, Turkey was the fiber exporter, thread, fabric and ready-made importer. She began to be textile exporter after 1950s, ready-made exporter after 1970s (Gazanfer, 1995). In 1950s textile sector, and in the first years of 1970s ready-made sector began to develop according to the internal market demand and conditions. With their cheap raw material and labor advantages they began to have an important position in Turkey's economy in 1980s. Free trade and the adoption of the right trade policies resulted textile and ready-made sectors to the locomotive sector.

Historical Development of Commerce in Textile & Ready-Made Industries

-9.8 percent of Turkey's 1995 yearly total export income come from textile sector and 28.6 percent come from ready-made sector. The share of textile and ready-made sectors in Turkey's total export is 38.5 percent.

-Turkey is the second textile and ready-made supplier of European Community. According to 1994 statistics Turkey is the 8th supplier in textile and 15th supplier of ready-made of USA.

-Textile and ready-made production is 14 percent of total manufacturing industry, 13 percent of total industrial production and 10 percent of Turkey's total production.

-There are 123 textile and ready-made from among the 500 biggest industrial organisations in Turkey.

-As of September '96 there are 143 out of 800 foreign capitalised firms in textile and ready-made sectors (110 textile-made, 33 textile). Ready-made sector is the second production in industry branch that foreign capitalised firms are interested after food industry.

The strong position held by textile and ready-made sectors is not only a success of modern integrated factories and firms that can market their own trade-marks but also then thousands of small and middle sized firms. These firms with a labor force under 200 are generally located in İstanbul, İzmir, Denizli and Adana organised under Textile and Ready-Made Exporters Unions. There are seven textile and ready-made Exporters Unions in İstabul, Izmir, Denizli, Bursa, Antalya, Mersin, Gaziantep and they are bound to the Foreign Trade Councillor (Gazanfer, 1995; Istanbul ready- Annual Report, 1996; Annual Report, 1997). The development of textle and ready-made trade is also a result of the development of transportation and communication facilities in Turkey, the education on textile and ready-made and the growth of a generation that knows at least one foreign language.

2.1. Some Macro Economic Factors of Textile Sector

Textile and ready-made products has significant place in Turkish economy. The best indicator of this is the share of textile and ready-made sector in Gross National Income. Textile and ready-made has share of about 10 percent in Gross National Income (Turkish - Annual Report, 1997).

Table 1: Ratio of Textile and T-Ready-Made Industry in Total GNP (Url-6; Istanbul ready- Annual Report, 1996)

	1993 billion TL	1994 billion TL	1995 billion TL
Production of Textile & Ready-Made Industry	161,074	172,008	183,965
Total GNP	1,929,250	1,855,938	937,600
Ratio of Textile & Ready-Made Industry in Total GNP	8.3%	9.2%	9.4%

Textile and ready-made industry has also a vital share in total export of Turkey. This share is 36.6 percent in 1994 and 38.5 percent in 1995. According to the big share in Gross national ıncome and big ratio in total export, macro economic developments affect textile and ready-made industry significantly, similarly developments in textile and ready-made industry significantly, similarly developments in textile and ready-made industry affect total economy. So the performance of the industry is closely related with the macro economic policies.

2.2. Sector Development and Position According to the Production and Consumption of Countries

In Turkey textile ready-made industry developped based on cotton cultivation and artificial-synhetic fiber prodcution follows this. Cotton has always been a strategic raw material for Turkish economy. In the threshold of year 2000, the importance given total human health and environmental issue increases the value of textile prodcuts made of natural raw materials. For this reason, Turkey has an inherent advantage by being the sixth producer of cotton in world (Url1; Url 4; Basal and Sezener, 2012).

As seen in Table 2, while cotton production in the world has been around 26 million tonnes in recent years, In the 2015/16 season, it has decreased by 13% to 22.6 million tons compared to the previous year. In this hypothetical, Especially the decline in the production of countries such as China, USA and Pakistan has been effective. As it happens, that the largest cotton production in the world has been in China for many years, this situation has changed with the increase of cotton cultivation areas in India in recent years.

Table 2: Main Cotton Producing Countries- World cotton Production (Url1; Url4)

Country	2011/2012 (1000 tons)	2012/2013 (1000 tons)	2013/2014 (1000 tons)	2014/2015 (1000 tons)	2015/2016* (1000 tons)
India	6.001	6.095	6.770	6.510	6.240
China	7.400	7.300	6.929	6.480	5.260
USA	3.391	3.770	2.811	3.350	2.820
Pakistan	2.294	2.204	2.076	2.310	1.610
Brazil	1.884	1.261	1.705	1.550	1.550
Ozbechistan	880	1.000	940	890	860
Turkey	750	858	760	847	779
Australia	1.225	1.018	890	450	470
Others	3.459	3.332	3.402	3.543	3.051

Resource: ICAC Cotton (*) Forecast ICAC Turkey Country Report

At present, India is the largest producer with 6.2 million tons of cotton production amount- value. This country is followed by China, USA, Pakistan. The result of the decline in production in Australia in recent years, Turkey has risen to seventh place in world cotton production (Url 4; Basal and Sezener, 2012).

Table 3: Main Cotton Consuming Countries- World cotton consumption (Url2; Url 4)

Country	2011/2012 (1000 tons)	2012/2013 (1000 tons)	2013/2014 (1000 tons)	2014/2015 (1000 tons)	2015/2016* (1000 tons)
China	8.635	8.290	7.531	7.520	7.330
India	4.340	4.817	5.042	5.360	5.520
Pakistan	2.217	2.416	2.271	2.510	2.220
Turkey	1.495	1.350	1.400	1.486	1.500
Brazil	888	910	879	800	800
USA	718	762	773	780	780
Others	4.748	4.504	5.599	5.994	5.911

Resource: ICAC Cotton (*) Forecast ICAC Turkey Country Report

When Table 3 is examined, it is seen that the countries that consume the most cotton in the world are the countries that produce the most cotton again. Whether, world cotton consumption in the last 5 periods is in the range of 23-24 million tons, consumption in the 2015/16 season is estimated at 24.3 million tons. The biggest share of world cotton consumption is China, with 7.3 million tons (30% share). This country is followed by India and Pakistan. Turkey, on the other hand, is the fourth country that uses the most cotton with an estimated consumption value of 1.5 million tons. The decrease in sowing areas of cotton related to supporting policies and purchase price resulted fluctuation in cotton yield. Hence cotton supply is not enough (Debnath et al., 2016; Url 4; Table 4, State Institute of Statistics)

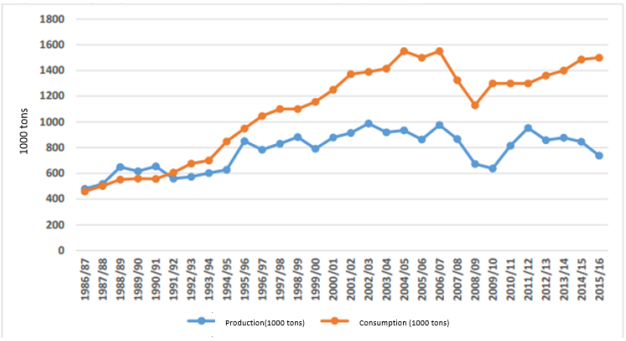
Table 4: Cotton Production & Consumption in Turkey (Url2; Url 4; State Institute of Statistics)

Years	Stock	Production (1000 Tons)	Consumption* (1000 Tons)	Difference	The ratio between Production and Consumption (%)
2011/2012	2.580	954	1.300	-346	73,3
2012/2013	2.320	858	1.360	-502	63
2013/2014	2.250	877	1.400	-523	62,6
2014/2015	2.350	846	1.486	-640	56,9
2015/2016*	2.050	738	1.500	-762	49,2

Resource: TUIK(Turkish Statistical Institute) (*) Forecast

According to TURKSTAT data, in our country in 2015/16 season, with the production of 2.050 tons of cotton stock, it is estimated that the amount of cotton for this amount is 738 thousand tons. According to the average of the last 5 production seasons in Table 4, country cotton production accounts for 60% of consumption. While This rate is 73% in 2011/12 season, because of decreasing the amount of production and increasing the amount of consumption, it has dropped to 49% in the 2015/16 season. For longer years, the trend of production and consumption has been shown in the graphic above.

Figure 1: The Trend of Cotton Production and Consumption in Turkey (Url 4; Url 5; State Institute of Statistics)



In this figure 1, The amount of cotton production of Turkey is presented. Until the mid-1990s, Cotton production and consumption in Turkey show an increasing course at a level close to each other, since 1991/92 season, the consumption of cotton started to overtake production, with the 1995/96 period, the production and consumption seems to have begun increasing levels (Url 4; Url 5). After the 2002/03 period cotton production generally followed declined trendline-cruising, especially in 2008/ 2009, the negative effects on Cotton prices of the global economic crisis, it caused a sharp decline in the country's production, After the crisis period, production reached pre-crisis levels, but in the last three periods, a recurring trend-course has shown (Url 4; Url 5). In the 30-year period, fiber cotton production in Turkey by 54%, Consumption has been increased by 226%. On the other hand, according to the 2002/03 season where production is close to 1 million tons, while production is reduced by 25% in 2015/16, consumption is seen to increase by 8% (Url 4; Url 5).

2.3. Textile Made of Cotton : Current Situation of Textile Sectors in Turkey

Figure 1 and Table 5 show that until the period of global economic crisis of Turkey's cotton consumption, the textile sector has been increasing continuously due to its growth-expansion (State Institute of Statistics, Url 4). “Supporting Premium System” (paying producer 60 cents premium for one kg production) decreased the cost of cotton and increased the cotton export. While there is cotton scarcity in 93/94 season, the 60 cents fund for cotton export paid in 94, draw-backed the increase of the scarcity. And the increase over world market prices. The fund for cotton export is abolished by law in december 1995. The export of cotton that is not carded and doffed is restricted by a quota 150.000 tons.

Table 5: Cotton Production & Consumption in Turkey (Periods of 1990-1996)

Years	Stock	Production (Ton)	Consumption (Ton)
1990/91	102813	654600	540000
1991/92	100697	561227	625000
1992/93	87406	573706	625000
1993/94	212347	602238	700000
1994/95	124219	628286	850000
1995/96	137572	851487	900000

Cotton thread sector, weaving and ready-made sectos were affected negatively between 93 december and 95 november. This situation hurt the competitiveness of Turkey in internaitonal market especially for “basic” products like T-shirts. The production of cotton thread is 517,600 tons in 1993; 585,000 tons in 1994 and expected 1995 prodcution is 610,000 tons which is enough to meet the demand (ITKIB- Istanbul ready, Annual report, 1996)

2.4. Artificial and Synthetic Fibers and Textile -Textile Made of Cotton

The history of artificial and synthetic fiber ans textile begins with the production of regenerate cellulose at 1938. After the huge capacity increases between 1984 and 1987, the investments are decreased until 1990. In 1994 capacity increased by 20 percent over the former year. It is expected that in the next several years fiber usage will increase by 3.5 percent yearly and polyester has the highest potential. At present, Turkey has the 9th biggest capacity of synthetic fiber and thread production in world. But in artificial-synthetic weaving area the supply does not meet the demand.

Wool production of Turkey is 4.2 percent of the world fiber prodcution and Turkey is one of leader countries in wool production. But the produced wool şs not thin enough. Because of the intends to the fiber that have wool features, comfortable and sport garments called “casualwear” and the need of care for wool, the price of wool decreased the demand of wool. But a research made shows that over 795 of Turk consumer knows “woolmark” etiquette and 81 percent trust it. It is expected that, production and consuming will increase according to the improvements in international wool market. There will be an inclination for wool and cellulose fibers mixture.

In the beginning of 1986, it is argued that whether Turkish textile and ready-made sector can show the same performance in the following years or not. Especially it’s argued that keeping investing on increasing capacity is beneficial for textile and ready-made sector or not. Turkish entrepreneur are criticised for being irrational in investment decisions in 12th International Textile Machinery Association (ITMA) fair at the end of 1995. Most of these arguments and criticisms may be true, but they are not scientific or they are not the results of searches. The searches about worldwide textile and ready-made products consumption and production tendencies are done by the Economist Intelligence Unit and published in Textile Outlook International periodical. These researches can help Turkish entrepreneurs because they try to estimate how the production and cınsumption balances would be in the following ten years. The researches of Textile Outlook International depend on the data of International Cotton Advisory Committee and the development in the consumption of ready-made products and investments o the textile machinery according to the development in the income level (Keane and Velde, 2008; Url4; Turkish Cl.As., Annual Report, 1997).

3. GENERAL PERSPECTIVE OF DEMAND FOR FIBER IN THE WORLD POPULATION

Textile and ready-made products meet primary or secondary needs of human kind. For instance, people use textile and ready-made products for prodcution against weather conditions, showing-off and decoration. Every individual conusmes textile and ready-made products according to the religion and tradition of the society. Textile and ready-made products such as T-shirt, shirt, sock, blanket, inside-wear and etc. contain amount of raw materials. Total fiber demand in the world was 29.9 million tons in 1980, but it reached to 37.5 million tons in 1992. This corresponds to 2.1 per cent increase per year. With his trend it is expected that demand will reach 51.5 million tons in 2004. The increase in the consumption of fiber is closely related with the increase in consumption of fiber in west Europe or developing countries.

Total consumption of fiber will increase as fiber consumption per person increases. Now we are faced with a question of “How can we meet this increasing demand?”. Undoubtedly supply must increase parallel with the increasing demand. A matter of fact, world fiber supply increases approximately 2.27 percent per year and this increase meets the total demand. Since countries’ total fiber consumption increases as result of increasing fiber consumption per person, thses countries have to meet the increasing demand by increasing the fiber supplies or importing fiber. This situation forces us to examine self capabilities of countries. If a country can meet the demand for thread and fiber from its own establishments and factories, this country is called self capable country (Url 4; Url 5; Keane and Velde, 2008; Sawatzki and Sirtioglu, 2016).

3.1. Encountered Problems in the Sector

The textile and ready-made sectors have new dynamics; they also meet with much more problems. These problems are summarized as follows:

-Excess amount over export is sold in internal market. This decreases the prices of internal market. So companies which sell their products into internal marlet are accomplished negatively.

-The excess amount of new firms causes price competition for foreign markets.

-The lack of energy and substructure cause many problems for the quality of textile industry.

-Since many firms in sector are small and middle sized, they are not financially strong. Because of this reason, banks do not give credit for these companies easily.

-There are rapid changes in the export markets. So many foreign importers do not buy in big lots, lot sizes are decreased. The companies that do not follow these rules are encountered with different problems.

-Bureaucracy and stationary are still among the significiant issues. Changes in the exchange rates influence the firms negatively that export amount of percent of their products (Keane and Velde, 2008; Dadashian et al., 2007; Hax and Majluf, 1991).

Textile and Ready-Made sectors are the applicants of dynamic sectors in short term.

-Marketing and delivery problems of small and middle size producers must be excluded by Sectorial Foreign Trade Firms.

-There is low risk of raw material insufficiency. Turkey is one of the few countries that has enough raw material capacity in the world.

-It is seen that being active in the marketing and delivery channels is a crucial necessity. For this reason, new investing on marketing and distribution channels in destination markets are increasing (Keane and Velde, 2008; Dadashian et al., 2007; Hax and Majluf, 1991).

4. OVERVIEW FOR EVALATION OF QUALITY REQUIREMENT

A Quality-Oriented Approach to the Assessment Quality Oriented Assessment supports the standards of organization and involves requirements, or instructions for an orderly peception to a specific facility. This evaluation may involve product design requirements, test methods, classifications, recommended practices, and other examinations. The advantages of this assesment are sorted like this. Firstly, it defines reliability requirements aimed to decrease the risk of production. It provides to build a level of performance for goods and products. The steps of this assessment are a structure for quality processes. This methodology tries to reduce cost and save money (Faridul et al., 2016; Keane and Velde, 2008).

Quality Control bases on these significant issues. A group of activities or methods whose objectives is to provide that all quality conditions are being considered via observing of processes and to construct performance issues through determining and testing. Components of Quality Control are to develop required quality goods and product, to provide customer satisfaction, to evaluate the demand of customer, to decrease the cost of production, to decrease the amount of waste and to support to earn maximum profit according to the minimum cost (Jeyaraj et al., 2012). Targets of testing causes for textile testing are checking the quality and convenience and reliability of raw material and selection of material of production, monitoring and observing of production stages, process control, evaluation of final-end user product, whether the quality is acceptable or refuse-rejected, and investigation of defective materials (for instance, assessment of customer grumbles, detection of failure modes in machine, research and development of products).

The purpose of study is to determine the best strategy and the improvement of the textile sector to evaluate this business and it will be analyzed via an engineering perspective. In this research scope, all steps of this industry will be involved. The other goal of this study is to define its strenghts and weaknesses, opportunity and threats components of this sector. Focused on the chosen significant criteria of the SWOT matrix is constructed.

4.1. Determination the Importance of Textile Industry via Swot Analysis

The SWOT variables are declared according to the prior experience of the researchers, experts of this sector and also withdraw support of brainstorming technique. The Swot analysis provides to make suggestion and decisions about problems in the framework of development of quality level of production in textile sector.

This application contributes to assess the problems generating real and actual business threats-risk and benefits conditions with quality problems and cumulative sustainability. The same way is applied as rating type questionnaire survey. All questionnaires are determined and assessed in qualitative nature. The questionnaires are presented with facilities reports of business and interview of experts. (Bernroider, 2002; Koo et al., 2008 and Hannah Koo et al., 2011) owing to all staff members in top management, managerial level and board level of the company (Gorsuch, 1983). They are consulted to evaluate and estimate the variables on Likert 5-point scale (1-Extremely unimportant, 2-Slightly unimportant, 3- Neither unimportant nor important, 4-Slightly important, 5- Extremely important). The figure 2 shows the strength, weakness, opportunity and threat criteria of the improvement the quality of textile industry of Turkey.

Figure 2: Evaluation the Quality Development of Textile Industry by Swot Analysis (Modified from Jeyaraj et al., 2012; Dadashian et al., 2007; Hax and Majluf, 1991; Keane and Velde, 2008 by author)

STRENGTHS	-Better brand vision -Export target in textile at USD 200 Billion by 2016. -Strong Research anad Development for steps of process of production -Low per capita consumption in this country (2.8 vs. global average of 6.8). -Cost competitiveness. -Support from the management for the managerial decision making -Increasing consciousness among customers and consumers -Decrease water consumption of manufacturing process - Decrease greenhouse gases (GHGs); -Modify to climate variability, extremes -Improve alternatives for insufficient appropriate resources (energy, materials) -Removal of quota restrictions to give a major boost. -Start with the ecological challenges which provides to exist future markets and declare areas where such challenges, trends support requirements in organizations	WEAKNESS	-Lack of regulations and guidance -Lack of technical experts and experience. -High cost of new spare parts and high cost of maintenance -Lack of ecological knowledge. -Fragmented Industry, -Effect of Historical Government Policies, -Technological Obsolescence -Expenditure of ecological friendly packaging -Slow speed of sample development
OPPORTUNITY	-Eliminate chemical, air, water pollution; health risks (wastemanagement, toxics) -Increased use of CAD to develop designing capabilities to increase the quality level. -Reduce Reuse Recycle – reduces cost -Low cost dyes and chemicals -Reduce Risks -Maximize reuse and minimize waste -Environmental challenges requires innovation and collective action by multiple parties across the value chain and beyond. -Effective environmental measures. -Ensure supply of freshwater (water quality, quantity) -Focusing on Product Development -Requirement of Mass production capacity -Investing in Trend Forecasting to enable the growth of industry -The growth opportunities exist in following areas: Medical textiles Construction textiles Packaging textiles, Home textiles, toilet and kitchen linen, Curtains, interior blinds, Furnishing articles	THREATS	-Increasing the carbon, water and energy footprints -Influence of external forces involved cannot be controlled -Government influence. -Lack of participation of top managment in acquiring green business, green production. -Competition in Domestic Market (Market competition) -High quality standard expected from international customers -Need to revamp Consumer Consciousness -High water consumption -rapid changes in the export markets -increasing amount of new firms causes price competition for foreign markets -Changes in the exchange rates influences the firms

5. EVALUATION AND CONCLUSION

This research is an evaluation of the quality level of textile industry. It is about the problems encountered the industry and the suggestions submitted during the research. Some suggestions are applied using the experience obtained. This study contributes the future studies on this topic. In the literature survey phase of the study, the main issue was the lack of publications about textile industry. Meanwhile, the available books were too old that they are far behind the new technological improvements. In the overview of textile industry phase, it is recognised that the data used in reports are not enough. These reports are prepared and used by authorised foundations such as ITKIB (İstanbul Tekstil ve Konfeksiyon Ihracatçılar Birliği), TCA (Turkish Clothing Association). Furthermore, the statistics in these reports are inconsistent. Since these reports are used for strategic decision making, this framework can cause to follow the wrong strategies. This conditions either can hide problems or cause inefficient usage, utilization of the capability.

Although Turkey has sufficient raw material and labor resources. But, this potential has not been used effectively. As Turkey already entered custom union of European Community, the textile industry gained more importance. In order to compete in international markets, Turkye's textile indusrty has to use its potential effectively and take necessary precautions. For a total effective structure, the firms should be related either in the form of financial associations and heartened by incentives to keep their flexible subcontracting configurations. This will lead to simultaneous coexistence of small and large prodcution units and will encourage the industry to organise itself at different production scales. In order to stay competitive and to surpas in this changing environment, the Turkish textile industry should develop its market policies by better design abilities, promote-advance the quality of the products and distribute into more demanding and exacting markets. Organisations of private sector and government should consider and work together for reconstitute programmes to increase activities and developments of fashion and design, marketing research and implementations of marketing, furtherance-publicity of country and brand image promotion.

Private non-profit organisations should be established to work on design intelligence, formal education, research and applications, Cad/Cam advance technologies, export promotion, employee and management training.

The textile industry is recognized for its wide and different range of products, different type and scale of firms and variety in the market. These factors lead to the authority of small and medium size firms which focus on comparatively few product classifications. This is a significant constraint, restricted factors on the process of technical change in the industry.

During this study, it is observed that production planning is the most important problem in textile industry. The structure of the production units in Turkey should be modified with the regulation of the government through an industrial investment incentive strategy. Quality control is also a crucial issue in textile industry. Since the last step of all processes is quality control, it is very significant stage. But inspection of detects in textile is subjective. Factories have risk when the final-end product is examined and delivered to the customers. Because customer satisfaction is a vital factor to sustain the textile industry. With the new developing technology, descriptions of fault and defect are made to the computer and computers are used for inspection. But this technology does not have the ability of making strategic decisions. Furthermore, this technology is no wide spread in Turkey. Thus, the concept of the quality of production should be expanded to the whole factory and textile industry and Total Quality Management (TQM) should be implemented.

Turkey, being a household textile development country, has been fortunate in sustaining its production groundwork despite different problems accomplished during the latest decades. The scope of textile production has been inauspiciously influenced by the government assessments, particularly during the last decades. Despite all these troubles and problems, Turkey has provided to sustain sixth in the world in cotton-textile production. The investments in textile manufacturing, which started off as cotton threads and cotton fabric production, and with significant expansion into other textile areas as well as to clothing sectors. This development supported for Turkey on the fifth order in the highest cotton consuming countries and sixth order in the clothing exporting countries. Although there occurs a well functioning scope interchange market for cotton, dating back to more than 100 years (Url 4).

REFERENCES

Basal, H., Sezener, V. 2012, Turkey Cotton Report, https://www.icac.org/wp-content/uploads/2012/11/03-HBASAL_TURKEY-COTTON-REPORT2.pdf

Bernroider, E. 2002, "Factors in SWOT analysis applied to micro, small-to-medium, and large software enterprises: An Austrian study", European Management Journal, vol. 20, no.5, pp.562 - 573.

Dadashian, F., Shakibfar, S. and Fazel Zarandi, M.H. 2007 "Strategic alliance for core competencies improvement in textile industries", International Journal of Management Science and Engineering Management, vol. 2, no. (2), 98-107.

Debnath, D., Thompson, W., Helmar, M. and Orman, T. 2016, Economics and Marketing, "Effect of the Southeastern Anatolia Project (GAP) on Cotton Markets of Turkey and the World", The Journal of Cotton Science vol.20, pp.46–55 (2016).

Faridul, H.K.M., Md. Shipan M.Md., Ashaduzzaman, Mostafizur, R.Md., Ullah, A., Ullah, M.S. 2016, "Role of Textile and Clothing Industries in the Growth and Development of Trade & Business Strategies of Bangladesh in the Global Economy", International Journal of Textile Science 2016, vol. 5, no.3, pp. 39-48 DOI: 10.5923/j.textile.20160503.01.

Gazanfer, S., Aegean Textiles and Raw Materials Exporters Unions Production and consumption of cotton in Turkey, Board Member

Gazanfer, S. "Assessing the establishment of a cotton futures market in Turkey" Paper presented at the Technical seminar at the 51st Plenary Meeting of the ICAC, Liverpool, United Kingdom

Gazanfer, S. 1995, "Guideliness for facilitating access to risk management markets through the stimulation of local and regionbal exchanges: the case of cotton in the Near East/CIS/Pakistan" UNCTAD/Com/65, 19 September, 1995.

Hax, A.C. and Majluf, N.S. 1991, "The strategy concept and process: A pragmatic approach", Prentice-Hall, London.

Istanbul ready made garment and apparel exporters' association (İstanbul Tekstil Konfeksiyon Ihracatçılar Birliği), Annual Report, 1996.

Istanbul ready made garment and apparel exporters' association (İstanbul Tekstil Konfeksiyon Ihracatçılar Birliği), Six Months Export Evaluation of Textile & Ready-Made Industry, Annual Report, 1997.

Jeyaraj, K.L., Muralidharan, C., Senthilvelan, T., Deshmukh, S.G. 2012, "Application of SWOT and Principal Component Analysis in a Textile Company - A Case Study", International Journal of Engineering Research and Development, vol. 1, no. 9 (June 2012), pp.46-54.

Koo, L.C., Koo, H., Luk, L. (2008), "A pragmatic and holistic approach to strategic formulation through adopting balanced scorecard, SWOT analysis and blue ocean strategy - A case study of a consumer product manufacturer in China", International Journal of Managerial and Financial Accounting, vol. 1, no.2, pp.127 - 146.

Koo, H., Chau, K.-Y., Koo, L.-C., Liu, S., Tsui, S.-C. 2011, "A structured SWOT approach to develop strategies for the government of Macau, SAR", Journal of Strategy and Management, vol. 4, no.1, pp.62 – 81.

Miller, E.1995, Properties and Behaviour in Clothing Use, London: B T Batsford Ltd.

Potter, D. and Corbman, B. P., 1967, Textiles Fiber to Fabric, New York: McGrawHill.

Sawatzki, K., Sirtioglu, I. 2016, Turkey Cotton and Products Annual Report, Turkey 3/23/2016, GAIN Report Number, TR6013, https://gain.fas.usda.gov/Recent%20GAIN%20Publications/Cotton%20and%20Products%20Annual_Ankara_Turkey_3-23-2016.pdf

State Institue of Statistics, Turkey in Statistics, 1997.

State Planning Organisation, Main Economic Indicators, 1998.

Turkish Clothing Association, Annual Report, 1997.

Url 1- Cotton production by country worldwide in 2015/2016 (in 1,000 metric tons) https://www.statista.com/statistics/263055/cotton-production-worldwide-by-top-countries/

Url 2-DataBank, http://databank.worldbank.org/data/home.aspx,

Url 3-Global cotton production volume from 1990 to 2016 (in 1,000 bales) https://www.statista.com/statistics/259392/cotton-production-worldwide-since-1990/,

Url 4-T.R. Customs and Trade Ministry General Directorate of Cooperative- T.C. Gümrük ve Ticaret Bakanligi Kooperatifcilik Genel Müdürlügü, 2015 yılı Pamuk, Feb. 2016, http://koop.gtb.gov.tr/data/56e95b3a1a79f5b210d9176f/2015%20Pamuk%20Raporu.pdf

Url 5-Türkiyede Pamuk üretimi ve Tüketimi, http://www.mkatextile.com/turkiye-pamuk-uretimi-ve-tuketimi.php.

Url 6- The World Bank, 2012. "Gross Domestic Product per Capita by Country in Current US$,

"http://data.worldbank.org/indicator/NY.GDP.PCAP.CD, son erişim tarihi: 23.12.2012.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

INVESTIGATING THE POTENTIAL OF DATA ANALYTICS SOLUTIONS IN MARITIME INDUSTRY

DOI: 10.17261/Pressacademia.2017.573
PAP-WCTIE-V.5-2017(11)-p.67-70

Omer Soner[1] , Metin Celik[2]
[1] Department of Maritime Transportation & Management Eng., Istanbul Technical University, Tuzla 34940, Istanbul, Turkey
[2] Department of Marine Eng., Istanbul Technical University, Tuzla 34940, Istanbul, Turkey

ABSTRACT

Recently, enhancing the data collection and utilization in maritime field is one of the key issue. At this stage, the maritime authorities seek for applicable solutions through improving the operational processes of key stakeholders such as ship operators, shipyards, offshore structures, ports & terminals, etc. This study reviews the existing data analytics solutions derived by research groups, classification societies, and technology providers. Considering the operational level integration requirements, the potential of the existing solutions in safety, efficiency, environmental sensitiveness and other performance indicators is determined. In conclusion, this study contributes to identification of the challenges and opportunities in order to extent data analytics applications to maritime field.

Keywords: Maritime industry, data analytics, operations management

1. INTRODUCTION

It's a common believe that data centred decision making is one of the leading factors that changing all aspects of the modern industries. Data analysis have been considered as the competitive differentiator, since an increase in output and productivity becomes more critical for many industries, where fierce competition takes place (Wang et al. 2015). New technologies have increased the capacity of capturing, storing, processing, presenting, and visualising of data. These data is used for many purposes, to illustrate; performance monitoring, condition-based and predictive maintenance, optimisation across systems, and system overview of operations (Løvoll & Kadal, 2014). A research shows that companies, who take advantage of the data analytic models, have improved their output and productivity rate about 5-6 % (Brynjolfsson et al. 2011). Therefore; it is not surprising that data analysis has already had a considerable impact on many industries, such as; maritime, aviation, health, and energy.

The maritime industry is subject to continuous demand in terms of cost, energy efficiency, environmentally friendly, timely delivery and economies of scale. Hence; International Maritime Organization (IMO) have introduced numerous conventions, and regulations to achieve these fundamental aims. According to International Convention for the Safety of Life at Sea (SOLAS); Voyage Data Recorder (VDR) is required to gathering specified data on board ship (SOLAS, 1974). In the environmental perspective; the MRV (Monitoring, Reporting, and Verification) regulation aims to quantify and reduce CO2 emissions from shipping (EU, 2014). Moreover, the e-navigation system is another striking development in the maritime industry that improves the collection and integration of shipboard data. The International Association of Lighthouse Authorities (IALA) is defined e-navigation as the unified collection, integration, exchange, and analysis of marine information on board and ashore to enhance safety and security at sea and protection of the marine environment (IALA, 2007). Even if the maritime authorities have imposed the data based improvement in all aspects of the industry, an inadequate number of research and project have been conducted within the maritime literature and maritime industry. Therefore; investigating the potential of data analytics solutions in maritime industry has a prominent important. The motivation of this paper is to review the data analytics related existing attempt in maritime domain. This paper organized as follows; motivation behind the study is presented in this section. In the second section, data centred studies and projects in maritime literature are investigated. Next, data analytics challenges and opportunities in maritime domain are fairly

discussed. In the last section, the potential of data analytics are underlined and possible solutions are suggested to overcome certain obstacles.

2. MARITIME DATA ANALYTICS

In this section, data analytics related studies and projects have been investigated. Data analytics applications and studies in maritime transportation can be divided following main areas; technical operation and maintenance, energy efficiency, safety performance, management and monitoring of accident and environmental risks from shipping traffic, commercial operation, and automation of ship operations (Løvoll & Kadal, 2014). In energy efficiency consideration; Baldi et al. (2014) have developed a methodology to analyse the energy flows of an entire ship with the goal of providing a better understanding of how energy is used on board ships. Perera and Mo (2016-a) have proposed data analysis techniques to understand marine engine operating regions as a part of the ship energy efficiency management plan (SEEMP). In order to improve ship energy efficiency under various emission control measures, SEEMP requires to collecting and analysing vessel performance and navigation data. Coraddu et al. (2017) have focused on the issue of fuel consumption prediction and trim optimisation of a vessel in real operations based on the data that gathered on board ship.

Regarding the operational efficiency related studies; Mak et al. (2015) have devoted a research for analysing the operational data for speed optimization in calm sea states and the development of a trim optimization software. Dobrkovic et al. (2015) have addressed the problem of "long term prediction" – the ability to estimate future position of a vessel at least 24 hours in advance for a specific region (the North Sea connecting the Netherlands and United Kingdom). Perera (2016) have developed the statistical filter based sensor and data acquisition (DAQ) fault detection for ship performance and navigation information. Perera and Mo (2016-b) have specified the marine engine operating regions under principal component analysis (PCA) to evaluate ship performance and navigation behaviour. Environmental aspects is another sub-topic that have been investigated, such as; Leonardi and Browne (2010) have proposed a method for the calculation of the maritime sector' carbon footprint, Perera and Mo (2016-c) have presented an overview of emission control based energy efficiency measures in the ship operation, and Bilgili and Celebi (2016) calculate the total emissions produced during a total 80,000 DWT load by two different bulk carrier ships.

Moreover; the FRAM Project is a voluntarily joint project that aimed to become more energy efficient and environmentally friendly by measure, report and verify CO2 emissions on ships. The Hull Risk Profiler is one of the maritime industry data centred applications. The Hull Risk Profiler enables to immediately assess the risk associated with an owner, manager, fleet or an individual vessel. On the other hand, the OstiaEdge Monitoring Suite enables improved use of on board data to help implement a condition-based maintenance. Certain Logistics firm are annually saving about the 100 million dollars since real-time analysis helps them to make contingency plans for different situations from route planning to general reductions in speed (Løvoll & Kadal, 2014). The AMINESS platform has aimed to integrate information from real-time vessel movements, weather data, traffic patterns, information on type and cargo of vessels, environmental data to suggest environmentally optimal safe route planning and deliver real-time alerts for ships (Giannakopoulos et al. 2014). Smart Ship Application Platform Project (SSAP Project) is one of the joint industry projects that consists of 27 member organizations. Main motivation of the project has utilized such application services to achieve optimum ship operation in terms of safety and energy efficiency (Ando, 2014). The studies that have been investigated clearly shows that in maritime domain there is still room for the expansion of the data centred applications and research. Moreover, data analytics have offered substantial opportunities that increase the maritime transportation' competitiveness. Therefore, identifying the challenges and opportunities have become a critical issue in order to extent data analytics applications in the maritime field.

3. CHALLENGES AND OPPURTUNIES

Most of the times, maritime transportation is considered as a conservative industry, so it takes a time to convince a traditional industry to fully integrate with the new data-driven age. However, required perspective should be able to understand this new age' challenges and opportunities. Therefore, maritime authorities and maritime researchers have investigated this issue. Koga (2015) have identified four major data analytics challenges, which are; sound competitive conditions, human resources, technology, and security. In the first category, the privacy issue, near monopoly of data providers and data governance have been investigated. At the second category, human resources with certain expertise for dealing with data analytics shortage is underlined. Next, technology perspective is analysed with respect to the availability of powerful tools, flexibility, data integration, correlation and causation, and collection biases. In the end, security perspective is investigated considering the unlawful control of device/machine and abusive insertion, update, and deletion of data. Rødseth et al. (2016) have investigated the obstacles to efficient collection of ship data on board. The study have emphasized cost and quality of on-board sensors and data acquisition systems, data ownership and technical obstacles to the effective collection and use of data on ships. Although the new technologies have made possible to collect all the data on board ship, these data are not readily available to analyse. Therefore; the owner or operator is likely to pay a high price for getting access to them. At this point, incompatible data formats for many types of data acquisition systems on a ship

makes difficult to observe and treat these data in a unified manner. E-navigation has offered a possible solution, which called “Common Maritime Data Structure” (IHO, 2010). On the other hand, high sea satellite communication still limits the ship-shore communication, while land-based communication systems offer relatively competitive low costs and high capacity. Even if the all challenges that maritime shareholders have encountered with the data analytics applications, it’s still have huge potential to improve the output and productivity in the maritime industy. Data analytics application can be used not just the real-time decision support systems, but also long term performance monitoring whether ship or fleet. Besides, maritime shareholders can also take advantage of the data analytic to increase the energy efficiency and reduce the environmental effects on the marine environment. Condition-based and predictive maintenance are also current issues that data analytics have great potential to deal with them.

5. CONCLUSION

The study have investigated the potential of data analytics in the maritime industry. In order to reveal the actual potential, the study have analysed challenges and opportunities that maritime shareholders have been encountered with data analytics application. Data acquisition problems are one of the fundamental challenges that maritime researchers constantly have to deal with. Even if the certain data is available on ship, accessing these data is required to strict regulatory development, and high cost infrastructure systems. Data analyst shortages is another considerable problems in the maritime industry. Securing the data and its provider, data management related problems have also generated essential restriction on the data analytics applications. However, data analytics have offered huge potential to increase the maritime competitiveness. Data analytics can be used to performance monitoring, condition-based and predictive maintenance, optimisation across systems, and system overview of operations in maritime industry. Therefore, data centred decision making related studies, and projects are urgently needed in the maritime industry.

REFERENCES

Ando, H. (2014). Smart Ship Application Platform Project (SSAP Project).

Baldi, F., Johnson, H., Gabrielii, C., & Andersson, K. (2014). Energy analysis of ship energy systems–the case of a chemical tanker. Energy Procedia, 61, 1732-1735.

Bilgili, L., & Celebi, U. B. (2016). Emission Routing in Maritime Transportation. In Energy, Transportation and Global Warming (pp. 837-849). Springer International Publishing.

Coraddu, A., Oneto, L., Baldi, F., & Anguita, D. (2017). Vessels fuel consumption forecast and trim optimisation: A data analytics perspective. Ocean Engineering, 130, 351-370.

Dobrkovic, A., Iacob, M. E., & van Hillegersberg, J. (2015, October). Using machine learning for unsupervised maritime waypoint discovery from streaming AIS data. In Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business (p. 16). ACM.

E. Brynjolfsson, L. M. Hitt, and H. H. Kim, "Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?" SSRN Electronic Journal, 2011.

EU, 2014. Proposal for a Regulation of the European Parliament and of the Council on the Monitoring, Reporting and Verification of Carbon Dioxide Emissions From Maritime Transport and Amending Regulation (EU) No. 525/2013–Political Agreement. Available at: 〈 http://register.consilium.europa.eu/doc/srv?l=EN&f=ST%2016238%202014%20INIT〉.

Giannakopoulos, T., Vetsikas, I. A., Koromila, I., Karkaletsis, V., & Perantonis, S. (2014, May). Aminess: a platform for environmentally safe shipping. In Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments (p. 45). ACM.

IALA. The IALA definition and vision for e-Navigation. E-NAV2-output 11 (March 2007)

IHO, S. (2010). 100, Universal Hydrographic Data Model. Robert Ward, Barrie Greenslad.

International Convention for the Safety of Life at Sea, 1974. IMO. (1974).

Koga, S. (2015). Major challenges and solutions for utilizing big data in the maritime industry.

Leonardi, J., & Browne, M. (2010). A method for assessing the carbon footprint of maritime freight transport: European case study and results. International Journal of Logistics: research and applications, 13(5), 349-358.

Løvoll, G., & Kadal, J. C. (2014). Big data-the new data reality and industry impact. DNV GL.

Mak, L., Seo, D. C., Kuczora, A., & Sullivan, M. (2015, May). Vessel Performance Analysis and Fuel Management. In ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering (pp. V011T12A052-V011T12A052). American Society of Mechanical Engineers.

Perera, L. P. (2016). Statistical Filter based Sensor and DAQ Fault Detection for Onboard Ship Performance and Navigation Monitoring Systems. IFAC-PapersOnLine, 49(23), 323-328.

Perera, L. P., & Mo, B. (2016-a). Data analysis on marine engine operating regions in relation to ship navigation. Ocean Engineering, 128, 163-172.

Perera, L. P., & Mo, B. (2016-b). Marine Engine Operating Regions under Principal Component Analysis to evaluate Ship Performance and Navigation Behavior. IFAC-PapersOnLine, 49(23), 512-517.

Perera, L. P., & Mo, B. (2016-c). Emission control based energy efficiency measures in ship operations. Applied Ocean Research, 60, 29-46.

Rødseth, Ø. J., Perera, L. P., & Mo, B. (2016). Big data in shipping-Challenges and opportunities. In Proceedings of the 15th International Conference on Computer and IT Applications in the Maritime Industries (COMPIT 2016), Lecce, Italy, May 2016, (361-373).

Wang, H., Osen, O. L., Li, G., Li, W., Dai, H. N., & Zeng, W. (2015, November). Big data and industrial internet of things for the maritime industry in northwestern Norway. In TENCON 2015-2015 IEEE Region 10 Conference (pp. 1-5). IEEE.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

FROM IDEATION TOWARDS INNOVATION: PILLARS OF FRONT-END IN NEW PRODUCT DEVELOPMENT

DOI: 10.17261/Pressacademia.2017.574
PAP-WCTIE-V.5-2017(12)-p.71-79

Onur Emre Kurt[1], Caglar Ucler[2], Ozalp Vayvay[3]
[1] Marmara University, onurkurt7529@gmail.com
[2] Ozyegin University, caglar.ucler@ozyegin.edu.tr
[3] Marmara University, ozalp@marmara.edu.tr

ABSTRACT

Companies need to deliver innovative new products in order to compete in the global market successfully. Customer preferences as well technologies are highly dynamic that continuous ideation is required in the fuzzy Front End of product development, where suitable ideas have to be selected to initiate object innovation. The fuzzy Front End is mainly unorganized in companies, and there is a lack of a consistent taxonomy in the literature covering all aspects in the fuzzy Front End for successful product innovation. Consequently a model describing the pillars of the fuzzy Front End with respect to object innovation is synthesized here based on evidence from the industry supported by a systematic literature review, which shall support practitioners and researchers in the successful evaluation of innovation ideas.

Keywords: Innovation, new product development, front end, fuzzy front end.

1. INTRODUCTION

Companies can reach distant markets in the global business environment; however, this also elaborates the level of competition. In addition, companies need to update their product portfolios to cope with the changing consumer preferences. Therefore change and continuous development are fundamental for companies, which is leading to innovation (Iyer et al. 2006), the widely accepted paradigm securing a growing market share (Gronlund et al., 2010; Bear, 2006; Andrew et al., 2010). Innovation is the introduction of something new; a new idea, method, or device: novelty (Innovation, 2017). The pragmatic as well the academic usage of innovation implies the generation of value to capture surplus (Jacobides et al., 2006). It delivers a competitive advantage in product/ service offerings (Tidd et al., 2005). Innovation shall be viewed as an end-to-end process within a value chain in interaction with resources and the environment (Hansen and Birkinshaw, 2007). Nagji and Tuff (2012) extended this concept by stating that smart firms use “total innovation” instead of depending on their future to collection of ad hoc efforts and apply a systematic approach. Thus innovative organizations are different in “strategy and vision, leadership, culture, processes, and physical work environment” (Stamm, 2003). According to OECD (2005) there are two main typologies, subject and object innovation. While subject innovation is focusing on business processes, object innovation focuses on new products/services. According the degree of the novelty, it is also designated incremental and radical innovation for minor, incremental improvements to radical changes (Tidd et al., 2005).

Koen (2004) described innovative product development process as composed of 3 phases; (i) Front End (FE), (ii) New Product Development (NPD) and (iii) Commercialization. Product innovation initiates from FE phase of NPD process, which covers idea generation and idea selection. FE covers the activities that precede structural New Product and Process Development (Koen et al., 2001), where project- /product definition is made (Achiche et al., 2013). According to Jou et al. (2016), FE is very important for the success of innovation projects, since the project is selected and then its quality, costs and schedule are defined. Sandmeier et al. (2004) implied that almost two thirds of the total cost of new product development is spent in FE phase of innovation process. Hirunyawipada et al. (2015) indicated further that when the idea is not understood deeply, lead time and cost are increased as a result of unnecessary iterations between FE and NPD phases.

The solution of design problems in NDP or commercialization phases cost much more time and money then in the FE (Thomke and Fujimoto, 2000).

Despite its importance, FE is less researched than other phases of product development. On the top of it, FE is fuzzy (Smith and Reinertsen, 1991). Fuzziness is simply the uncertainty. FE has fuzziness sources inborn, such as changing customer requirements or available technology for the time being. Zhang and Doll (2001) also indicated in (2001) the nature of competition and changing technology in this fuzzy set, and studied the consequences of fuzziness as well: the product is not perfectly defined. This is referred as the Fuzzy Front End (FFE) in the literature. According to Brem and Voigt (2009) consequently the FFE includes unknowable and uncontrollable factors. In addition to that, the FFE has generally an unorganized structure in companies. At this point, Koen et al. (2004) introduced "front end innovation" term instead of using "fuzzy front end" term, and created a framework for FE and called this framework as New Concept Development (NCD) model.

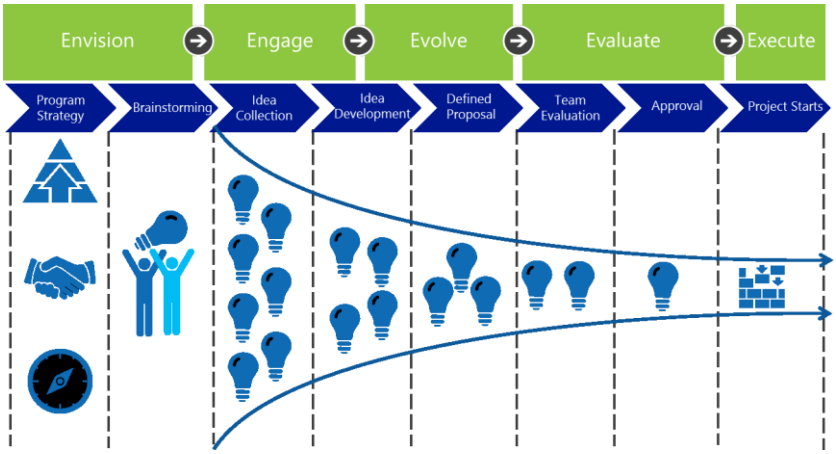
However; while only a few researchers are studying F-/FE, even less deal with the structure of it. Although importance of FE in product innovation is known, research about how to resolve the complexity in FE is limited. There is also a lack of a consistent taxonomy in the literature covering all aspects of the fuzzy extend in the FE. The selection of ideas is a vitally important, which needs a structured approach. Consequently FFE, selection criteria and best practices for the selection of NPD ideas have to be looked at in detail that related dimensions have to be isolated, which constitute the pillars of the FFE with respect to object innovation as introduced next.

2. REVIEW OF FRONT END METHODOLOGIES FOR INNOVATION

One might think that unstructured FE could avoid loss of innovation project ideas because of the formalization killing the creativity. However, CEN/TS (2013) explained that structured FE is an enabler of the innovation stream delivering a continuous basis for innovation projects. In companies, generally, expert knowledge and judgment are used in FE phase for decision making purpose because of fuzziness of this phase. If FE is not structured, especially idea evaluation capability of the company will depend on individuals and would not be consistent in long term. Innovation is a process itself and must be implemented throughout the organization. Morris (2011) described the internal innovation process as a systematic process, involving strategic thinking, portfolio management and research to internalize explicit and tacit knowledge. Therefore strategy, portfolio management, R&D, sales and marketing enable continuous growth, driven by the generation and selection of quality ideas in the FE. Consequently the FE tasks do cover product strategy formulation, opportunity identification, idea generation, product definition, product planning and reviews (Murphy and Kumar, 1997). Koen et al. (2001) classified this as opportunity identification/analysis, idea genesis/selection and concept/technology development. Deppe et al. (2002) divided FE into four stages as preparation for the idea generation, idea generation, idea screening and evaluation and the first concept, by designating the first concept as the beginning of the NPD. The FE model of Sandmeier et al. (2004) include market and technology opportunities phase, product and business ideas phase, and draft concept of product and business plan phase. The management of the project portfolio is important rather than the management of a single project for competitive advantage, and the ideation portfolio management in FE and project success do impact each other (Heising, 2012; Kock et al., 2015).

The first step in the FE process is the ideation. It is the generation, gathering, and assessing of ideas within a collaborative network. Only ideas with a high potential shall be selected for innovation success. According to Stevanovic et al. (2015) companies generally carry out the decision by using ad hoc or intuitively. However a managed and systematic approach is needed for game changing new product ideas (Cooper and Edgett, 2008), but the idea management processes are not generally managed in an organized and systematic way (Stevanovic et al., 2016). Successful ideas for NPD contain three dimensions which are novelty for customer, usefulness for customer and usefulness for companies (Hirunyawipada et al., 2015). There are two main approaches to structure the FE: the New Concept Development (NCD) model (Koen et al., 2001; Koen, 2004) further fuzzified by Achiche et al. (2013) and the Stage-Gate Model (Cooper and Edgett, 2008). The research of Koen et al. (2001) revealed that organizational attributes including senior management involvement, vision, strategy, resources and culture are the most important factors for FE performance. In the Stage-Gate model (see Figure 1) ideas are simply put through a funnel with gates after certain stages letting only successful ideas pass the designated gates. However, the requirements for the gates are not defined and kept as company specific (Riel et al., 2013). These models both apply subject innovation to reach object innovation, and consequently innovation management is applied therewith.

Figure 1: Innovation Management Model (Microsoft, 2013)



To isolate requirements the FE elements have to be understood. Chamakiotis et al. (2016) tried to isolate unique characteristics of virtual project teams with respect to collaboration in the FFE. Jou and Yuan (2016) proposed a novel method, combination of Crawford and Di Bendetto's model and Cooper's model to support the decision making in FE, and with this method they divided FE phase into four stages which are; market exploration and technology forecasting, idea generation and segmentation, portfolio analysis, and technology road mapping. Gronlund et al. (2010) proposed to combine the open innovation activities and Stage-Gate Model and called it Open Stage-Gate Model, to enable externalization. Ahmed-Kristensen and Daalhuizen (2015) proposed that the integrated usage of agile and stage-gate models as well. Sommer et al (2015) created a hybrid model, a combination of Stage-Gate and Agile Models, but only initial idea development was covered. All in one there is research indicating the importance of ideation and idea selection in the FE. Some tools and techniques for the selection of these ideas are introduced embedded in methodologies, but there is no generalized taxonomy classifying/clustering the required aspects, which have to be respected during the idea selection process. Available literature represents a wide spread around several methodologies, but is not structured. Consequently, to remedy the complexity a harmonization shall be applied in order to deliver a set of important aspects to be respected in the FE to achieve a potential innovation outcome. In other words; the pillars of FE in NPD are to be isolated. There is enough evidence and data in the literature, which has to be synthesized to a model as discussed next.

3. DATA AND METHODOLOGY

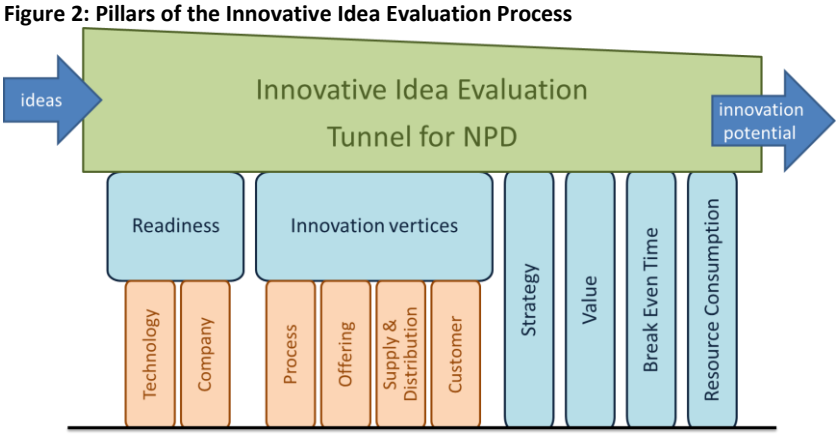
This study originated from the need to select correct ideas for the innovation stream. Thus the main objective of this paper is to quantify associated dimensions affecting on the innovation output. Therefore a systematic literature review has been accomplished by using key words as [innovation; fuzzy front end; idea selection; new product development] in scientific databases including but not limited to EBSCO, Scopus and WOS Resources. This literature review ranges from surveys to papers and dissertations. A total of 58 papers, 1 graduate thesis, 6 books and 1 survey have been considered as relevant based on the evidence from industrial consulting expertise of the authors. Data gathered thereof is used to synthesize a model describing the pillars of fuzzy front-end of NPD.

4. FINDINGS AND DISCUSSIONS

Innovation ideas shall be collected frequently, evaluated and then executed (Stamm, 2003). More generally such ideas are to be bred, evolved and finally selected in FE phase, which consequently shall go through NPD to become an innovation output via successful commercialization. Consequently FE shall respect during concurrent ideation and evaluation all aspects of the product life cycle to avoid costly iterations. However; limited research is available for FE, and there is a gap in the literature for models' selection criteria. Despite its criticality, the idea selection phase generally depends on the area experts' subjective assessments or decisions according to their knowledge (Bear, 2006). While subjective assessment can lead towards failure; objective assessment can lead to unbiased judgement and create more opportunities. Scholars proposed to structure innovative NPD by using range from lean, agile models etc. alone to in combination of these models with the NCD and Stage-Gate models. Furthermore, idea selection is generally investigated under idealization, project portfolio management, inter-firm collaboration, inter-personnel relationship, open innovation, Agile/ Stage Gate Hybrid methods; however it is generally mentioned as a word. Barely some researches investigate selection process alone. There is no generalized taxonomy with respect to the classification of required aspects for idea selection process. Moreover, despite available methodologies in the literature, the structuration of ideation and idea selection were found to be a weak link in

the innovation chain. In order to remedy the complexity of FE, mainly because of ideation, pillars of FE are isolated and a model in which these respected pillars synthesized is applied to secure harmonization in the FE phase. In this study, Cooper`s Stage-Gate Model was selected as a structure or backbone. Thus subject innovation was applied to reach the object innovation by ideas driven through gates, when meeting the criteria supported by the pillars. Consequently instead of the idea funnel approach, which is widely used in the literature, an evaluation tunnel is described (see Figure 2), where the respected pillars are supporting the idea evaluation model.

Figure 2: Pillars of the Innovative Idea Evaluation Process



This has a specific meaning. While during the funnel approach the ideas are simply passing through the funnel, or are eliminated, in this model the ideas have to be driven through the evaluation tunnel, i.e. they have to be further evolved. Ideas with correct justifications, but wrong development direction are not simply dropped out. They are developed to achieve successful innovation. This is an extension of the model of Stevanovic et al. (2015), by which idea efficacy of innovation was evaluated by means of technical, customer, market, financial and social attributes. According to preliminary research, pillars of the idea evaluation were consolidated to six groups as (i) Readiness, (ii) Innovation Vertices, (iii) Strategy, (iv) Value, (v) Break Even Time and (vi) Resource Consumption, which will be looked at in detail next. An idea might be the brightest, but cannot be converted to innovation successfully, when the product realization is not possible due to missing technology or capabilities, i.e. it shall be determined whether the innovation idea unveils a fit with competencies (Salomo et al., 2007). Thus maybe the most important pillar is the Readiness, which consequently covers *Technology* and *Company* sub-pillars. These sub-pillars have an interaction with each other. Tidd et al (2005) indicated that when technology interacts with both the market and the company, technology related innovation projects can be achieved. The competitiveness of especially high tech companies shall and can be enhanced by successful management of technology (Kropsu et al., 2009). Therefore peripheral vision, disruptive technologies, patent mapping and idea capture internally are supporting innovation (Cooper and Edgett, 2008). In order to target and reach innovation, the nature of the idea has to be understood thoroughly. It might be that there are no breakthrough ideas, but selected ideas shall support innovation in one or another way. Tsekouras et al. indicated in (2014) four vectors as offering, process, distribution, and customer, which were also mentioned by Sawhney (2006). In consequence, the Innovation Vertices were selected as Process, Offering, Supply & Distribution and Customer, wherefore the innovation contribution shall be questioned during the idea evaluation as explained further below.

On the one hand, innovative organizations differ in their *Processes* (Stamm, 2003), which describe how companies are working (Cronemyr and Danielsson, 2013) by defining activities with inputs and outputs for specific goals (Cao et al., 2013). Koen et al. (2001) indicated that creating more innovation starts with organizational attributes rather than team performance, opportunity identification or ideation. Problems arising during product realization shall leverage the redesign of processes (Tenner and DeToro, 1996). Likewise the preventive reengineering of processes by utilizing new approaches or technologies also leads to subject, i.e. process innovation (Tidd et al., 2005; Eliens, 2015; OECD, 2005). This is tightly related to knowledge management (Cao et al., 2013). Eliens (2015) indicated that process and distribution innovation as in (OECD, 2005) are not enough researched in literature. Considering that innovation can be achieved by the iteration and circulation of shared knowledge (Nonaka et al., 2008) and that process innovation involves knowledge management, the importance of the process pillar of innovation is further underlined. On the other hand, the object innovation is focused on in the *Offering* sub-pillar. Here the offering shall be evaluated with respect to its contribution to the object innovation. Both, products and services are covered in there (CEN/TS, 2013; Tsekouras et al., 2014). The way how the offering is organized, as well the products/services themselves are the focal point of this sub-pillar. The organization of the offering can be e.g. platform type, where common architecture of design is done and different types of product are developed on this common architecture (Koen, 2004). However, when transforming the voice of customer (VOC), the products might become

standardized and the contribution to innovation might be reduced. Thus market research has to be done continuously (Stamm, 2003). Therefore the ideation shall be capable to look beyond the market of today and shall bring a strategic insight as well, by tracking the change today. Nevertheless Kock et al. implied in (2016) that ideation is interdepended from market environment turbulence. Consequently short time trends shall not sway the evaluation, but riskiness, portfolio size and interdepended projects have to be examined under this pillar by applying consistent portfolio management (Morris, 2011), to evaluate ideas with respect to their contribution to innovation potential as products.

The product realization is accomplished within integrated supply chains to deliver competitive advantage (Ursino, 2015), where exchange of tacit knowledge and mutual knowledge generation may not be well-preserved, which are prerequisites of innovation (Casanueva et al., 2013). However; if object innovation is supported by concrete not only internal but also external linkages, value created by can be further increased. This dilemma is described in the literature as open innovation (Chesbrough, 2006) in the collaborative virtual product realization environment (Ucler and Kristensen, 2016). Consequently this pillar is simply named here as the *Supply & Distribution* in a generic manner, emphasizing that supply and distribution do contribute to the Innovation network (Tidd et al., 2005). Fuzziness can be decreased by supplier integrated FE (Wovak et al., 2016), where iterative learning cycles among integrated product development stakeholders increase the efficiency and effectiveness of the FE (Sandmeier et al., 2004). Temporary virtual project teams may have unique characteristics over FE activities (Chamakiotis et al., 2016), thus the model shall be kept as simple as possible in a collaborative and lean manner (Tuli and Shankar, 2015). Moreover, exporting and importing know-how may be inter-organizational as in open innovation, but also across various departments. Thus the internal linkages in an organization are also important in innovation (Tidd et al., 2005; Cornetto et al., 2016), where the balance between inter-firm and intra-firm structure types in FE phase should be managed properly to succeed in the market (Jörgensen et al., 2011). Innovation can be delivered by how supply and distribution is made (Tsekouras et al., 2014). Furthermore partners, vendors, external scientific/ technical communities, external finished product designs and external ideas can contribute to the open innovation (Cooper and Edgett, 2008). Thus importing external know-how and exporting internal know-how reduce cost and development (Gronlund, 2010) for the benefit of the company and the customer, which shall be evaluated here.

One aim of innovation is creating differentiated customer experience or a new clientele that *Customer* is selected as another sub-pillar. If early involvement of the customer is gained, product type or service needs can be understood in the early phases in the FE (Khurana and Rosendhal, 1998; Eliens, 2015). Stamm studied in (2003) the confliction between customer preference and the preference to be offered, furthermore, concluded that misinterpretations can be occurred while transforming the VOC. Moreover Cooper and Edgett (2008) include the VOC as an important idea source for innovation. But the customer vertex includes much more than the integration of the customer experience in the FE. According to Tsekouras et al. (2014) the customer as in the fourth vector of business innovation does represent new value capture through new revenue streams within new segments as well. So innovation ideas have also to be questioned with respect to new customers and the revenue stream in short term and long term.

Strategy is a key driver for product innovation (Stevanovic et al., 2016). Morris implied in (2011) that strategy and innovation form each other. Eliens (2015) and Khurana and Rosendhal (1998) indicated that strategy pillar implies the coherence between product innovation project and firm's business strategy, where internal and external environment shall be respected (Tidd et al., 2005). First the innovation strategy of the organization shall be defined (Cassiman and Veugelers, 2006) and then the ideas shall be selected according to the coverage of their business plans to the corporate strategy (Salomo et al., 2007). All strategies of the company shall be covered by the selected ideas thus the portfolio management is vitally important. Kock et al. (2015) studied the effects of ideation portfolio management on FE success, and created a research framework in which ideation portfolio management is divided into three elements which are ideation strategy, process formulization and creative encouragement, concluding that there are significant effects between these elements regarding front end success and eventually project portfolio success.

Today "high-level science and engineering are no more important than the ability to use them in mid- and ground-level innovations" (Bhide, 2009). Within the Value vertex the innovation idea shall be tested for main value drivers (Salomo et al., 2007). Therefore the practical application and the value generation shall be focused on, not how good the idea is. Especially added value can be achieved easily by incremental innovation, while other innovation types might generate totally new horizons, but imply risks. Value creation can be in different areas such as by product innovation or by business model innovation (Bhidé, A. 2009; Amit and Zott, 2012). Jacobides et al. (2006) implied that innovation value capture surplus. Hansen and Birkinshaw, (2007) studies the innovation value chain in interaction, and defined the innovation value chain as idea generation, conversion and the diffusion. Sometimes the users value of IT innovations cannot be captured with traditional accounting (Grant et al., 2013) that Porter's value chain analysis for differentiation can be used for innovation assessment as well (Ucler and Gok, 2015).

The assessment for the Break Even Time is a very important consideration. Easy to realize innovation with an early return may be preferred to increase innovation volume. Furthermore the appropriate product introduction time within the

innovation portfolio and thus the schedule can be determined therewith (Calantone et al., 2014; Palmberg 2006). Park et al. (2016) studied the break even time analysis in NPD, and concluded that depending on businesses' cost structure and innovation speed, there is a difference in break even time of industries. Nevertheless companies shall benchmark in their segments and evaluate business cases of innovation ideas accordingly. Finally the Resource Consumption is the last important vertex. NPD projects for innovation might require many resources individually, but collectively they can consume a large amount of resources, thus a balanced portfolio management is a must (Cooper, 2003). The allocation can be adjusted (Loch and Kavadias, 2002) to have a spread over certain projects, but shall not jeopardize existing value streams. The sustainability context may be included there by sustainable innovation as well (Wever and Boks, 2007; Dewulf, 2013), but shall be better included throughout the whole evaluation process.

All in one these pillars span the solution space of the questioning towards innovation. Instead of using business specific evaluation pillars, generic idea evaluation pillars are generated to provide a general overview to the problem. These pillars might have weights associated, depending on the strategic objectives of the organization. Moreover depending on the business type minor adjustments can be done to increase the effectiveness of the innovation project, but the delivered taxonomy based on these pillars does deliver a robust evaluation respecting all aspects.

5. CONCLUSION

Although the importance of FE in product innovation is known, research about solving the complexity of FE is in progress. Moreover the fuzziness of the FE is increasing this complexity via rapidly changing variables in market conditions, technology and customer requirements. Nevertheless; a product has to successfully address the requirements by transforming technology under available market constraints. This is not an easy task; especially it is not easy to foresee the success in the FE phase. On the top of this, companies have to innovate and deliver added value to capture surplus. This automatically implies that new technologies are adopted and new ideas are transformed into products. This builds up the unknown further and the question remains open, which ideas are the best candidates for NPD, i.e. which ideas do provide the best innovation potential, but do streamline with the organizations' strategy and capabilities. Ideation is the exploratory part of FE, where NPD ideas are generated and then evaluated. The evaluated ideas, products in spe, play an indispensable role in the success of the company in the market. When the idea generation relies on persons, there is no collaboration. The outcome of such an unstructured approach is arbitrary evaluation, not complying with the innovation goals of the company. Furthermore the lack of a system for ideation, as well for evaluation and tracking does risk the continuity of the innovation. As a result the consequent selection after ideation in FFE is focused on in this study, in order to harmonize the literature and provide a model synthesized by the aspects of selection, which have to be respected in the FFE of NPD.

The New Concept Development Model and the Stage-Gate Model are popular for structured ideation and consequent evaluation throughout NPD, allowing a consistent flow of ideas through a funnel, where innovation contributors can be selected. These can be converted then into new products. In order to eliminate subjective assessments and especially to enable a multi-dimensional approach for the evaluation, the pillars of the FE were isolated, which have to be respected during evaluations. Criteria from the literature were consolidated to six groups of (i) readiness, (ii) innovation vertices, (iii) strategy, (iv) value, (v) break-even time, and (vi) resource consumption, which are proposed to be used in combination with a modified Stage-Gate Model, where ideas are driven that their evolvement is supported as well. Thereof the generic idea evaluation pillars do not depend on the nature of the business case. When after ideation these pillars are respected in the evaluation, the structured coverage eases innovation outcome by streamlining ideas, strategy and capabilities. The FE is inborn fuzzy because of the given change, but this objective, structured approach leads the way to manage the fuzziness as well.

This work contributes to the literature first by reviewing and merging FFE and idea selection related innovation literature and practices, which were found to be unorganized with a divergent spread. Then taxonomy is delivered within. Thus the practical implication is that NPD and innovation professionals can use this delivered model to cover all required aspects in their evaluations. Moreover, to attain continuous innovation, companies need to have a quality idea evaluation process, which basics are supported herewith, since this work combines many sources for idea evaluation. This work is of course of general nature and despite guiding the practical implementation it is reflecting a research in progress. Thus a robust model utilizing these pillars will be constituted within future research, minimizing the intrapersonal interactions by guiding the evaluators.

ACKNOWLEDGEMENT

The authors would like to thank to the *Unit of Scientific Research Projects of Marmara University* (BAPKO) for supporting this research numbered as FEN-D-250608-0182.

REFERENCES

Achiche S., Appio F.P., McAloone T.C., and Minin A.D. (2013), "Fuzzy decision support for tools selection in the core front end activities of new product development", *Research in Engineering Design*, vol.24, iss.1, pp.1-18.

Ahmed-Kristensen S. and Daalhuizen J. (2015), "Pioneering the combined use of agile and stage-gate models in new product development: cases from the manufacturing industry", *in proceedings of the 22nd Innovation & Product Development Management Conference*. Copenhagen, Denmark.

Amit R., Zott C., (2012) "Creating Value Through Business Model Innovation", MIT Sloan Management Review, Spring 2012, vol.53, no.3

Andrew J.P., Manget J., Michael D.C., Taylor A., and Zblit H. (2010), "A return to prominence – and the Emergence of a New World Order", April 2010, The Boston Consulting Group.

Bear, D.J. (2006), "The Quest for Innovation: A Global Study of Innovation Management, 2006-2016", American Management Association, [online] http://www.amanet.org/images/hri_innovation.pdf

Bhidé, A. (2009), "Where innovation creates value", *McKinsey Quarterly*, Vol. February 2009, [online] http://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/where-innovation-creates-value

Brem A. and Voigt K.I. (2009), "Integration of market pull and technology push in the corporate front end and innovation management: Insights from the German software industry", *Technovation*, vol.29, iss.5, pp. 351–367.

Calantone, R. J., Randhawa, P. and Voorhees, C. M. (2014), "Breakeven Time on New Product Launches: An Investigation of the Drivers and Impact on Firm Performance", Journal of Product Innovation Management, vol.31, pp.94–104.

Cao, Q., Thompson M.A. and Triche, J. (2013), "Investigating the role of business processes and knowledge management systems on performance: A multi-case study approach", *International Journal of Production Research*, vol.51, iss.18, pp. 5565–5575.

Casanueva, C., Castro , I. and Galán, J.L. (2013), "Informational networks and innovation in mature industrial clusters", *Journal of Business Research*, Vol. 66, Iss. 5, pp. 603–613.

Cassiman, B. and Veugelers, R. (2006), "In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition", Management Science, vol.52, iss 1, pp.68-82.

CEN/TS (2013), *CEN/TC 389 Innovation Management, Part 1: Innovation Management System*. Technical Specification (TS) of the European Committee for Standardization (CEN), Nr. CEN/TS 16555-1:2013.

Chamakiotis, P., Boukis, A. and Panteli, N. (2016), "The fuzzy front-end of innovation in temporary virtual project teams", *in Proceedings of the 10th Mediterranean Conference on Information Systems (MCIS 2016)*, 4-6 September, Paphos, Cyprus.

Chesbrough, H.W., *Open innovation: The new imperative for creating and profiting from technology.* Harvard Business Press:. Boston, Massachusetts, 2006.

Cooper R. and Edgett S. (2008), "Ideation for Product Innovation: What are the best methods?", *PDMA Visions Magazine*, wp#29, vol.1, iss.1, pp.12-17.

Cooper, R.G. (2003), "Your NPD portfolio may be harmful to your business's health", *Research Technology Management*, vol.47, iss.1, pp.31-43.

Cornetto T., Nisar A., Palacios M., Meunier-FitzHugh K., and Labadie J.G. (2016), "Organizational linkages for new product development: Implementation of innovation projects", *Journal of Business Research*, vol.69, pp.2093-2100.

Cronemyr, P. and Danielsson, M. (2013), "Process Management 1-2-3 - a maturity model and diagnostics tool", *Total Quality Management & Business Excellence*, vol.24, iss. 7-8, pp. 933–944.

Deppe L., Kohn S., Paoletti F., and Levermann A., (2002), "The holistic view of the front end of innovation", In*: Proceeding of IMTs and New Product Development Conference*, Mantova, Italy, 17-18 October 2002.

Dewulf, K.R. (2013), "Sustainable product innovation: the importance of the front-end stage in the innovation process." Chapter 7 in *Advances in Industrial Design Engineering*, edited by DA Coelho. InTech, 2013.

Eliens L. (2015), *Disentagling the fuzzy front-end: an integrative literature review.* Dissertation, Faculdade de Engenharia da Universidade do Porto. [online] https://repositorio-aberto.up.pt/bitstream/10216/79627/2/35840.pdf

Grant, K., Alefantos, T., Meyer, M. & Edgar D. (2013) "Capturing and measuring technology based service innovation – A case analysis within theory and practice", *International Journal of Information Management*, Vol. 33, pp. 899– 905.

Gronlund J., Sjodin D.R. and Frishammar J. (2010), "Open Innovation and the Stage-Gate Process: A revised Model for New Product Development", *California Management Review*, vol.52, no.3, spring 2010.

Hansen, M. T. and Birkinshaw, J. (2007), "The innovation value chain", *Harvard Business Review*, vol.85, iss.6, 121.

Heising W. (2012), "The integration of ideation and project portfolio management – A key factor for sustainable success", *International Journal of Project Management*, vol.30, iss.5, pp. 582-595.

Hirunyawipada , T., Paswan , A.K. and Blankson, C. (2015) "Toward the development of new product ideas: asymmetric effects of team cohesion on new product ideation", *Journal of Business & Industrial Marketing*, vol.30, iss.7, pp.855-866.

Innovation. [Def. 1,2] in *Merriam Webster Online,* Retrieved March 15, 2017, [On-line] https://www.merriam-webster.com/dictionary/innovation

Iyer, R., Laplaca, P.J. and Sharma, A. (2006), "Innovation and New Product Introductions in Emerging Markets: Strategic Recommendations for the Indian market", *Industrial Marketing Management*, vol. 35, iss.3, pp. 373–382.

Jacobides, M. G., Knudsen, T., & Augier, M. (2006), "Benefiting from innovation: Value creation, value appropriation and the role of industry architectures", *Research Policy*, vol.35, iss.8, pp. 1200-1221.

Jou G.T. and Yuan B.J.C. (2016), "Utilizing a Novel Approach at the Fuzzy Front-End of New Product Development: A Case Study in a Flexible Fabric Supercapacitor", *Sustainability*, vol.8, iss.8, 740.

Jörgensen J.H., Bergenholtz C., Goduscheıt R.C., and Rasmussen E.S. (2011), "Managing Inter-Firm Collaboration in the Fuzzy Front-End: Structure As A Two-Edged Sword", *International Journal of Innovation Management*, vol.15, no.1, pp.145-163.

Khurana A. and Rosenthal S.R. (1998) "Towards holistic front ends in new product development", *Journal of Product Innovation Management*, vol.14, iss.2, pp. 252-269.

Kock A., Heising W., and Gemünden H.G. (2015), "How Ideation Portfolio Management Influences Front-End Success", *Journal of Product Innovation Management*, vol.32, iss.4, pp.539-555.

Kock A., Heising W. and Gemünden H.G. (2016), "A contingency Approach on the Impact of Front-End Success on Project Portfolio Success", *Project Management Journal*, vol.47, iss.2, pp.115-129.

Koen, P. A. (2004), "The Fuzzy Front End for Incremental, Platform, and Breakthrough Products", in the *PDMA Handbook of New Product Development*, Second Edition (ed K. B. Kahn), John Wiley & Sons, Inc.: Hoboken, NJ, USA.

Koen P., Ajamian G., Burkart R., Clamen A., Davidson J., D'Amore, R., Elkins C., Herald K., Incorvia M., Johnson A., Karol R., Seibert R., Slavejkov A., and Wagner K. (2001), "Providing Clarity and A Common Language to the "Fuzzy Front End", *Research-Technology Management*, vol.44, iss.2, pp.46-55.

Kropsu-Vehkapera, H., Haapasalo, H., Rusanen, J.-P. (2009), "Analysis of Technology Management Functions in Finnish High Tech Companies", *The Open Management Journal*, vol.2, pp. 1-10.

Loch, C.H. and Kavadias, S. (2002), "Dynamic portfolio selection of NPD programs using marginal returns", *Management Science,* vol. 48, iss.10, pp.1227-1241.

Microsoft (2013), *Best Practices for Innovation: Microsoft's Innovation Management Framework.* Whitepaper, Microsoft Corporation, [online] http://aka.ms/t1ufzy

Murphy, S.A. and Kumar V. (1997), "The Front End of new product development: a Canadian survey", *R&D Management*, vol.27, iss.1, pp.5-15.

Morris, L. (2011), *The innovation master plan: the CEO's guide to innovation*, Innovation Academy: Walnut Creek, CA, USA, [online] http://innovationmanagement.se/wp-content/uploads/2011/08/Innovation_Master_Plan_Intro.pdf

Nagji, B. and Tuff, G. (2012), "Managing Your Innovation Portfolio", *Harvard Business Review*, vol.90, iss.5, pp.66-74.

Nonaka, I., Toyama R. and Hirata, T., *Managing Flow: A Process Theory of the Knowledge–Based Firm.* Palgrave Macmillan: New York,2008.

OECD, *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data.* 3rd Edition, OECD Publishing: Paris, France, 2005.

Palmberg C., (2006), "The sources and success of innovations – Determinants of commercialization and break-even times", *Technovation*, vol.26, iss.11, pp. 1253-1267.

Park W., Lee K., Doo S., Yoon S., (2016), "Investments for New Product Development: A Break-even Time Analysis", Engineering Management Journal, vol.28, iss.3, pp.158-167.

Riel A., Neumann M., and Tichkiewitch S. (2013), "Structuring the early fuzzy front-end to manage ideation for new product development", CIRP Annals – Manufacturing Technology, vol.62, iss.1, pp.107-110.

Salomo, S., Weise, J. and Gemünden, H.G. (2007), "NPD planning activities and innovation performance: the mediating role of process management and the moderating effect of product innovativeness", *Journal of product innovation management*, vol. 24, iss. 4, pp.285-302.

Sandmeier, P., Jamali, N., Kobe, C., Enkel, E., Gassmann, O., and Meier, M. (2004) "Towards a Structured and Integrative Front-End of Product Innovation", in: R&D Management Conference (RADMA), 6 July 2004 Lisbon, Portugal.

Sawhney, M., Wolcott, R.C. and Arroniz, I. (2006), "The 12 different ways for companies to innovate", *MIT Sloan Management Review*, vol.47, iss. 3, p.75.

Smith, P.G. and Reinertsen, D.G., *Developing Products in Half the Time*. New York: Van Nostrand Reinhold, 1991.

Stamm, B., *The Innovation Wave: Meeting the Corporate Challenge*, Chichester: Wiley & Sons Ltd, 2003.

Stevanovich M., Marjanovic D. and Storga M. (2015), "A model of idea evaluation and selection for product innovation", *in DS 80-8 Proceedings of the 20th International Conference on Engineering Design (ICED 15)*, vol.8: Innovation and Creativity, Milan, Italy, July, 27-30, 2015.

Stevanovic M., Marjanovic D. and Storga M. (2016), "Idea Management in Product Innovation – The Emprical Research Results", *Technical Gazette, vol.*23, iss.5, pp.1285-1294.

Sommer A.F., Hedegaard C., Dukovska-Popovska I. and Steger-Jensen K. (2015), "Improved Product Development Performance through Agile/Stage-Gate Hybrids", *Research Technology Management*, vol.58, iss.1, pp.34-45.

Tenner, A.R. and DeToro, I.J., *Process Redesign: The Implementation Guide for Managers.* Addison Wesley Longman: 1996, Harlow, UK.

Thomke S. and Fujimoto T. (2000), "The Effect of "Front-Loading" Problem-Solving on Product Development Performance", *Journal of Product Innovation Management,* vol.17, iss.2, pp.128-142.

Tidd, J., Bessant, J. and Pavitt, K., *Managing Innovation: Integrating Technological, Market and Organizational Change.* 3rd Edition, John Wiley & Sons: 2005.

Tsekouras G., Meier C., Sutkus A. and Wolf, A. (2014), "The smE-MPOWER Business Innovation Roadmap methodology: A methodological approach for establishing in a participative way a roadmap for business innovation in SMEs", *Technical Results Paper, smE-MPOWER Project of the EU (ETIS-CT-2005-023401)*, 12 pages, [online] www.sme-mpower.eu, CC-BY-SA.

Tuli P. and Shankar R., (2015), "Collaborative and lean new product development approach: a case study in the automotive product design", *International Journal of Production Research*, vol.53, iss.8, pp.2457-2471.

Ucler, C. and Gok, O., 2015, "Innovating General Aviation MRO's through IT: The Sky Aircraft Management System-SAMS", *Procedia-Social and Behavioral Sciences,* vol.195, pp.1503-1513.

Ucler, C. and Kristensen, K. (2016), "CONTEXT – A Novel Approach For Collaborative Virtual Product Realization and Its Disclosure Using the Business Model Canvas", in the 22nd ICE Conference & 2016 IEEE Technology Management Europe Conference, Trondheim, Norway.

Ursino G. (2015), "Supply Chain Control: A Theory of Vertical Integration", *Journal of Economic Analysis & Policy*, Vol. 15, No. 4, pp.1831-1866.

Wever, R. and Boks C. (2007), "Design for Sustainability in the Fuzzy Front End", *in Proceedings of Sustainable Innovation 07*, Farnham, UK, uuid:7dd9685d-9374-4bd1-8c19-ce0582f3301a

Wovak K.D., Craighead, C.W., Ketchen, D.J., Tomas, G. and Hult, M. (2016) "Toward a "theoretical toolbox" for the supplier-enabled fuzzy front end of the new product development process", *Journal of Supply Chain Management*, vol.52, iss.1, pp.66-81.

Zhang Q. and Doll W.J., (2001), "The fuzzy front end and success of new product development: a causal model", *European Journal of Innovation Management*, vol.4, iss.2, pp. 95 – 112.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

USING TEXTUAL FEATURES FOR THE DETECTION OF VANDALISM IN WIKIPEDIA: A COMPARATIVE APPROACH IN LOW-RESOURCE LANGUAGE SECTIONS

DOI: 10.17261/Pressacademia.2017.575
PAP-WCTIE-V.5-2017(13)-p.80-87

Arsim Susuri[1], Mentor Hamiti[2], Agni Dika[3]
[1]South East European University, Faculty of Contemporary Sciences and Technologies , arsimsusuri@gmail.com
[2]South East European University, Faculty of Contemporary Sciences and Technologies , m.hamiti@seeu.edu.mk
[3]South East European University, Faculty of Contemporary Sciences and Technologies , agnidika@yahoo.com

ABSTRACT

This study investigates the impact of using textual features for the detection of vandalism across low-resource language sections in Wikipedia. For this purpose, we propose new features that allow the machine learning-based text classifiers to better distinguish vandalism and to improve the detection rates of vandalism across languages, based on textual features applied in previous researches. These features enable us to compare the contributions of the bots against vandalism, stressing the differences between bots and editors with regards to the detection of vandalism. We propose a new set of efficient and language independent features, which has the performance level similar to the previous sets. Three Wikipedia sections will be used for this purpose: Simple English (simple), Albanian (sq) and Bosnian (bs). We will show that our set of textual features has similar and, in some cases, better vandalism detection rates across languages than previous research.

Keywords: Wikipedia, textual features, low-resource languages, vandalism.

1. INTRODUCTION

Vandalism is a great challenge for Wikipedia, with humans being the main cause, through various illegitimate acts leave traces in computer systems. Our hypothesis is that vandalism can be characterized through models of article views of vandalized Wikipedia articles and that vandalism behaviour is similar across different languages. In the past, a similar research was done by West (2013) and Tran and Christen (2013). This paper is an extension of work originally presented in Susuri et al. (2016), by addressing the issue of using textual features detecting vandalism in Wikipedia's articles across languages. According to our hypothesis, a model developed in one language can be applied to other languages. If successful, this would drop the costs of training the classifiers separately for each language. Applying this model of vandalism detection across different languages shows similar results. In this paper, we will explore the possibility of applying the detection of vandalism across languages through textual features. We combine these data sets in order to analyse any gains in terms of language independency of certain features.

For this purpose, we compare performances of standard classifiers for identifying vandalism in three Wikipedia data sets (Simple English, Albanian and Bosnian). On top of this, we compare the performances of classifiers in one language and the other one and in the combined, data set. Vandalism detection bots have and are changing the way Wikipedia identifies and prevents vandals (Geiger and Ribes, 2010). However, contributions from these bots are rarely discussed, despite their importance for maintaining the reliability of Wikipedia, some of whom have become the most active editors (Adler et al., 2008). The increase of responsibilities of bots in the detection of vandalism raises several research issues:

- How big is the difference in values between the bots of vandalism detection and editors?
- How do they differ across different languages?

Our research is based on learning to detect vandalism bots together as well as from the editors, and then implemented these models in three different languages: English (simple), Albanian (sq) and Bosnian (bs).

We propose a new set of efficient and language independent features, with similar or better performance levels in comparison to previous textual features. We will show that bots and users have similar reference values in terms of identifying vandalism when their models apply to other sets of classification in cases of vandalism. The contributions of this paper are:

- Development of new textual features used for the detection of vandalism, independent of language, with better efficiency in comparison to previous researches
- Determining the differences between bots and users regarding vandalism identified by bots and users
- Demonstration of application of classification models across languages without losing out on quality classification

The paper is organized as follows. Section 2 describes data sets used for this purpose, and the textual features used for the purpose of the detection of vandalism across languages. Section 3 reports and discusses the empirical results. And the final section gives conclusions.

2. DATA SETS AND FEATURES

2.1. Data Sets

Wikipedia saves all article revisions in the so-called history dumps in XML or SQL format. For the purpose of this paper will use the history dumps of Wikipedia in three low-resource language section: history dumps in Albanian[1], Simple English[2] and in Bosnian[3]. We use the Wikipedia history dumps of edits dated 29.10.2015, for Simple English, Albanian and Bosnian language sections. History files (*pages-meta-history*) include article revision history, including additional data related to the revision and labeling text.

The reason for choosing these data sets is that, to our knowledge, no one else has analysed the history dump of Wikipedia in Albanian in the context of detection of vandalism. Two other history dumps were selected to validate our hypothesis about the independence of the textual features, to validate detection of vandalism across languages, and because the sizes of these data set are similar. In terms of the volume of articles, similarities exist between these three editions of Wikipedia, as described below:

- Wikipedia in Simple English contains 119,183 articles[4]
- Wikipedia in Albanian contains 55,894 articles[5]
- Wikipedia in Bosnian contains 69,765 articles[6]

Therefore, we can compare the possibility of applying our textual features in the detection of vandalism. We split these data sets into training sets (revisions before 2015) and testing sets (revisions in 2015). The testing sets contain 8-28% of all revisions for each of the three language sections. We will distinguish contributions of bots from contributions of editors (users), and compare vandalism that one classifier can learn from vandalism repairs made from bots or users.

2.2. Feature Creation

We create our textual features stemming from differences in the content of the repaired revision of the article in comparison to the previous revision, containing vandalism. Using the *diff* algorithm (Hunt and Mcllroy, 1974) we obtain unique rows in terms of revision before the repair, unique rows in terms of revision after the repair and revised rows during the process of repair. We do not take into account common rows, due to precise designation of changes in the content. Common rows show us the ratio of vandalized content and the legitimate content, but in cases of mass deletions, the size of unique rows in relation to the repaired revision is sufficient to indicate vandalism.

The features are shown in Table 1.

We then apply the process of determining the difference at the word level with the aim of extracting vandalized words that were repaired. For the processing of the text, we use Unicode (UTF-8) and alphabets for the English, Albanian and Bosnian language, respectively. Along with the appropriate description and the average time for generating features, we also use the statistical test of Kolmogorov-Smirnov (K-S) (Massey, 1951). Our features are applied on changed words instead of changed revisions (as in previous researches). We use textual features from PAN 2010 and PAN 2011 Workshops (Mola-Velasco, 2010; West and Lee, 2011), where we find the first application of textual features for the purposes of detecting vandalism.

[1] https://dumps.wikimedia.org/sqwiki/
[2] https://dumps.wikimedia.org/simplewiki/
[3] https://dumps.wikimedia.org/bswiki/
[4] https://simple.wikipedia.org/wiki/Main_Page
[5] https://sq.wikipedia.org/wiki/Faqja_kryesore
[6] https://bs.wikipedia.org/wiki/Početna_strana

Features from V01-NR1 to V10-NTF are created from the revisions before and after the repair. Features from V11-FP to V18-KV are created from changed words during the repair process, which isolate possible vandal words and include distribution of words in the repair. Features from V19-RMV to V24-SGV are applied on each of the repaired words, where values indicating vandalism are chosen.

2.2.1. Modifying Features

The above-mentioned features are suitable for detecting changes in the content of the articles.

Features from V01-N1 to V04-NRN2 are enumerations of types of rows from the *diff* algorithm. High values of unique rows in the vandalized revision (before repair) could be a sign of massive insertions, while higher values in the repaired revision (after repair) could be a sign of massive deletions. The number of rows changes may be a sign of small changes as a sign of vandalizing insertions or textual changes.

Table 1: List of Features Created Before (1) and After (2) Repair

Features	Description	Time (ms)	K-S test (failures)	K-S test (PAN) (failures)
V01-NR1	Number of Unique Rows in (1)	0.043	10%	0%
V02-NR2	Number of Unique Rows in (2)	0.043	0%	50%
V03-NRN1	Number of Unique Revised Rows in (1)	0.043	10%	50%
V04-NRN2	Number of Unique Revised Rows in (2)	0.043	10%	25%
V05-DGTF	Difference in Total Length of Unique Words in (1) and (2)	0.522	0%	25%
V06-DNTF	Difference in Total Number of Unique Words in (1) and (2)	0.472	0%	0%
V07-RGTF	Ratio of Total Length of Unique Words in (1) and (2)	0.522	11%	25%
V08-RNTF	Ratio of Total Number of Unique Words in (1) and (2)	0.472	11%	25%
V09-NFV	Number of Unique Words	0.005	12%	0%
V10-NTF	Number of Total Words	0.003	125%	0%
V11-FP	Pronoun Words	0.011	50%	100%
V12-FZ	Slang Words	0.007	30%	50%
V13-FV	Vulgar Words	0.007	50%	100%
V14-FSM	Capitalised Words	0.006	10%	0%
V15-FAN	Alphanumerical Words	0.006	10%	0%
V16-SV	Single Letters	0.007	80%	100%
V17-NN	Single Digits	0.005	22%	75%
V18-KV	Single Characters	0.006	80%	100%
V19-RMV	Highest Ratio Uppercase-Lowercase	0.170	0%	25%
V20-RSS	Highest Ratio Digits-All Letters	0.170	0%	25%
V21-GFP	Length of the Most Repeated Word	0.200	12%	50%
V22-GKG	Length of the Longest Character Repeated	0.175	10%	50%
V23-FGV	Longest Unique Word	0.045	10%	25%
V24-SGFV	Sum of Length of the Unique Words	0.045	10%	0%

These features are similar to the feature byte change in the West and Lee (2011), but the difference here is the possibility of clarifying the impact of changes in these features. Features from V05-DGTF to V10-NTF are similar to the enumeration of rows. Through these features the process of enumerating changes in words, before and after repair, is implemented. These changes indicate cases of minor vandalism that modify specific words. The difference in length of specific words, and the difference in numbers of specific words determine the relative size changes needed for revision repair. Lengths and numbers of special words determine the relative change in size and the sheer number of changes needed to repair the vandalism. These combinations of features can identify repairs made by bots in cases of minor vandalism.

2.2.2. Features from Wiktionary

We borrow features from Wiktionary[7] in English[8], Albanian[9] and Bosnian[10] languages. At the same time, we are borrowing the workshops features of PAN Workshops, adapted for our testing purposes. By using features from V11-FP to V13-FV we analyse the ratio letter / word considering three types of actual vandalism: pronouns, slang and vulgarism. In total, there are 40 pronouns, 1130 slangs and 347 and vulgarisms. We seek these words in the *diff* algorithm at the sentence level. For example, if vulgarism in the English language is used in the Albanian language, these vulgarism features are numerated into

[7] https://www.wiktionary.org/
[8] https://simple.wiktionary.org/wiki/Main_Page
[9] https://sq.wiktionary.org/wiki/Faqja_kryesore
[10] https://bs.wiktionary.org/wiki/Početna_strana

the features for revision in the Albanian language. Visual inspection shows that slang and vulgarism are not usually common words in other languages. Features from V14-FSM to V18-KV are used to enumerate different types of sentences. By analysing the letters of each word, some indications of possible vandalism can include words that start with uppercase, words with numbers and words that consist of only one letter. These features are common indicators of vandalism in other studies (Mola-Velasco, 2010; West and Lee, 2011), as well.

2.2.3. Features from Words

Features presented in Table 1 are based on preliminary research but have been modified in order to apply in our analysis at the level of words. In analysing the change of sentence, the idea is that an unusual appearance a presentation of unusual words point to vandalism, therefore, not apply a total or average value based on the fact that the vandaliser can attempt to avoid detection of vandalism by disguising the vandalism with legitimate question but not related to sentences vandalized. Features from V19-RMV to V20-RSS are used to determine the ratio letter/word. These features are based on research conducted by Mola-Velasco (2010), with a difference in that the application of our experiments is at the level of words and not at the level of the document. Minimum or maximum values are a clear indicator of vandalism. Features from V21-GFP to V22-GKG are used to determine the length of the longest repeated character in a word, as applied in Mola-Velasco (2010), as one of the clear indicators of vandalism. Feature from V23-FGV to V24-SGFV are used to separate longest unique words and total size of unique words in the changed words. These features are based on Mola-Velasco (2010) and West and Lee (2011), but with a modified application.

3. EXPERIMENTS

We apply separation of the Wikipedia data sets into training sets (all reviews prior to October 2015) and into testing sets (after October 2015). Since the data set is not balanced, we apply under-sampling in order to balance the legitimate revisions with vandalised revisions. Through this preparatory procedure we enable the Random Forest algorithm to improve performance with more balanced sample trees.

3.1. Classification Results

We use the Random Forest classifier in Weka[11]. This classifier is better in terms of performance, as shown in previous relevant research (Adler et al., 2011), therefore this is the reason of the application of this classifier in our tests. To get the most out of the performance of this classifier, we apply a ten-fold cross validation on the training data, using a number of parameters of this classifier, such as the number of the forecasters (the trees in the forest), the maximum number of features, the minimum number of leaf samples, the minimum number of samples for separation and minimal density. The results obtained are measuring values of Area Under Curve – Precision Recall (AUC-PR). The reason for using AUC-PR is that these values obtained are not affected by the design of the data (under-sampling in our case) (Davis and Goadrich, 2006).

3.1.1 Combining Classification Languages

For complete sets of Wikipedia (simple and sq), Table 2 presents the results of the combination of training and testing data. Within the same language and type of users (diagonal values), the classifier shows higher performance in comparison with language combinations. Except in the case of the bots' values in Albanian, where the classifier trained on the data from bots in Wikipedia in English shows better results. This suggests that the bots of Wikipedia in English can identify more cases of vandalism in the Albanian section of Wikipedia than bots of Albanian section of Wikipedia.

Table 2: Classification Across Languages and Users for the Random Forest Classifier

AUC-PR	**Testing**	**simple**		**sq**		**bs**	
Training	***Type***	Bots	users	bots	users	bots	users
simple	Bots	0.923	0.821	0.871	0.732	0.913	0.778
	Users	0.912	0.844	0.849	0.755	0.905	0.785
sq	Bots	0.899	0.766	0.864	0.730	0.854	0.764
	Users	0.893	0.785	0.847	0.749	0.834	0.783
bs	Bots	0.912	0.801	0.881	0.787	0.918	0.768
	Users	0.905	0.823	0.847	0.802	0.907	0.790

For users in three languages, we find consistent performance in the detection of vandalism in three languages. This suggests that users search for similar patterns of vandalism just as bits do. Users of Simple English Wikipedia identify, proportionally, more cases of vandalism across languages than users of other languages. This suggests that with more users (editors), more vandalism models can be identified.

[11] http://www.cs.waikato.ac.nz/ml/weka/

3.1.2 Combined Training Data

We combine training data from bots and users for both languages, for each type of editor, and for two languages and two types kinds of editors. The classification results are presented in Table 3. The purpose of this test is to try to apply the learning of vandalism without distinguishing the contribution from bots or users. By learning from the bots and users for each language, we find some differences in the classification performance. Bots follow common rules and structures of vandalism that learning algorithms learn fast, providing more accurate results and higher AUC-PR value. On the other hand, cases of vandalism detection from editors have greater variation in vandalism types as there are cases of vandalism which cannot be detected by bots. Similarly, as in the previous case, we find no statistically significant difference when comparing the rows between Tables 2 and 3. This shows that there are no differences in learning vandalism from bots or users across languages.

So, the combination of observations from bots and users does not cause performance improvement in the vandalism detection rates identified by users. It makes us realize that, on one hand, users really identify a wider array of vandalism types, while on the other hand, the contributions of bots do not change with the changing of languages, but nevertheless improve the classification performance. Although these improvements are small, in reality it means hundreds or thousands of cases (depending on the section of Wikipedia) of vandalism are automatically detected.

Table 3: Classification Across Languages and Users

AUC-PR	***Testing***	**simple**		**sq**		**bs**	
Training	***Type***	botët	përdoruesit	botët	përdoruesit	botët	përdoruesit
simple	Bots/users	0.932	0.808	0.902	0.732	0.911	0.785
sq	Bots/users	0.924	0.798	0.905	0.735	0.896	0.782
bs	Bots/users	0.927	0.803	0.915	0.734	0.934	0.784
Total	Bots	0.937	0.793	0.931	0.730	0.931	0.783
Totai	Users	0.927	0.815	0.953	0.741	0.937	0.791
Total	Bots/users	0.935	0.816	0.947	0.754	0.934	0.795

3.1.3. Combined Training and Testing Data Sets

To complete the learning across languages and to have comparable data with previous research, we combine both types of editors for training and testing data. In Table 4 we have presented the results of the classification across languages for each language and combined training for the world and for three languages user total data and training. Results of the relevant training and testing data of the languages in question (Simple English, Albanian and Bosnian) show AUC-PR values being between the values in Table 3 and 4. This enables us to understand that by using all training data, detection rates benefit significantly better statistically in both languages.

Table 4: AUC-PR values across three languages

AUC-PR	**Testing**		
Training	simple	sq	bs
Simple	0.874	0.734	0.804
Sq	0.865	0.732	0.811
Bs	0.872	0.722	0.833
Bots	0.871	0.723	0.802
Users	0.871	0.734	0.754
Total	0.884	0.743	0.834

3.1.5 Results of Data Modelling

We have modeled over-fitted legitimate reviews by the Random Forest classifier because it enables us to build decision-making tree on the classifier in order to distinguish vandalism, reduces the size of the model and data needed for training, and reduces the time of the learning. However, modeling the data raises questions of bias in performance. Table 4 shows values according to the ratio 1:1 of legitimate reviews with vandalised reviews (over-fitted values). In order to have more convincing results, we repeated our experiments based on modelled ratios of 2:1, 5:1, 9:1 and 13:1. The latter two values are applied in the experiments because the share of the vandaised reviews in the PAN-WVC-10 and PAN-WVC-11 data sets is approximately 7% (93% reviews legitimate - 7% vandalised reviews) which coincides with the ratio 13: 1. Ratio 9: 1 reflects the range of vandalism found in previous studies (Potthast, 2010).

We compare the values obtained in Figure 1, for the classification within the same language (diagonal values in Table 4). For classification across languages, in Figure 2 we have shown average PR-AUC values, along with mean standard error. Based

on figures and applied tests, we can conclude that the re-modelling process has little negative effet on the values of classification. In Figure 1 we have shown the results of the training applied to the balanced data,with a ratio of 1: 1, together with the results of applying on unbalanced testing data sets with ratios 8:1 and 13:1. These results simulate real consequences of learning in the balanced data set and applying in the unbaanced data set, such as full sections Wikipedia. In Figure 3 we have shown the results of the classification within the same language, as well as in Figure 4 the classification across languages, with average AUC-PR values, along with mean standard error. Experiments applied in PAN data sets are included in the Figures 3 and 4 for comparative purposes.We can see from the figures, that the proposed features have very good results, compared with previously applied features. Similarly, for the classification across languages, the proposed features have very good and comparable values to previous research. On this basis, we can say that there is a stable trend of vandalism within a language, which is applicable to a large extent, across languages.

4. FINDINGS AND DISUSSIONS

The benefit from the application of machine learning across languages for the detection of vandalism is the generalization of classification models for different sections of Wikipedia, without learning from many cases of vandalism. Our results show that learning the language that has many cases of vandalism, such as the section on

Figure 1: Comparison of Different Values of Classification across Languages

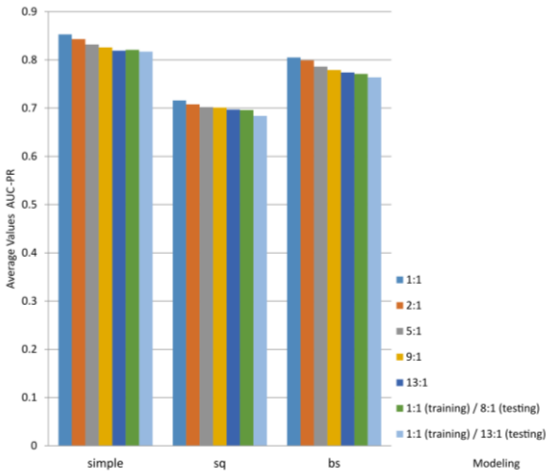


Figure 2: Comparison of Different Features of Classification within Languages

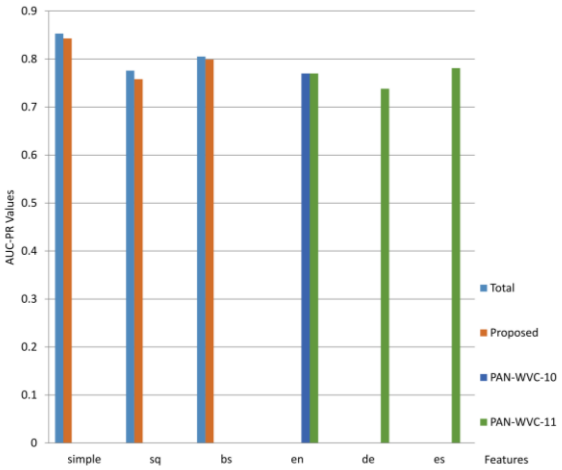
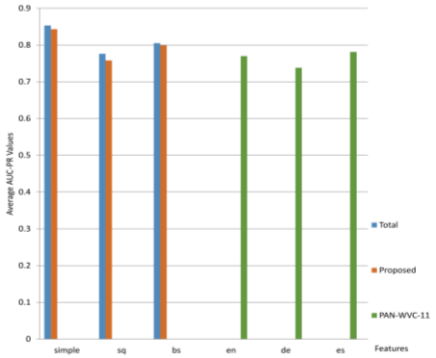
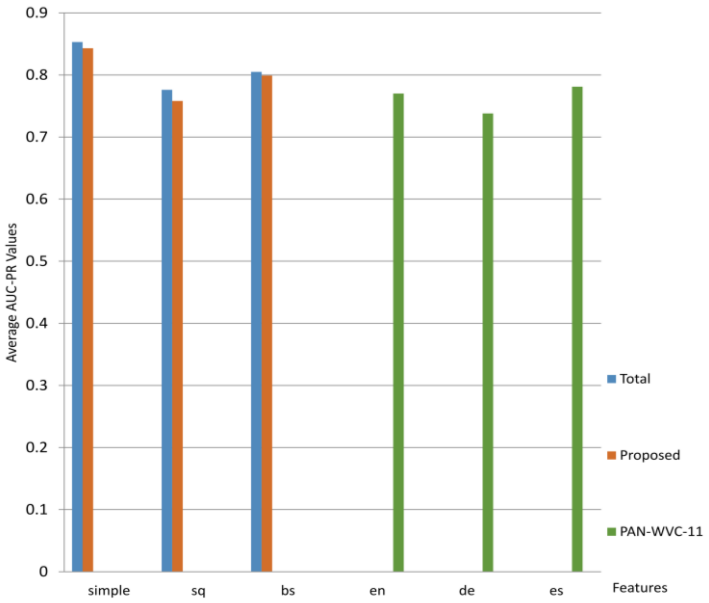


Figure 3: Comparison of Combination of Different Features of Classification across Languages



Simple English may generalize very well to smaller sections of Wikipedia, such as the sections in Albanian and in Bosnian. One advantage of our approach is the immediate and comparative analysis of the current revision of the text of the previous review, with the aim of determining the potential vandalism.

Figure 4: Comparison of Combinations of Different Features of Classification across Languages



We do not need to add meta data, and user profiling to determine vandalism.

Our textual features show performance which is comparable to the previous research and improves performance based on the preliminary review models in Wikipedia. Our features are designed with the aim of generalizing the language in experimentation, which is reflected in the classification performance.

A limitation of our research is to support the textual features that cannot easily detect vandalism which can detect whether we use metadata or user features reputation. Our classification method uses under-fitting as a balanced approach to distinguish between the majority class (legitimate reviews) and minority class (vandalizing revisions), and to reduce the training data set. Although this method can add noise in the data sets, results in Section 3.1.5 tell us that under-fitting does not affect the classification results significantly, if we repeat the experiments with different proportions of training and testing. We have presented only one classifier performance which, although in these sections Wikipedia has shown consistent results, may not be the best in other sections of Wikipedia.

5. CONCLUSION

In this chapter we applied the comparison between bots and users, in terms of identifying based on the detection of vandalism in low-resources sections of Wikipedia. We have developed textual features that include features commonly used for the detection of vandalism and we used a classifier for the features, listed above, important to bots and to users in

three language sections of Wikipedia. We present and discuss the differences between the bots and users in terms of identifying vandalism in three sections of Wikipedia: Simple English, Albanian and Bosnian languages. Comparison with previous research has shown that our techniques are comparable and sometimes better than these researches. Our contributions showed that we can apply the learning of vandalism in one language section of Wikipedia and then apply a classifier in other sections with little loss in quality of classification.

In the future, the focus of the research should be anonymous users' contribution in identification of vandalism. This is because of difficulties in determining their identity. Creating an online system for the purpose of evaluating the efficiency of the characteristics previously submitted and to assess the performance impact of these features in the detection of vandalism, will further improve the detection of vandalism.

REFERENCES

Adler B. T, de Alfaro L., Pye I., 2008, "Measuring author contributions to the Wikipedia. In: WikiSym '08, Porto, Portugal, 8-10 September 2008. New York: ACM.

Adler B. T., de Alfaro L., Mola-Velasco S. M., Rosso P., and West A. G., 2011, "Wikipedia vandalism detection: Combining natural language, metadata, and reputation features". In Proceedings of the 12th International Conference on Computational Linguistics and Intelligent Text Processing - Volume Part II, CICLing'11, pages 277 - 288, Berlin, Heidelberg, Springer-Verlag.

Davis J. and Goadrich M., 2006, "The Relationship Between Precision-Recall and ROCCurves". In Proceedings of the 23rd International Conference on Machine learning (ICML), 2006.

Geiger R. S. and Ribes D., 2010, "The Work of Sustaining Order in Wikipedia: The Banning of a Vandal". In Proceedings of the 22nd ACM Conference on Computer Supported Cooperative Work (CSCW).

Hunt J. W., Mcllroy M. D, 1974, "An Algorithm for Differential File Comparison", Computer Science Technical Report, Bell Laboratories.

Massey F. J., 1951, "The Kolmogorov-Smirnov Test for Goodness of Fit". Journal of the American Statistical Association, 46.

Mola-Velasco S. M., 2010, "Wikipedia Vandalism Detection Through Machine Learning: Feature Review and New Proposals". In CLEF (Notebook Papers/Labs/-Workshops).

Susuri A., Hamiti M. and Dika A, 2016, "Machine Learning Based Detection of Vandalism in Wikipedia across Languages". In proceedings of the 5th Mediterranean Conference on Embedded Computing (MECO), Bar, Montenegro.

Tran K.N., Christen P., 2013 "Cross-language prediction of vandalism on wikipedia using article views and revisions". Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD).

West A. G., 2013, "Damage Detection and Mitigation in Open Collaboration Applications", Ph.D. thesis, University of Pennsylvania.

West A. G. and Lee I., 2011, "Multilingual Vandalism Detection using Language-Independent & Ex Post Facto Evidence". In CLEF (Notebook Papers/Labs/Workshops).

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

INDOOR LOCALIZATION FOR WIRELESS SENSOR NETWORK AND DV-HOP

DOI: 10.17261/Pressacademia.2017.576
PAP-WCTIE-V.5-2017(14)-p.88-95

Merve Kanmaz[1], Muhammed Ali Aydin[2]
[1]Istanbul University, merve.kanmaz@istanbul.edu.tr
[2]Istanbul University, aydinali@istanbul.edu.tr

ABSTRACT

Indoor positioning is a significant problem for many work areas. Indoor positioning due to breakage and disappearance of signals due to walls and other obstacles in indoor requires a different solution to the outdoor Therefore, the GPS, which is successful for outdoor positioning, can not work in the same way as internal positioning. Positioning can be done by using anchor nodes that have their own position information located in wireless sensor networks In this study, DV-hop method, one of these methods, has been examined. It has been investigated in three different methods presented as an alternative to the classical Dv-Hop method and its advantages and disadvantages are discussed.

Keywords: Indoor localization, wireless sensor network, positioning, Dv-Hop

1. INTRODUCTION

With the development of technology, academic and industrial field requirements have changed. Knowing the location of the users is one of the most important of these needs. The applications that are being worked on aim to make life easier for users, especially in areas such as health, military pursuit and security [1]. GPS is the most used and most successful system in outdoor positioning [2]. However, GPS is not as successful as outdoor positioning in terms of indoor positioning. Because especially in the interiors walls are an obstacle to the transition of any object signal. Objects that are not completely obstructive to the passage cause the signals to be broken, which cause power loss and therefore the wrong result. However, the necessity of installing a GPS transmitter on every object to be tracked by GPS increases the cost considerably [3].

Wireless sensor networks are also one of the technologies used for positioning. Reducing the cost of used sensors, making them less energy-consuming, and wider and more effective networks can be achieved after the scaling of their dimensions. In this way, more convenient information can be collected and processed from the environment. When positioning in sensor networks, it is necessary to mention the anchor nodes. These nodes are nodes that know their location. By using these nodes, the position of the nodes that do not know their position is determined. Several methods have been developed to do this positioning. These are classified as distance based, anchor node based, and hybrid constructions. In this report, after the information about some developed technologies related to this subject is given, the ongoing Dv-Hop method will be detailed. Alternative Dv-Hop methods will be compared with the developed methods.

2. WIRELESS SENSOR NETWORK POSITIONING

Wireless networks are produced as an alternative to wired networks and are technology that enables communication with specific standards from the air. Wireless sensor networks are called wireless networks that contain sensor nodes. Sensors can communicate wirelessly and have features such as signal processing and propagation. Sensors have a certain amount of energy and coverage. Therefore, they can serve at a certain capacity. More sensors are needed to build a more efficient structure in wireless sensor networks. This is a factor that increases the cost. However, the coverage area and the amount of energy in the sensors are of a certain level. Therefore, a network structure with more sensors must be built in order to create a structure that will lead to more efficient results. This will increase the cost and the amount of energy required. This

should not be overlooked in the systems to be installed. It has been reported that a sensor for a viable system should cost less than $1 [4]. Satisfying positioning requires certain conditions. These:

- **Accuracy**: It is necessary that the estimated location information should has acceptable error rate
- **Sensitivity**: The result should be consistent in multiple experiments.
- **Scalability**: The number of nodes in the study area should be observable.
- **Energy Consumption**: System resources should be used efficiently.

3. LOCATION ESTIMATION TECHNIQUES

3.1 Infrared

They are electromagnetic signals with a coverage of about 5 m [3]. They are used in systems such as remote control. They are reflected back on wall-like obstacles. When positioning is performed, each object to be monitored is provided with an infrared device, and during certain periods, the signals sent from this device are processed by a central server to determine the position.

3.2. Cellular Network

They are networks formed by radio cells. It operates at a higher capacity than normal networks and requires less energy. When using cellular networks in the positioning process, more than one base station is placed in the area to be controlled so that the signals are better tracked. Accuracy is high. It has been shown to give up to 2.5 meters of accuracy in a work carried out with cellular networks [5].

3.3. Bluetooth

It is a technology that many devices already have, with low cost and low energy requirement. The disadvantage is that the coverage area is limited and delayed responding.

3.4. Wireless Local Area Network

Although it is not manufactured for positioning, it allows positioning due to the technology it is housed on. It uses the 802.11 infrastructure. Uses the RSSI (Receiver Signal Strength Detector) value when positioning. It saves the signal strengths received from the node to a file called Radio Map and performs the positioning operation by checking this file when positioning. This approach is also called fingerprinting.

3.5. Wireless Sensor Network – WSN

Sensors are heat, temperature, noise or nerves sensitive devices. The wireless sensor is a network that consists of these devices. When positioning in sensor networks, it is necessary to mention about the anchor nodes. Anchor nodes are nodes that know their location within the network. Inside the network, there are nodes that do not know their location called blind nodes. When positioning the blind nodes anchor nodes being used. Once the distance or angle estimation has been performed with the related algorithms, physical positioning is performed by applying one of the positioning techniques. Techniques used in physical positioning process are:

- **Lateration**: Positioning is performed by using the distance information of at least 3 points on the 2-dimensional plane.
- **Angulation**: Positioning is performed on the 2-dimensional plane by using at least 2 pieces of angle information of the same non-axial point.

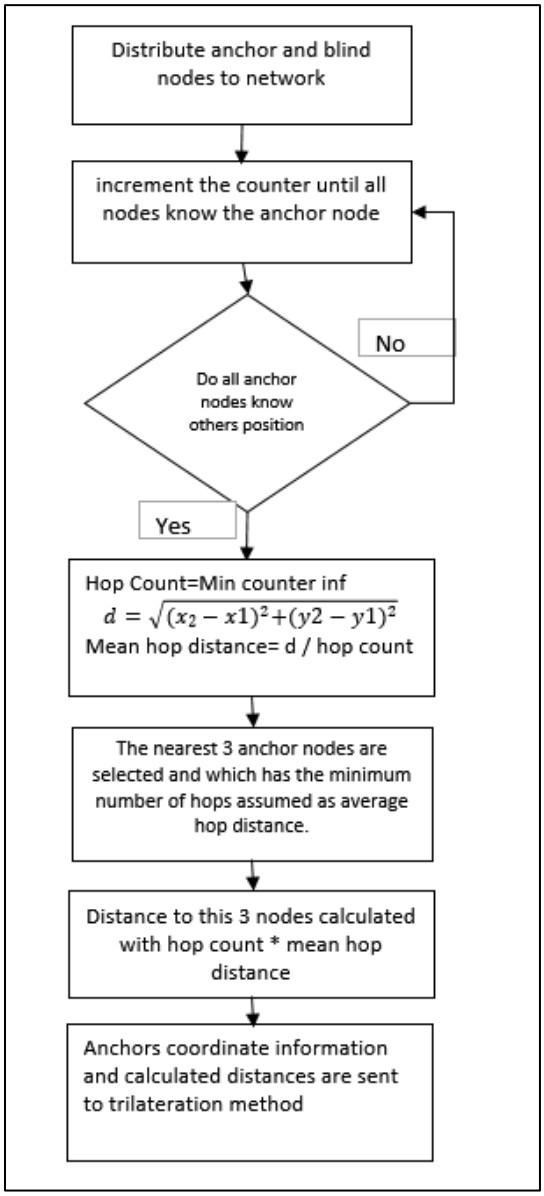
4. POSITIONING ALGORITHMS

Many algorithms have been developed for the problem of indoor positioning. Some of them are based on signal strength and some on the basis of distance. In addition, it is used for positioning in the same way in hybrid structures which combine multiple algorithms. Each algorithm has advantages and disadvantages.

4.1. Dv-Hop Algorithm

Algorithm is the basic Dv-Hop algorithm that was introduced in 2001. The mean hop distance value for each anchor node in the network is used and this value is used in the positioning process [6]. The flow chart of the algorithm is as follows:

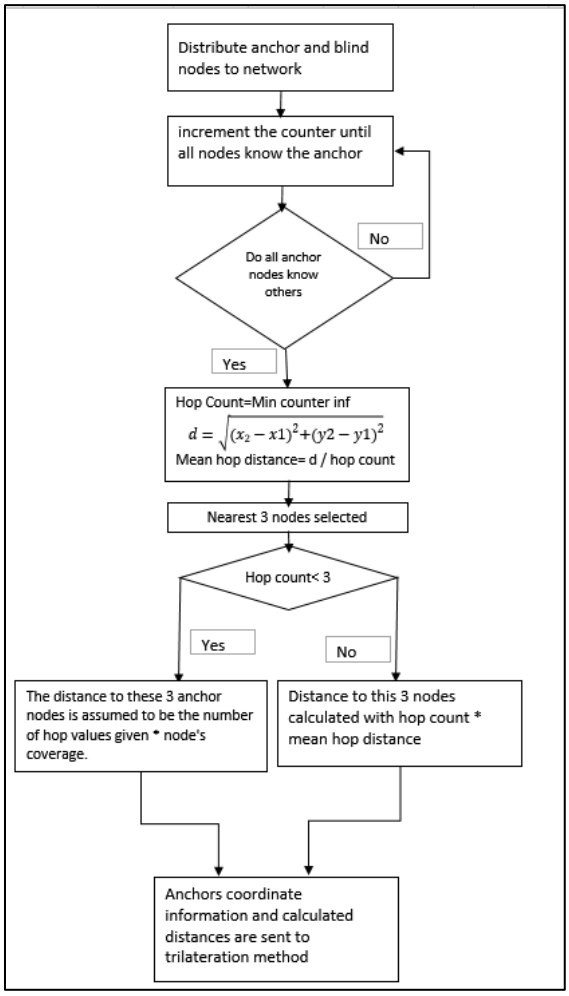
Figure 1: Dv-Hop Algorithm Flow Chart



4.2. RSSI Based DV-Hop Algorithm

The algorithm was developed in 2007 by combining the Dv-Hop algorithm with RSSI (Received Signal Strength Indicator) to obtain better results. It has been proposed that a different calculation would yield more accurate results when calculating the location of unknown node position while calculating the positions of the nodes that can take the RSSI message more than 3 and more hop counts [7]. The flow chart of the algorithm is as follows:

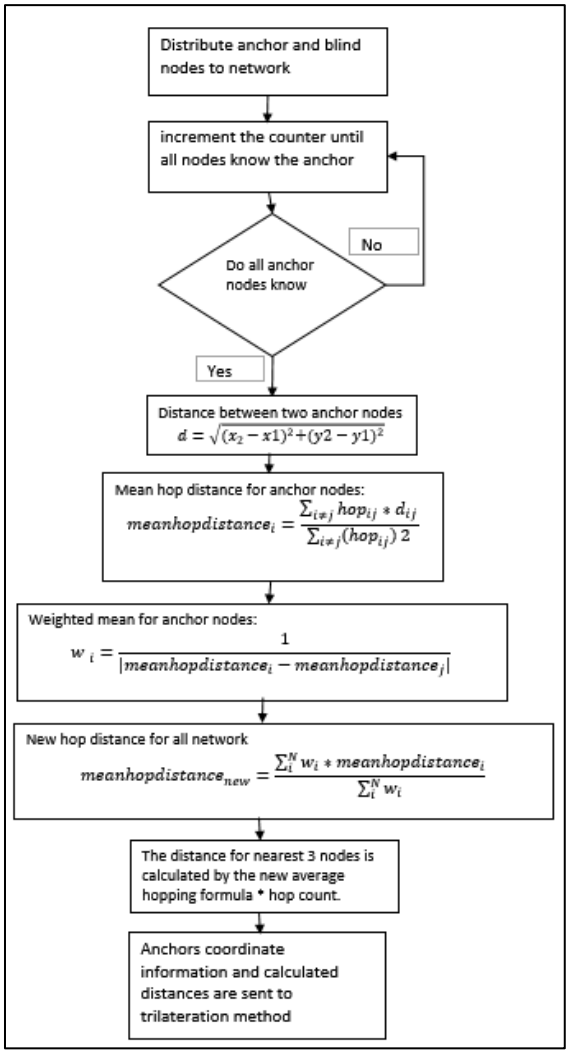
Figure 2: RSSI Based Dv-Hop Flow Chart



4.3. Average Hop Weighted Mean Dv-Hop Algorithm

In this method, which emerged in 2014, the average hop distance calculation for each anchor node used in the Dv-Hop algorithm are changed and the average hop distance for all network is calculated. This hop distance is used for the positioning of all nodal points whose position is unknown [8]. The flow chart of the algorithm is as follows:

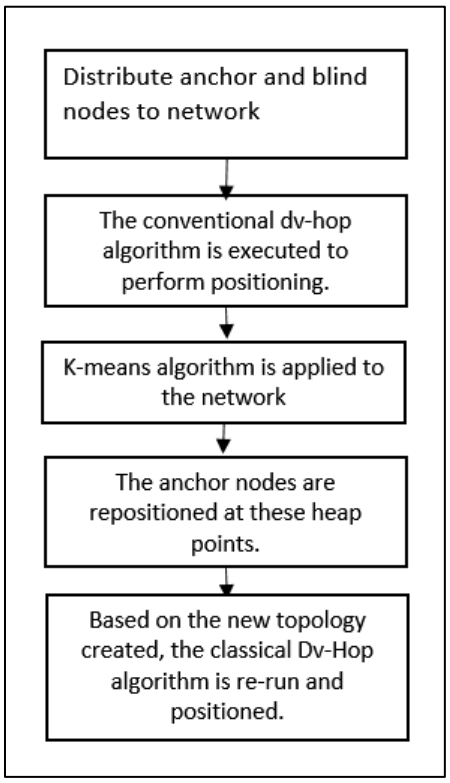
Figure 3: Average Hop Weighted Mean Dv-Hop Algorithm



4.4. Dv-Hop with Clustering Method

In this method, once the Dv-Hop algorithm is executed, it is considered to change the positions of the anchor nodes in the network by running the k-means clustering algorithm on the network. With K-means algorithm, the anchor nodes are placed in order to send the signal to the other nodes more easily at the center point when the clustering is done. According to the last settlement, the Dv-hop algorithm is run again and compared against the first case [9]. The flow chart looks like this:

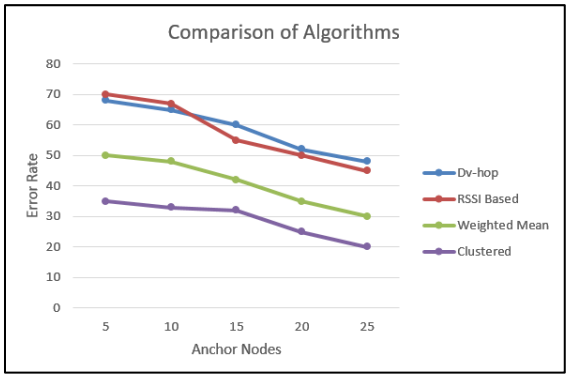
Figure 4: Dv-Hop with Clustering Method



5. CONCLUSIONS AND RECOMMENDATIONS

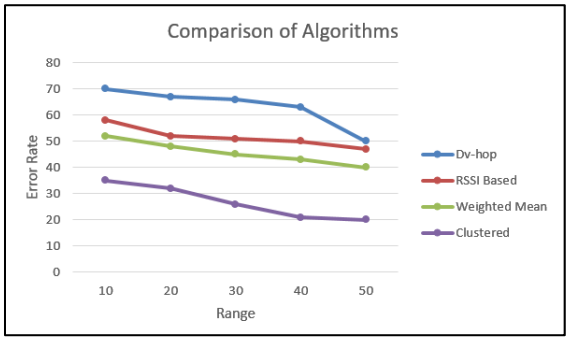
In this study one of anchor based approaches DV-Hop method and alternative Dv-Hop methods are compared to each other and which cases the accuracy can be increased have been investigated. To compare the methods, the simulation environment was first prepared with the C # programming language. The working environment consists of 50 nodes with 100x100m area, each with a coverage area of 25m. First, the parameter monitored is the number of anchor nodes in the network. This was determined as 5, 10, 15, 20 and 25, respectively, with 10%, 20%, 30%, 40% and 50% depending on the number of nodes in the network. The methods were run 100 times in the determined working environment and average results were obtained. According to the results obtained, it looks like the graph.

Figure 5: Comparison of Algorithms -1



The second parameter monitored on the same topology is the coverage area. The coverage area has been increased to 10, 20, 30, 40, 50m respectively. The results obtained are as in the chart.

Figure 6: Comparison of Algorithms – 2



According to the graphical results:

- In 4 algorithms, number of nodes in the network increases, the accuracy rate increases. However, it should be taken into consideration that this situation will also increase the cost.
- It has been observed that the dv-hop method lags behind other methods because it is based on the number of hops and distance.
- In the case of RSSI-based DV-hop method, the algorithm does not always perform better than Classic Dv-hop method in each topology because it behaves differently when the hop count is less than 3.
- The Average Hop Weighted Mean Dv-Hop method calculates the average hop distance by looking at the network in general, and it produces better results compared to the conventional Dv-hop method and the RSSI based Dv-Hop method.
- The clustering Dv-Hop method has observed a significant increase in accuracy due to the fact that the anchor nodes facilitate access to the unknown nodes since they once clustered on the nodes in the network and then moved the anchor nodes to these calculated center points [10].
- If the coverage area is too great, one-to-one transmission will take place with a very small number of hops and this will also adversely affect these methods, which depend on the number of hops.
- Observed results give the expected results when appropriate topologies are established. It has been observed that even the best algorithms in different topologies give unexpected results.

As a result, the application of better clustering algorithms to the network is expected to give better results as it improves the network structure. In future studies, the results can be evaluated using untested clustering algorithms. In addition, these results can be tested on real environments. On observing that the method yields much better results in the appropriate topologies, the following is achieved: using better clustering algorithms is better. The monitoring methods that give much better results on the appropriate topology has concluded that the clustering algorithms to be implemented better method you use will give better results. Future studies can use better untested clustering algorithms and evaluate the results. Moreover, the test environment can be applied in real environment and the results can be evaluated.

ACKNOWLEDGEMENTS

This work was a part of the master thesis titled “Precise Localization With Efficient User Mobility” at Istanbul University, Institute of Physical Sciences.

REFERENCES

Pahlavan, K., Li, X., Ylianttila, M., Chana, R.S. , Latva-aho, M., An overview of wireless indoor geolocation techniques and systems, 2000, Proceedings of the IFIPTC6/European Commission International Workshop on in Mobile and Wireless Communications Networks, 1–13.

Enge, P. , Misra, P., Special issue on GPS: The global positioning system, 1999, Proceedings of the IEEE, 3–172.

Doğancı, Y.U. , 2008, 802.11 Standartlarını kullanarak pozisyon Tespiti, Yüksek Lisans, TOBB Ekonomi ve Teknoloji Üniversitesi Fen Bilimleri Enstitüsü.

Akyildiz, I. F., Su, W., Sankarasubramaniam ,Y., E. Cayirci, 2002, Wireless Sensor Networks: A Survey , Computer Networks, 38 (4), 393-422.

Otsason, V., Varshavsky, A., LaMarca A., E. de Lara, 2005 , Accurate gsm indoor localization, Ubicomp (M. Beigl, S. S. Intille, J. Rekimoto, and H. Tokuda, eds.), vol. 3660 of Lecture Notes in Computer Science, 141–158.

Niculescu ,D., Nath, B., 2001, Ad-hoc Positioning System, IEEE Global Telecommunications Conference, San Antonio ,vol.5, 2926 – 2931.

Tian S. , Zhang X. , Liu P. , Wang X., 2007, A RSSI-based DV-Hop Algorithm for Wireless Sensor Network, Wireless Communications, Networking and Mobile Computing 2007, DOI: 10.1109/WICOM.2007.636,2555-2558.

Hadir, A. , ZineDine, K. , Bahhouya, M. , El Kafi, J., ,2014 , An optimized DV-hop localization algorithm using average hop weighted mean in WSNs, Codes, Cryptography and Communication Systems (WCCCS), 2014 5th Workshop on, IEEE, 25 – 29.

Kulaib, A. R., Shubair, R. M. , Al-Qutayri, M. A., Jason, W. P. , 2015, Improved DV-hop localization using node repositioning and clustering, Communications, Signal Processing, and their Applications (ICCSPA), 2015 International Conference on, 10.1109/ICCSPA.2015.7081314,1-6.

Kanmaz, M., 2016, Etkin Kullanıcı Hareketliliği ile Kesin Konumlandırma, Yüksek Lisans, İstanbul Üniversitesi Fen Bilimleri Enstitüsü.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

OUTLIER DETECTION METHOD BY USING DEEP NEURAL NETWORKS

DOI: 10.17261/Pressacademia.2017.577
PAP-WCTIE-V.5-2017(15)-p.96-101

Olgun Aydin[1], Semra Erpolat Tasabat[2]
[1]Mimar Sinan Fine Art University. olgun.aydin@ogr.msgsu.edu.tr
[2]Mimar Sinan Fine Art University. semra.erpolat@msgsu.edu.tr

ABSTRACT

Detecting outliers in the data set is quite important for building effective predictive models. Consistent prediction can not be made through models created with data sets containing outliers, or robust models can not be created. In such cases, it may be possible to exclude observations that are determined to be outlier from the data set, or to assign less weight to these points of observation than to other points of observation. Lower and upper boundaries can be created to exclude outliers from the dataset, and models can be created using the data between those boundaries. In this study, it was aimed to propose a different perspective on outlier detection methods by creating upper bounds with the aid of deep neural networks using skewness, kurtosis and standard deviation values obtained from the dataset with trained models.

Keywords: Deep neural network, outliers, outlier detection, modelling, predictive modelling

1. INTRODUCTION

It is inevitable that there are deficiencies or some problems in the data obtained in real life. When analyzing data sets containing a large number of variables, researchers are faced with unusual observations that may have potentially harmful effects on the results of the problem and are called outliers. Since outliers increase the value of the error variance, they also have an impact on the power of statistical tests. While it is important to clear the data from the outliers, it is very important to determine the outliers and analyze them specially for some study areas. It is of utmost importance to examine outliers separately in situations such as weather forecasting, fraud detection, detection of unexpectedly responding patients to different types of drugs (http://www.cse.yorku.ca/~jarek/courses/6412/lectures/Outliers.ppt) Outliers can appear in almost any data set in any application domain. Due to measurement errors or momentary extreme conditions, the data set may contain outliers. For this reason, outlier detection is a research field that researchers place importance on.(Hawkins, 1980). In the scope of the study, a different method is proposed for the density based method which is a statistical based method used for outlier detection. With this motivation, data sets without normal outliers were created and outliers were added to these data sets in a controlled manner. It is aimed to determine outlier by using deep neural networks using skewness, kurtosis, standard deviation with generated data sets.

2. LITERATURE REVIEW

In this section, outlier detection methods used in the literature and summary information are given. In addition, summary information on deep neural networks is included.

2.1. Outlier Detection Methods

The method of outlier detection based on statistical distribution is based on the standard deviation method and the calculation of Interquartile Range (IQR) calculated before the boxplot is drawn. In the standard deviation method, values that fall outside the range of [Mean -2*Standard deviation, Mean +2*Standard deviation], or [Mean -3*Standard deviation,Mean+3*Standard deviation] are considered as outliers, if the data has normal distribution. In another method IQR, first quartile (Q1) followed by third quartile (Q3) is calculated. IQR is calculated from Q3-Q1. The data outside the

range [Q1 - 1.5 * IQR, Q1 + 1.5 * IQR] is marked as outlier.

One of the methods based on statistical tests is the Dixon method. In this method, the data set is sorted and the different test statistic is calculated according to the state of the smallest and largest value. The calculated test statistics is compared with the corresponding critical value to determine outlier values. Another method performed by statistical testing is Rosner. In this method, values are assigned in the range 0-9 according to the distance from the average starting from the farthest distance. 9 data is the most distant from the average. 0 is the closest to the average. The data considered to be outliers are removed from the individual data set and the average and standard deviation of each term is calculated. The test statistic is calculated with these calculated mean and standard deviations. These test statistics are compared with the critical values and it is decided whether the observation is outlier or not (Aggarwal 2013).

Clustering-based methods view small-sized clusters, including an observational dimension, as clustering outliers. Some examples for such methods include segmentation around medoids (PAM) and cluster large applications (CLARA) A modified version of these for spatial extensions, called CLARANS And a fractal size based method. Since the main objectives are clustering, these methods are not always optimized for outliers. In most cases, uncertain detection measures are implicit and can not be easily understood from clustering procedures (Ben-gal, 2005). Spatial methods are closely related to clustering methods. Lu et al. (Lu et al., 2003) define non-spatial values as spatially cited objects that differ significantly from local values (Ben-gal, 2005).

2.2. Deep Neural Networks

Deep Learning is a new field in Machine Learning research. Deep Learning; In general, studies that help to understand the meaning of images, sounds and texts. (http://deeplearning.net/tutorial/). Prior to 2006, attempts to train deep architects failed: a deeply supervised feedforward neural network training gives worse results (compared to one or two layers of hidden layer) (both training and test failure). In 2006, three studies led by Hinton's revolutionary work destroyed many memorabilia

- Hinton, G. E., Osindero, S. and Teh, Y., A fast learning algorithm for deep belief nets Neural Computation 18:1527-1554, 2006
- Yoshua Bengio, Pascal Lamblin, Dan Popovici and Hugo Larochelle, Greedy Layer-Wise Training of Deep Networks, in J. Platt et al. (Eds), Advances in Neural Information Processing Systems 19 (NIPS 2006), pp. 153-160, MIT Press, 2007
- Marc'Aurelio Ranzato, Christopher Poultney, Sumit Chopra and Yann LeCun Efficient Learning of Sparse Representations with an Energy-Based Model, in J. Platt et al. (Eds), Advances in Neural Information Processing Systems (NIPS 2006), MIT Press, 2007

The following three principles have been included in three studies.

- The learning of impressions as unchecked is used to train each layer.
- One layer at a time is being trained on the previously educated, unchecked.
- Use audited training to fine tune all layers. (http://www.iro.umontreal.ca/~pift6266/H10/notes/deepintro.html)

3. DATA AND METHODOLOGY

Controlled deviations were added to the data set generated from the normal distribution. In this process, we first derive 500 random numbers from the normal distribution with a mean of 10 and a standard deviation of 2. Random numbers were then derived from the normal distributions with averages of 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 and standard deviation 2. Then all these values are multiplied by 10 to obtain outliers. These operations were repeated 10,000 times. The mean, standard deviation, skewness, kurtosis values of data sets with outliers and without outliers were calculated and recorded at each step. All calculations were performed in R Studio environment by following script.

```
lbf <- NULL

ubf<- NULL

mxf <- NULL

dxf <- NULL

mxoutf <- NULL

dxoutf <- NULL
```

```
ubcf <- NULL
skwf <- NULL
krtsf <- NULL
for (i  in 1:10000)
{
 for(j in 1:10)
 {
        x <- rnorm(500,mean = 10,sd = 2)
        out_up<- abs(rnorm(50, mean = 10+5*j, sd = 2))*10
        x_out <- c(x,out_up)
        mxout <- mean(x_out)
        sdxout <- sd(x_out)
        mx<- mean(x)
        sdx<- sd(x)
        lb <- mx - 3*sdx
        ub <- mx + 3*sdx
        skw <- skewness(x_out)
        krts <- kurtosis(x_out)
        ubc <- (floor(min(out_up)) - mxout) / sdxout
        ubcf <- rbind(ubcf,ubc)
        lbf <- rbind(lbf, lb)
        ubf <- rbind(ubf, ub)
        mxf <- rbind(mxf, mx)
        dxf <- rbind(dxf, sdx)
        mxoutf <- rbind(mxoutf, mxout)
        dxoutf <- rbind(dxoutf, sdxout)
        krtsf <- rbind(krtsf, krts)
        skwf <- rbind(skwf, skw)
 }
}
```

Figure 1 shows the histogram of the 500 units data without outliers obtained from the normal distribution with a mean of 10 standard deviations of 2 in the above step. As you can see, the dataset is symmetrical.

Figure 1: Histogram of Variables without Outliers Generated from Normal Distribution

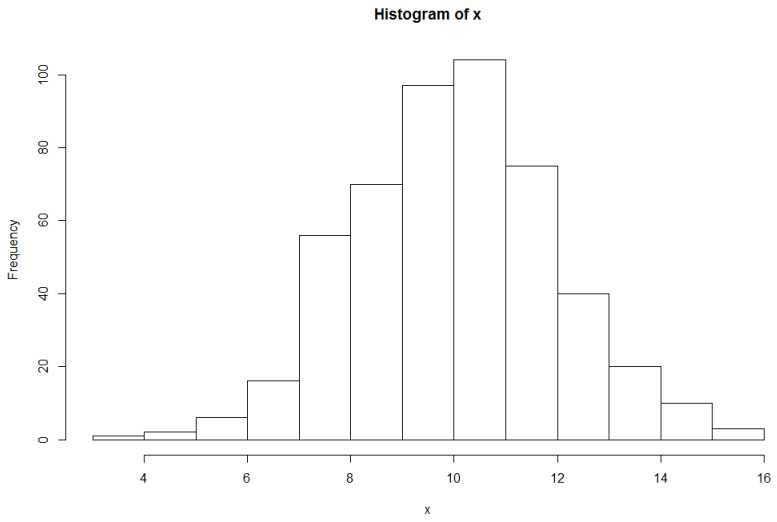
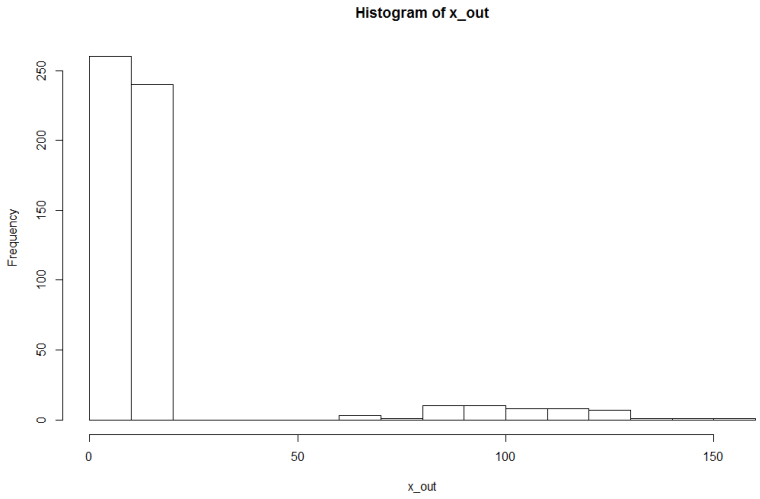


Figure 2 shows the histogram of 500 units of data added outlier to the data set obtained from the normal distribution with a mean of 10 standard deviations of 2 obtained in the above step. As you can see, the dataset is not symmetrical and is right skewed.

Figure 2: Histogram of Variables with Outliers Generated from Normal Distribution



As it is known that 99% of the data set with standard normal distribution is in the [Mean -3*Standard deviation,Mean+3*Standard deviation] With the help of this information, the outliers can be defined as those outside the range of [Mean -3*Standard deviation,Mean+3*Standard deviation]. +3. For this generalization, the upper limit coefficient can be defined as -3, and the lower limit coefficient can be defined as -3. If the data set is normal distribution, the data other than Mheat +/- 3 * standard deviation can be evaluated as outlier. However, the upper and lower limit coefficients are not defined as +3 or -3 when the data does not fit the normal distribution and the data is skewed to the right or to the left. These coefficients also vary according to the skewness, kurtosis and standard deviation of the data. If these coefficients can be estimated, values outside the range [Mheat -lower coefficient * standard deviation, Mheat + upper coefficient * standard deviation] can be defined as outlier. With this in mind, outliers were added to the data obtained from the normal distribution in a controlled manner and 100000 data sets were created to the right. Upper coefficients can also be calculated as it is known at which points the outliers are added. For this purpose, (floor (min (out_up)) - mxout) / sdxout formula is used. After all these calculations the final data set was created.

Table 1: Sample from Final Dataset (Independent Variables)

mxoutf	dxoutf	krtsf	skwf
22.52423	39.94923	6.767220	2.919109
26.88227	54.03466	6.659846	2.907539
32.05484	70.34434	6.342558	2.870150
36.33946	83.34157	6.283900	2.863794
40.65891	97.40719	6.206902	2.854884
45.31630	111.66793	6.167622	2.849995
49.81123	126.20862	6.145563	2.847488
54.20110	140.20838	6.147518	2.847924
59.38225	156.64846	6.129798	2.845660
63.70956	170.21744	6.116862	2.844120
23.05472	41.56753	6.844389	2.930907
27.27965	55.08199	6.617082	2.903642

A sample from the 100000x4 size final dataset is shown in Table 1. In the data set, mxoutf is about average of the data with outliers, dxoutf is about standard deviation of the data with outliers, krtsf is the kurtosis of the data with outlier, and skwf is the skewness of the data with outlier. Ubcf is about upper coefficient. In the study, ubcf is considered as a dependent variable, while dxoutf, krtsf, swkf are considered as independent variables. A deep neural network, LSTM, is used to detect long and short term dependencies. Before creating this neural network, 75% of the data is allocated as a train set and 25% as a test set. In train phase, epoch 50, batch size 30, mean squared error as loss function, and optimization algorithm as adam.

4. FINDINGS AND DISCUSSIONS

When the model training process is completed, it is seen that the final MSE is 0.0806. Figure 3 shows the variation of the loss function in each epoch in the train process. It was seen that the MSE was 0.0808 when the modeled train test set was applied.

Figure 3: Training Process

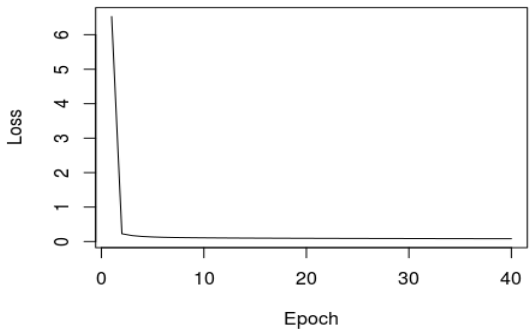


Table 2 also shows some estimation results by using . According to the results, it can be seen that the upper coefficients can be predicted consistently.

Table 2: Some Test Results

mxoutf	dxoutf	krtsf	skwf	ubcf (real)	ubcf (predicted)	Biass
-1,58549	-1,58345	3,20958	3,23297	2,99761	2,83683	0,16078
-1,21875	-1,21455	1,08475	1,05602	4,47566	4,62052	-0,14486
-0,85889	-0,86192	0,12233	0,13471	6,48482	6,65756	-0,17274
-0,51340	-0,50979	-0,21554	-0,20037	8,16181	8,43490	-0,27309
-0,17372	-0,17109	-0,27275	-0,28420	9,61548	9,69550	-0,08002
0,15581	0,16377	-0,55835	-0,55735	11,03645	11,63451	-0,59805
0,51716	0,51341	-0,54952	-0,55273	13,45952	13,39541	0,06412
0,87909	0,88744	-0,67357	-0,67475	15,16611	15,29549	-0,12938
1,21913	1,22310	-0,70991	-0,71117	16,88347	16,89834	-0,01487
1,54321	1,55068	-0,77332	-0,77348	18,91024	18,33228	0,57796

5. CONCLUSION

Before analyzing a dataset, it is very important to extract the dataset from the outliers. Models made with a data set containing outliers will not be able to achieve accurate information as a result of the inferences. In this study it is aimed to give a different perspective to outlier detection. As a result of the studies made, it is possible to detect the outliers and remove the dataset from the outliers through deep neural networks created using some descriptive statistics. In future studies, it is aimed to use other deep neural networks other than LSTM and to compare the results.

REFERENCES

Aggarwal, C.C. (2013), Outlier Analysis, Springer-Verlag New York

Hawkins, D. (1980), Identification of Outliers Chapman and Hall Hawkins, D. Identification of Outliers. Chapman and Hall.

http://www.cse.yorku.ca/~jarek/courses/6412/lectures/Outliers.ppt

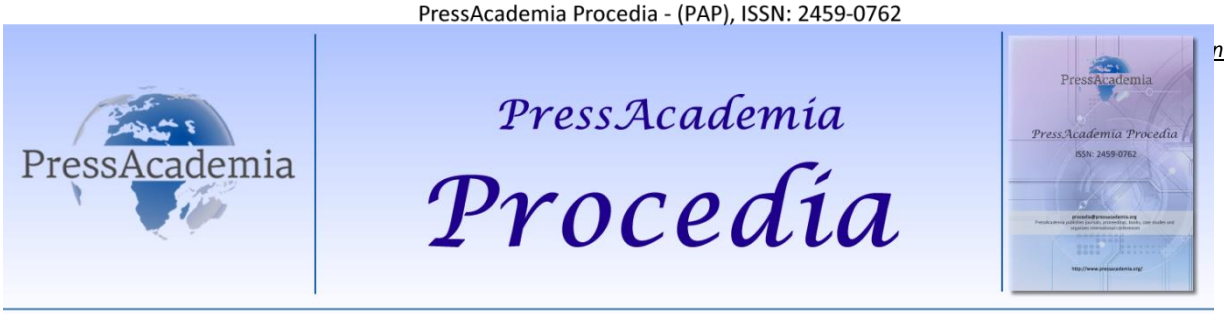
http://deeplearning.net/tutorial/

http://www.iro.umontreal.ca/~pift6266/H10/notes/deepintro.html

Ben-Gal, Irad. "Outlier detection." Data mining and knowledge discovery handbook (2005): 131-146.

Osborne, Jason W., and Amy Overbay. "The power of outliers (and why researchers should always check for them)." Practical assessment, research & evaluation 9.6 (2004): 1-12.

PressAcademia Procedia - (PAP), ISSN: 2459-0762



2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

DESIGN OF ZIGBEE BASED WIRELESS ONLINE MONITORING SYSTEM FOR PHOTOVOLTAIC POWER SYSTEMS

DOI: 10.17261/Pressacademia.2017.578
PAP-WCTIE-V.5-2017(16)-p.102-110

Hakan Bogan[1]
[1] Ege University, hb@hakanbogan.com

ABSTRACT

Within this project, it is intended to design and implement a web-based wireless photovoltaic monitoring system that can measure important parameters for photovoltaic power systems and perform performance calculations and comparisons of the modules. A web-based, low-cost wireless monitoring system using the Zigbee wireless communication protocol has been developed, taking into account the International Electrotechnical Commission (IEC) standards. Thanks to this developed Zigbee based monitoring system, the data in the power plant can be recorded without any loss without regard to weather conditions. It also has the necessary infrastructure for various emergency scenarios. At this point, the investor can access the system from any device with internet access, regardless of where in the world. As a result of the tests made, there is no data loss between the receiver and the transmitter in this developed system. Emergency systems, which are developed as solutions to electricity or internet interruption problems, prevent data losses.

Keywords: Zigbee, photovoltaic, monitoring, online, wireless sensors network.

1. INTRODUCTION

As the energy demand on the world increases with the growing world population, all countries attach great importance to the development of renewable energy day by day. People consider that the solar energy is most reliable and promising energy source among the renewable energy sources (Shariff, Rahim & Ping, 2015). There are many photovoltaic power plants that need to be managed and kept healthy state of these power plants by using information techniques. That is important point that needs to be dealt with for controlled and successful electricity generation (TaoHua, Zhenga, Tana, Zhua & Miao, 2015)

The Compound Annual Growth Rate (CAGR) of Photovoltaic installations was 42 % between 2000 to 2015. Concerning PV module production in 2015, China&Taiwan hold the lead with a share of 67 %, followed by Rest of Asia-Pacific & Central Asia (ROAP/CA) with 14%. Europe contributed with a share of 5 % (was 6% in 2014); USA/CAN contributed 3 %. In 2015, Europe's contribution to the total cumulative PV installations amounted to 40 % (compared to 48 % in 2014). In contrast, installations in China accounted for 21 % (compared to 17 % in 2014)(Philipps & Warmuth, 2016).

The total installed capacity of 1078 Photovoltaic Power Plants located in Turkey is 860,63 MW. In 2016, 1,020,000,000 kilowatt-hours of electricity generation was made with Photovoltaic Power Plants (Güneş Enerji Santralleri, n.d.). Turkey is a growing market with so much acceleration. It has the potential to become a leader in the world with accurate and conscious guidance.

Some environmental and electrical factors such as dust, dirt, bird droppings, weather conditions and accidents can reduce energy generation efficiency. Therefore, photovoltaic plant and power generation should be kept under constant surveillance. Constantly monitoring the system makes it possible to identify problems instantly and enables early intervention. So total production will be kept at the highest level (Woyte et al., 2013).

Wireless communication is used to facilitate data transmission in monitoring systems. Various types of wireless technology and networks allow devices to speak (send data) to each other and to the web (TCP/IP Networks) without cables. There are

many wireless communication protocols such as Bluetooth, ZigBee, WiFi and LTE used for wireless communication. Each protocol has its own advantages and areas of usage.

ZigBee is a wireless technology developed as an open global standard to address the unique needs of low-cost, low-power wireless Machine to Machine (M2M) networks. The ZigBee standard operates on the IEEE 802.15.4 physical radio specification and operates in unlicensed bands including 2.4 GHz, 900 MHz and 868 MHz. ZigBee enables broad-based deployment of wireless networks with low-cost, low-power solutions. It provides the ability to run for years on inexpensive batteries for a host of monitoring and control applications. Smart energy/smart grid, AMR (Automatic Meter Reading), lighting controls, building automation systems, tank monitoring, HVAC control, medical devices and fleet applications are just some of the many spaces where ZigBee technology is making significant advancements (XBee and XBee Pro, n.d.)

The analysis of photovoltaic power plants is made in the framework of the parameters determined in IEC 61724 standards and the necessary calculations are carried out. Wireless photovoltaic monitoring system using the Zigbee wireless communication protocol has been developed in this project, taking into account the International Electrotechnical Commission (IEC) standards.

2. LITERATURE REVIEW

In the past there are works related to data transfer and recording in photovoltaic power plants. Various monitoring systems were introduced and released. It seems that nowadays it is still necessary to develop the technology and reduce the costs in photovoltaic systems (Katsioulis, n.d.). Information about the data transfer technologies used in these studies can be found below.

2.1. Data Transfer

Past studies have been done in two ways as wired and wireless. RS232 or RS484 cables were generally used in wired systems and they are computer based (Anwari, Dom, & Rashid, 2011; Ayompe, Duffy, McCormack, & Conlon, 2011). Wired monitoring systems are not preferred because of their limitations and various negative aspects (Shariff, Rahim & Ping, 2015).

Monitoring systems include devices with various wireless data transfer technologies such as satellite (Krauter, 2004), GSM (Gagliarducci, Lampasi, & Podestà, 2007; Rosiek & Batlles, 2008) and ZigBee (Shariff, Rahim & Ping, 2015; Katsioulis, Karapidakis, Hadjinicolaou & Tsikalakis, 2011). Satellite technology has a high installation cost. Despite this, the data transmission speed is very low (Krauter, 2004). GSM is a really secure technology. Data transmission rate of GSM is approximate 100%. However, this technology has high operating costs (Rosiek & Batlles, 2008). Bluetooth may be a good alternative, but this technology has a short transmission range (Hua, Lin, Xu, Li, & Ouyang, 2009). Wi-Fi has very high data rate. However, cost and power consumption of this technology is high when compared to other technologies such as Bluetooth and ZigBee (López, Mantiñan & Molina, 2012). ZigBee has low data transfer rate when compared to others. However, it has huge network size, long range capacity, low power consumption and reliability (Shariff, Rahim & Ping, 2015). Therefore, ZigBee is popular.

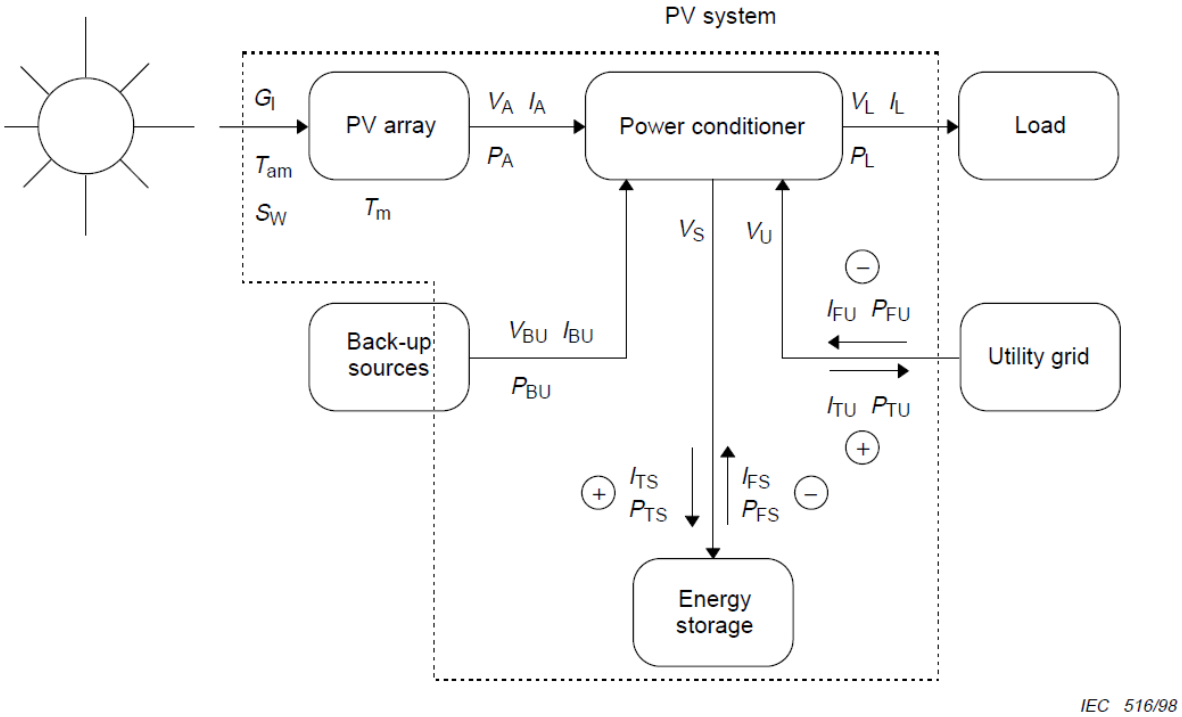
3. DATA AND METHODOLOGY

3.1. Parameters

The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). IEC promotes international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides.

Today, large investments are made in photovoltaic power plants. Monitoring these investments and production of plants is of great importance. It is necessary to make a conscious measurement taking into account the parameters to be measured specified in IEC 61724 standards.

Figure 1: The parameters to be Measured in Real Time



3.2. Materials and Equipment

3.2.1. Sensors

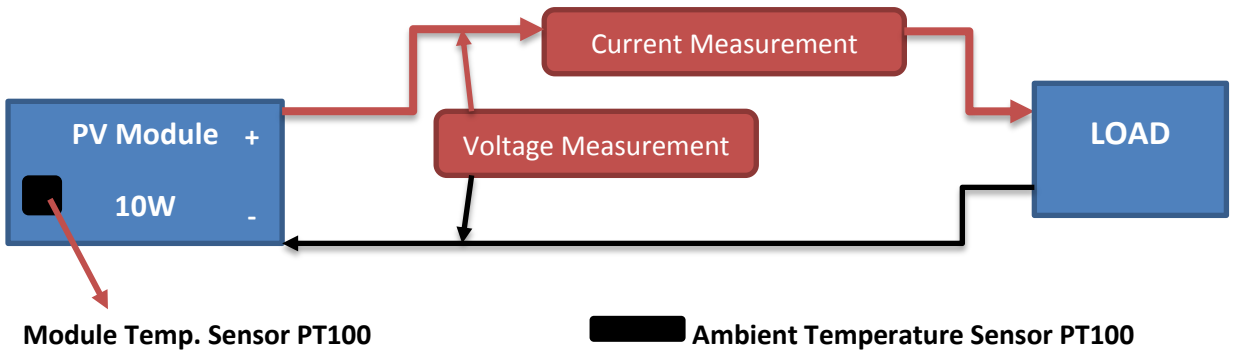
Some parameters should be measured such as module temperature, ambient temperature, irradiance, each current and voltage of components in real time.

2-wire PT100 temperature sensor is used for measuring PV module temperature in solar power plants. The PT100 was used in this project to measure the PV module temperature and also ambient temperature. This sensor can measure between -20°C and 100°C. These type of sensors are called resistance temperature detectors (RTDs).

ACS712 current sensor is used for current measurement. It is hall effect-based linear current sensor that can measure up to 30A. Its operating temperature range of -40°C to 85°C. Analog to digital conversion needs to be done to read value.

A voltage divider is used to measure the voltage. It is is a simple circuit which turns a large voltage into a smaller one. Using just two series resistors and an input voltage, an output voltage that is a fraction of the input can be created. Thus, the voltage can be measured easily. Voltage dividers are one of the most fundamental circuits in electronics. Analog to digital conversion needs to be done to read value.

Figure 2: PV Module Measurements



Tests are being done on 10W PV module. And also 25W and 11W aluminum and stone resistors are used as the load. To get that precision and accuracy out of PT100 RTD an amplifier that is designed to read the low resistance needs to be used. This situation also applies to the pryranometer.

Figure 3: Data Collection from the Sensors

Sensors connected to the microcontroller board measure data. This data is processed on the microcontroller and the wireless communication module is ready for transmission.

3.2.2. Data Processing

Arduino UNO was used as a end device board nearby the PV module and sensors to collect and process sensor information. It is a microcontroller board based on the ATmega328P. It has 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog inputs, a 16 MHz quartz crystal, a USB connection, a power jack, an ICSP header and a reset button.

ACS712 required for current measurement is connected in series to the output of the photovoltaic module. The data output pin of the sensor is connected to the analog input of the Arduino. To get a better measurement, 150 samples are taken and the average of all is taken afterwards.

A voltage divider is made using resistors of 30k and 7.5k. So that voltages in the range of 0-25v can be easily measured. The voltage divider output is connected to the analog input of the Arduino. The reference voltage was taken as 5 volts and the necessary operations were performed.

The output voltage of the PT100 RTD sensors is low and the amplifier board has been purchased to increase the output voltage. This card enhances the output signal of the sensor so that it allows precise measurement on the Arduino. The output of the amplifier board is directly connected to the digital input of the Arduino.

Arduino Mega 2560 was used as a coordinator board nearby the modem. It is a microcontroller board based on the ATmega2560. It has 54 digital input/output pins (of which 15 can be used as PWM outputs), 16 analog inputs, 4 UARTs (hardware serial ports), a 16 MHz crystal oscillator, a USB connection, a power jack, an ICSP header, and a reset button. Arduino Mega 2560 is preferred as a coordinator because the coordinator side needs much more capacity and RAM than the end device side. In addition, the data received from the end device is transferred through the Ethernet module on the Mega 2560. That Ethernet module has own SD Card module. So when the internet connection is interrupted, SD Card steps in and records the non-transmitted data.

3.2.2. Wireless Communication

ZigBee is a low-power wireless mesh network standard operating in the 2.4 GHz range. In this project, XBee modules were used to provide wireless data transmission. These modules are designed for a variety of uses, ranging from ZigBee to high-throughput, low latency applications. The XBee modules were engineered to meet IEEE 802.15.4 standards and support the unique needs of low-cost, low-power wireless sensor networks. The modules require minimal power and provide reliable delivery of data between devices. XBee used in this Project has 1.6 km outdoor data transmission range.

The most important reasons for choosing ZigBee in this project are; It supports IOT, provides energy efficiency and is reliable and scalable. ZigBee provides seamless interoperability among the widest range of smart devices related to home automation, energy management, and retail. ZigBee and ZigBee Pro consume 1/10,000th of the energy required by a Wi-Fi network. ZigBee is scalable, as it supports networks of thousands of nodes. Due to the use of security mechanisms like AES-128 encryption devices, along with network keys and frame counters, it is also one of the most secure devices out there.

Table 1: Comparison of the WiFi, Bluetooth and ZigBee Protocols

Category	WiFi	Bluetooth	ZigBee
Underlying Standard	802.11.b	802.15.1	802.15.4
Range (meters)	1 - 100	1 - 10+	1 – 100+
Data Rate	11 and 54 Mbits/s	1 Mbit/s	20, 40 and 250 kbits/s
Power Consumption	High	Low	Very Low
Frequency	2.4 and 5 GHz	2.4 GHz	868 MHz, 916 MHz, 2.4 GHz
Network Nodes	50	8	65535
Security	SSID	64 bit, 128 bit	128 bit AES
Integration Level & Reliability	Normal	High	High

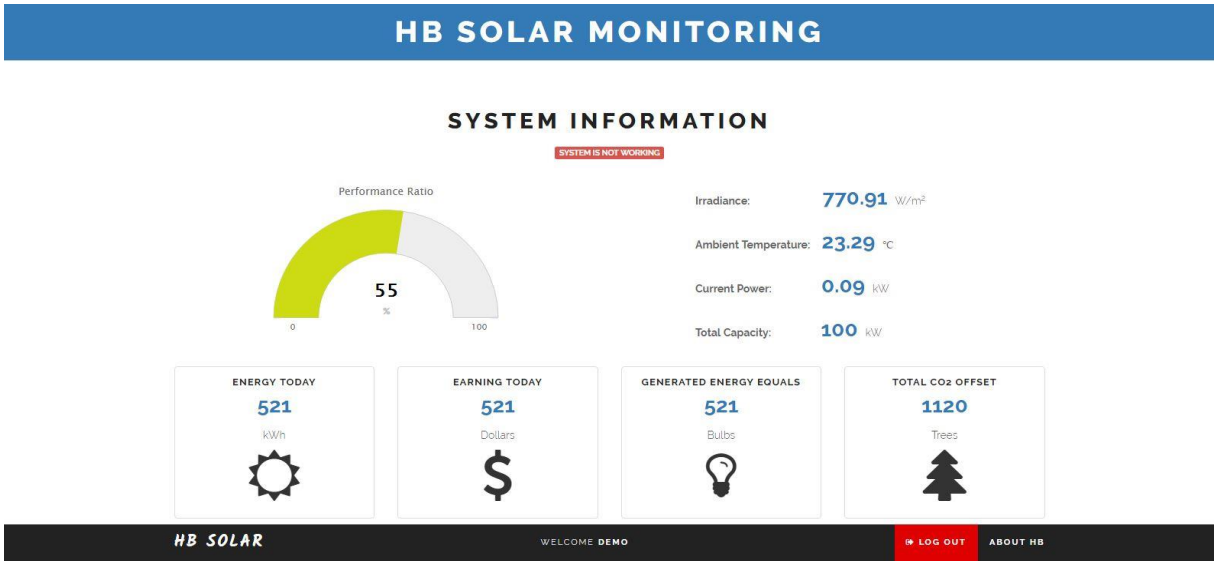
Cost of Terminal Unit	High	Low	Low
Prime Cost	Normal	Low	Low
Optimized for	Speed	Low Cost, Convenience	Reliability, Low Power Consumption, Low Cost, Scalability
Typical Applications	Wireless Local Area Network (WLAN) connectivity, broadband internet, security cameras	Inter-device wireless connectivity, e.g., phones, PDAs, laptops, headsets, cameras, serial cable replacements	Industrial control and monitoring, sensor networks, building automation, toys

The end device and the coordinator have their own XBee devices. The end device takes measurements every 2 minutes according to IEC standards and transmits the stored data wirelessly to the coordinator device. Moreover, thanks to continuous control, internet connection, wireless communication line and sensors are kept under control. The warnings are displayed on the LCD screen.

3.3. Monitoring Dashboard

The cloud hosting service was purchased to record the data of this system. Cloud hosting services provide hosting for websites on virtual servers that pull their computing resource from extensive underlying networks of physical web servers. Broadly speaking the client can tap into their service as much as they need, depending on the demands of their website, and they will only pay for what they use. All these services have reliability, physical security, scalability and flexibility, utility style costing and responsive load balancing.

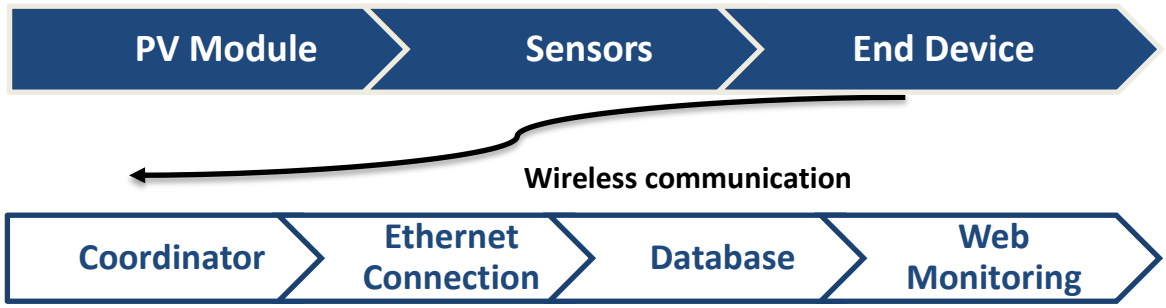
Figure 4: Admin Dashboard of the Project (solar.hakanbogan.com)



Monitoring dashboard is created from scratch with the bootstrap framework. Bootstrap was created at Twitter and it has become one of the most popular front-end frameworks and open source projects in the world. That monitoring dashboard has a completely responsive design. This means that it is compatible with all phones, tablets and computers. A login system is available and each investor has his / her own private page. Output of all data can be exported as xls, csv and pdf file from the web page. Thanks to this interface, investors can easily manage the photovoltaic power plant reports.

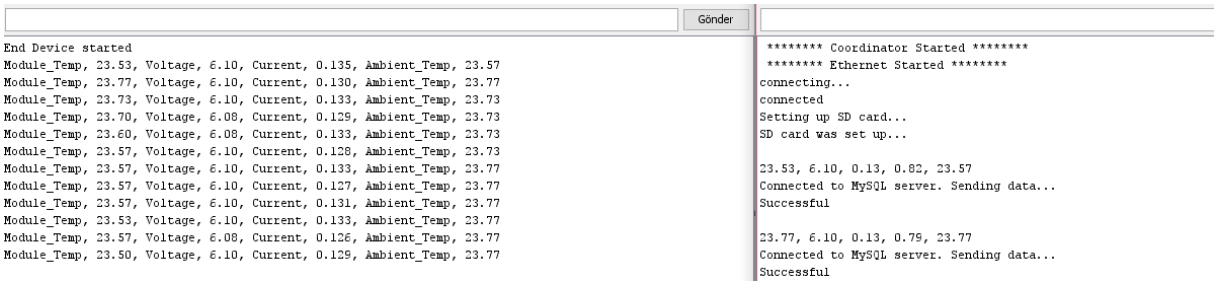
3.4. General Process

Figure 5: Block Diagram of the Web-Based Wireless Monitoring System



4. FINDINGS AND DISCUSSIONS

Figure 6: Wireless Communication between the Coordinator and End Device in the Serial Monitor

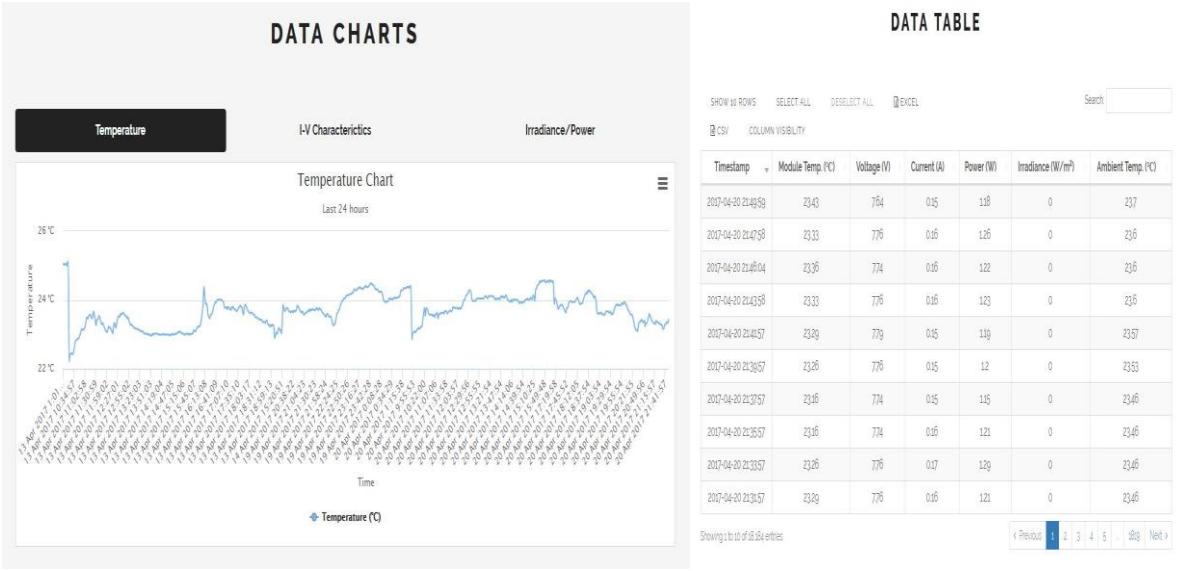


As a result of the tests, the system is able to transfer data in different weather conditions with no problems. Wireless transmission of data reaches 1.6 km (outdoor). The transmitted data was received by the receiving device with 100% and no data was lost. Within a variety of emergency scenarios, the device has continued to work successfully and record data. Thanks to the SD Card on the device, data loss is prevented.

Figure 7: Data in the Database

	id	timeStamp	temperature	voltage	current	power	irradiation	amb_temperature
Düzenle Kopyala Sil	18599	2017-04-20 20:42:09	23.4	6.64	0.15	0.97	0	23.6
Düzenle Kopyala Sil	18600	2017-04-20 20:43:56	23.43	6.64	0.14	0.95	0	23.57
Düzenle Kopyala Sil	18601	2017-04-20 20:47:56	23.36	6.64	0.14	0.93	0	23.53
Düzenle Kopyala Sil	18602	2017-04-20 20:49:56	23.36	6.64	0.14	0.92	0	23.57
Düzenle Kopyala Sil	18603	2017-04-20 20:51:56	23.4	6.64	0.14	0.94	0	23.57
Düzenle Kopyala Sil	18604	2017-04-20 20:53:56	23.23	6.64	0.14	0.94	0	23.46
Düzenle Kopyala Sil	18605	2017-04-20 20:56:02	23.33	6.64	0.14	0.93	0	23.46
Düzenle Kopyala Sil	18606	2017-04-20 20:57:56	23.33	6.64	0.14	0.92	0	23.5
Düzenle Kopyala Sil	18607	2017-04-20 20:59:56	23.4	6.64	0.15	0.97	0	23.53
Düzenle Kopyala Sil	18608	2017-04-20 21:01:56	23.5	6.64	0.14	0.94	0	23.7
Düzenle Kopyala Sil	18609	2017-04-20 21:03:56	23.6	6.64	0.13	0.88	0	23.67
Düzenle Kopyala Sil	18610	2017-04-20 21:05:56	23.53	6.64	0.15	0.97	0	23.63
Düzenle Kopyala Sil	18611	2017-04-20 21:07:56	23.46	6.64	0.14	0.92	0	23.6
Düzenle Kopyala Sil	18612	2017-04-20 21:09:57	23.4	6.64	0.14	0.96	0	23.57
Düzenle Kopyala Sil	18613	2017-04-20 21:12:03	23.36	6.64	0.14	0.94	0	23.57
Düzenle Kopyala Sil	18614	2017-04-20 21:13:57	23.33	6.64	0.14	0.95	0	23.53

Figure 8: Data Table and Temperature Chart of the Tests



All the data received by the coordinator is sent to the hakanbogan.com database via the modem to which the ethernet cable is connected. Very large data can be stored in the database. So there is no problem of data logging for many years. More than 18,000 measurements were made in the tests and the data of each was recorded with no problems. Data from any date can be exported as xls, csv or pdf through the monitoring dashboard. It has also been tested.

Figure 9: End Device and Coordinator

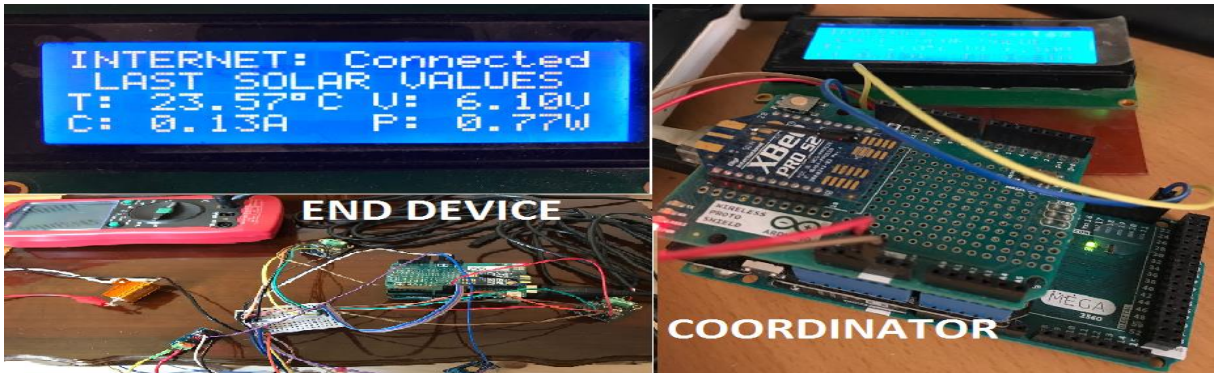
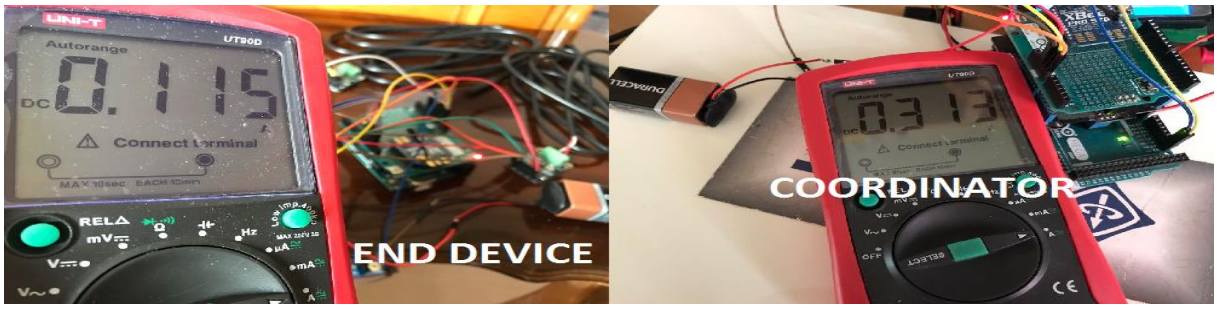


Figure 10:. Current Draws of the End Device and the Coordinator



As seen in Fig. 10, power consumption is very low. These current values were measured using a 9v battery. The following calculations related to monthly electricity consumption can be made.

Battery has 8.32V while the end device is connected and has 7.47V while the coordinator is connected.

End Device: 0.115A x 8.32V = 0.957 W (DC) Coordinator: 0.313A x 7.47V = 2.338 W (DC)

0. 957 W x 24h = 22.963 Wh/day (DC) 2.338W x 24h = 56.115 Wh/day (DC)

22.963Wh/day x 30 = 688.896 Wh/month (DC) 56.115Wh/day x 30 = 1,683.4392 Wh/month (DC)

Total consumption: (688.896) + (1,683.4392) = **2,372.3352 Wh/month (DC)**

The photovoltaic array, pryranometer, anemometer (optional) measurements must also be added to these measurements for the device to be adapted to real life. In order to further reduce power consumption and production cost, it is necessary to design the microcontroller card from scratch. Then a suitable box must be designed for the device.

5. CONCLUSION

Zigbee based wireless online monitoring system for photovoltaic power systems was developed and will continue to be developed. The system design features easy-to-obtain hardware and free software, making it accessible to any researcher. IEC 61724 standards were taken into consideration while the system was being developed. In the tests, 10W solar panel, 25W aluminum resistors, temperature, voltage and current sensors were used as photovoltaic equipment. In particular, it has been tested to ensure that the software used on the device and all electronic equipment work properly with each other. The system transmits each measured data wirelessly without any loss. The web page also displays all data without any problems. All data can be kept in the database. In emergency scenarios, data loss is prevented by recording the data to the SD card with continuous checks. Power consumption is also noticeably low. As a result, the system is working successfully.

In the future, this system needs to be developed. One of the most difficult stages to measure the pyranometer. Studies are running to make precise measurement of the pyranometer. When theoretically every stage is completed, the experiments will continue at a real photovoltaic power plant for a long time. Moreover, in the future, an unique card design will be made instead of Arduino. The device software will be updated with an emphasis on low power consumption. Thus, the cost and power consumption of the product are planned to be further reduced. The design will be improved by designing a box suitable for the device. Monitoring dashboard will also be updated to make it more functional.

REFERENCES

IEC 61724, 1988, “Photovoltaic system performance monitoring guidelines for measurement, data exchange and analysis”, International Electrotechnical Commission (IEC).

Shariff F., Rahim N. A. and Ping H. W., 2015, “Zigbee-based data acquisition system for online monitoring of grid-connected photovoltaic system”, Expert Systems with Applications, vol. 42, no. 3, pp. 1730–1742.

Rosiek S., Batlles F.J., 2008, “A microcontroller-based data-acquisition system for meteorological station monitoring”, Energy Conversion and Management, vol. 49, no. 12, pp. 3746-3754.

Hu T., Zheng M., Tan J., Zhu L. and Miao W., 2015, “Intelligent photovoltaic monitoring based on solar irradiance big data and wireless sensor networks”, Ad Hoc Networks, vol. 35, pp. 127–136.

Hua Jianfeng, Lin Xinfan, Xu Liangfei, Li Jianqiu and Ouyang Minggao, 2009, “Bluetooth wireless monitoring, diagnosis and calibration interface for control system of fuel cell bus in Olympic demonstration”, Journal of Power Sources, vol. 186, pp. 478-484.

Zahurul S., Mariun N., Grozescu I. V., Tsuyoshi H., Mitani Y., Othman M. L., Hizam H. and Abidin I. Z., 2015, “Future strategic plan analysis for integrating distributed renewable generation tos mart grid through wireless sensor network: Malaysia prospect”, Renewable and Sustainable Energy Reviews, vol. 53, pp. 978–992.

Fuentes M., Vivar M., Burgos J.M., Aguilera J. and Vacas J.A., 2014, “Design of an accurate, low-cost autonomous data logger for PV system monitoring using Arduino that complies with IEC standards”, Solar Energy Materials & Solar Cells, vol. 130, pp. 529–543.

Meliones A., Apostolacos S. and Nouvaki A., 2014, “A web-based three-tier control and monitoring application for integrated facility management of photovoltaic systems”, Applied Computing and Informatics, vol. 10, pp. 14–37.

Papageorgas P., Piromalis D., Antonakoglou K., Vokas G., Tseles D. and Arvanitis K. G., 2013, “Smart solar panels: in-situ monitoring of photovoltaic panels based on wired and wireless sensor networks”, Energy Procedia, vol. 36, pp. 535–545.

Almaliky A. (2014). Development of wireless monitoring system for photovoltaic panels. Retrieved from http://researcharchive.wintec.ac.nz/3362/

Katsioulis V. (2011). Design of a wireless monitoring system based on the zigbee protocol for photovoltaic systems (Doctoral dissertation). Retrieved from http://bura.brunel.ac.uk/handle/2438/5189

Batista N.C., Melicio R., Matias J.C.O. and Catalao J.P.S., 2013, "Photovoltaic and wind energy systems monitoring and building/home energy management using zigbee devices within a smart grid", Energy, vol. 49, pp. 306–315.

Andreoni M., Mantinan F. G., and Molina M.G. (2012). Implementation of wireless remote monitoring and control of solar photovoltaic (PV) system, Latin America Conference and Exposition, doi: 10.1109/TDC-LA.2012.6319050

Gagliarducci M., Lampasi D.A. and Podesta L., 2007, "GSM-based monitoring and control of photovoltaic power generation", Measurement, vol. 40, 314-321.

Ayompe L.M., Duffy A., McCormack S.J. and Conlon M., 2011, "Measured Performance of a 1.72 kW Rooftop Grid Connected Photovoltaic System in Ireland", Energy Conversion and Management, vol. 52, pp. 816-825.

Woyte A., Richter M., Moser D., Mau S., Reich N. and Jahn U., 2013, "Monitoring of photovoltaic systems: good practices and systematic analysis", The 28th European PV Solar Energy Conference and Exhibition, doi: 10.4229/28thEUPVSEC2013-5CO.6.1

Philipps S. and Warmuth W., (2016). Photovoltaics Report. Retrieved from Fraunhofer ISE Website: https://www.ise.fraunhofer.de/content/dam/ise/de/documents/publications/studies/Photovoltaics-Report.pdf

Krauter S.C.W. and Depping T., 2004, "Remote PV-system monitored via satellite", Solar Energy Materials & Solar Cells, vol. 82, pp. 139-150.

Anwaria M., Domb M.M. and Rashid M.I.M., 2011, "Small Scale PV Monitoring System Software Design", Energy Procedia, vol. 12, pp. 586-592.

Xiaolia X. and Huanb W., 2011, "The Wireless Sensor Network Construction of the Photovoltaic Power System——ZigBee", Procedia Engineering, vol. 15, pp. 2511-2515.

Katsioulis V., Karapidakis E., Hadjinicolaou M., and Tsikalakis A. (2011). Wireless Monitoring and Remote Control of PV Systems Based on the ZigBee Protocol, IFIP Advances in Information and Communication Technology, doi: 10.1007/978-3-642-19170-1_32

Rapacioli G. (n.d.). Reti wireless - Zigbee. Retrieved from http://www.swappa.it/wiki/Uni/RW-13Aprile

Why Zigbee is the Fastest Growing Trend in Wireless Technology. (2016, July 7). Retrieved from https://www.technavio.com/blog/why-zigbee-fastest-growing-trend-wireless-technology

XBee and XBee Pro Zigbee (n.d.). Retrieved from https://www.digi.com/pdf/ds_xbee_zigbee.pdf

Güneş Enerji Santralleri (n.d.). Retrieved from http://www.enerjiatlasi.com/gunes/

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurshi
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

REGARDING THE QUALITY ORIENTED DEVELOPMENT AND AN OVERVIEW OF SUSTAINABILITY OF THIS APPROACH FOR INDUSTRIAL/SERVICE SECTORS: A CASE STUDY FOR SUPPLY CHAIN AND PRODUCTION SYSTEM OF XB FIRM

DOI: 10.17261/Pressacademia.2017.579
PAP-WCTIE-V.5-2017(17)-p.111-123

Ayşenur Erdil[1], Mehtap Erdil[2]
[1]Marmara-Yalova University, erdil.aysenur@gmail.com
[2]Istanbul University, erdilmehtap@gmail.com

ABSTRACT

Six sigma provides to reduce cost by improving productivity and this method achieves this by focusing on the reduction of both defects and non-value added facilities. This method aims to reach zero defect in the theory improving supply chain processes. For extending and improving the effectiveness of supply chains, the critical processes of these systems must be developped. when the amount of users increase in the supply chain, processes and steps of systems will be more complex. There is an effective and usable method named as six sigma. This perspective aims to provide zero defect in the theory developping stages of supply chain. In the practical and implementation life, it is impossible because of changes, diversification in the life. This methodology has been chosen as a quality improvement technique to solve the problem in a production and supply chain system for industrial sector. By this way sustainability of Supply chains' and production processes can be provided and improved continuously. This study aims to present literature survey and overview about this quality method, six sigma and supply chain. This research includes some implementation stories, backgrounds of firms of services and industrial sectors. At the end of this study, it presents a case study of XB company and some suggestions can be offered about the case study according to the results.

Keywords: Quality oriented, six sigma, sustainability, supply chain, production

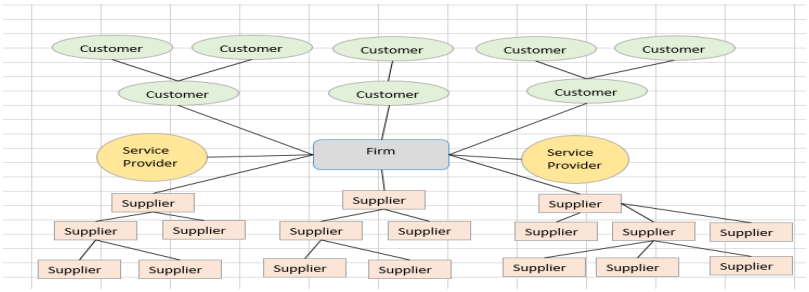
1. INTRODUCTION

Supply chain management (SCM) includes these important activities. These are sorted like this: delivering the right product to the right place, at the right time and at the right price is one of the most significant factors of business modification. Nowadays, it is one of the influential expense saving and profits sweetening strategies (Url1). According to the other definition, Supply chain management (SCM) is the process of planning, applying, and controlling the operations of the supply chain with the destination to resolve and supply customer requirements as expeditiously as possible. SCM comprehends the planning and management of all facilities including in sourcing, transition, and logistics management facilities (Url1; Url 2).

At the end stage of the continous of trends that started off from Business Process Reengineering, and Total Quality Management that have followed only the internal workings of an association, SCM purposes at combining the interval systems' structure of company to these of its suppliers, sharers and consumers (Erol and Ferrell, 2004). Technologies such as the Internet, transportation and software of warehouse management, programs of production and plant scheduling and also software that supports demand forecasting, SCM a adaptable strategy to acquire (Erol and Ferrell, 2004; Gunasekaran and Ngai, 2004).

The facilities in a supply chain regularly involve suppliers, customers, manufacturing-production centers, store houses, distribution locations and retail outlets (see figure 1). These facilities are interconnected via transportation and communicating relation on which raw materials, work-in process, inventory, end products, goods, and information flow stream structure. The supply chain comprises multiple business that occasion efficiently like a single firm with whole required information availability (Sheikh, 2003).

Figure 1: Framework of Basic Supply Chain (Erol and Ferrell, 2004)



Generally, the focus of supply chains is on specific functionalities such as purchasing, manufacturing, and shipping to support logistics operations. Presently, the effective management of technology and quality is the key point to increased quality level and supported competitive situation in global environment. Supply chain quality is a key component in achieving competitive advantage. The combination of logistics functions into integrated supply chain systems has increased the requirement for improved process quality. Improving the quality of all supply chain processes results in decreased costs, improved resource utilization, and improved process efficiency (Wing et al., 2005). The competitive environment of the 21th century requires the delivery of expense, effectiveness, high service levels, high quality of products and services (Benita et al., 1998; Wing et al., 2005). Some research studies have examined how quality management can be applied in SCM to develop the execution of the whole supply chain network. The continuous improvement issues accomplished by Deming, Juran, Feigenbaum and Crosby have developped overview for the measurement of supply chain management. Lately, the Six-Sigma development methodology has become extremely significant and popular. A review of literature and survey present that Six-Sigma methodology has been applied in scopes as human resource functional components and supplier-customer improvement. Six-Sigma could be applied to improve a structural model for appraisal, developing and controlling quality level in the supply chain network. Development in the quality of processes and stages of all supply chain influences to the simultaneous reducing of expenses, costs and sweetening of service (Wang et al., 2004). Six sigma seeks to develop the quality of the output of a process by defining and removing the reasons of defects and minimizing variability in processing and business processes. It needs a set of quality management technics, mainly empirical, statistical technics, and occurs a special infrastructure of human within the organization who are specialists in these technics (Tennant, 2001; Url 3). For organizations to stay competitive, they must continously reduce costs via improved productivity. Six Sigma achieves this by focusing on the reduction of both defects and non-value added activities. Its importance has been increased recently and it has serious effects on a company's success (Tennant, 2001). Therefore, in this study, six sigma has been chosen as a quality improvement philosophy to solve the problems in production system of a pipe application. The research topic is about the application of six sigma in usage and re-machining processes of the pipe production. During the research, rolls used in tube machines were taken into account as the main subject. This report summarizes the work that is done during the research. In the first part, objectives and literature survey will be given about Six Sigma tools. Then the system and its problems will be examined. Define-Measure-Analyze-Improve-Control (DMAIC) methodology of Six Sigma will be followed step by step. Priorities of all problems will be defined, the goals and objectives will be specified. Then collected data about current processes for future comparison will be presented. Analysis will be made about collected data and problems. At last, suggestions will be offered about the new system and the research will be concluded. Objectives in this research are to understand the six sigma approach and its tools in detail with the help of the literature survey, then making a successful pilot study in a firm that Six Sigma culture by appliying these tools on the related research topic.

2. LITERATURE REVIEW

2.1. General Background of Six Sigma

To understand the concepts in Six Sigma better and to apply accurately, literature survey has been made from different kinds of sources. Six Sigma methodology, its history and steps were investigated. The results of the survey were presented.

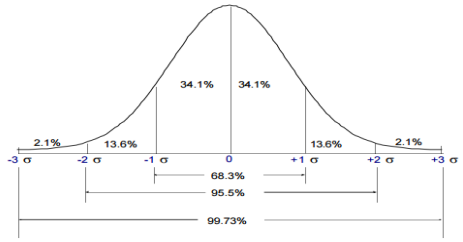
2.1.1. What is Six Sigma?

Six Sigma (SS) allows comparison of products, of varying complexity on basis. Also, it provides a common basis for benchmarking (competitors and non-competitors). The higher the sigma level, the better your operation is performing (Url 3, see Figure 2).

A standard deviation is the square root of the average squared differences between a set of data points and their common average according to the higher the sigma level and the lower the number of detects. It is less important, however, to understand the mechanics of the calculation than it is to understand what it means for a process to be operating at a certain sigma level (Url 3). Six Sigma is a classified and systematic technique for development of strategic stages and service

development which depends on statistical techniques and also the scientific method to support considerable deduction in defect ratio of consumer's voice. A significant stage in steps of Six Sigma development determines exactly what the consumer needs and specifies failures and defects in parts of requirements of customers critical to quality factors (Linderman et al., 2003; Url 2; Url 3).

Figure 2: Six Sigma Levels (Url 3)



Use of these unique metrics provides to explain objectives and support them well-defined. However, sometimes main data may not obtain for the system approach, as often develops with a new product, which causes it difficult to provide specific targets. In the steps of process development, the method is formed after the plan, do, check, act (PDCA) cycle. One popular technique utilizes define, measure, analyze, improve and control (DMAIC) as the five steps for the process development. Design step for Six Sigma is applied for incremental product design (define, measure, analyze, design and verify). Data and objective measurement have vital effect at each stage of the technique. Six Sigma guidelines determine a combination of proper tools at each step of the method. This careful integrating of tools with the statistical methods is unsurpassed to Six Sigma (Linderman et al., 2003; Url 2; Url 3).

2.2. Background of Six Sigma

The Greek Letter Sigma (α) has been the universally accepted symbol for standard deviation. Standard deviation is a measure of variation or spread. The quality tools and techniques that were named "Six Sigma" in 1986 by Motorola are not new. Six sigma(SS) occurred by Motorola and popularized by General Electric defines to a technic and set of tools that utilize statistical analysis to evaluate and develop an organization's performance, implements, and systems with prime aim of defining and removing difference to development quality. They were developed over the past 50 or more years, through the work of quality experts such as Deming Juran and others (Raisinghani, 2005; Url 4). Apart from mathematical field, SS is a company-wide transformation that has helped them and brought them for few years already, the success according to employees of Motorola, General Electric, Honeywell, Bambardier, Black and Decker, ABB, Polaroid and many other companies. The companies listed above can be reffered to as the "six sigma companies". Many other organizations have adopted a SS approach, but these particular companies have rejected their success and have emphasized the part played by SS in the achievement of this success. None more so than General Electric, as its Annual Report for 1998 states very clearly:

"… we plunged into SS with a company-consuming vengeance just over three years ago. We have invested more than a billion dollars in the effort, and the financial returns have now entered the exponential phase more than three quarters of a billion dollars saving beyond our investment in 1998, with a billion and a half in sight for 1999." (Url 5). As a result, SS has adopted, in a small number of large companies, benefits that present high interest to managers and to shareholders. One of these common characteristics of SS is a widespread focus on processes and the existence of a company-wide language for describing the capability of processes. The expected outcomes of SS efforts are faster and more robust product development, more efficient and capable manufacturing processes, more confident overall business performance (Sanders and Hild, 2000; Url 6; Url 7). The fundamental objective of the six sigma methodology is the implementation of a measurement-based strategy that focuses on process improvement and variation reduction through the application of Six Sigma improvement studies. This is accomplished through the use of two six sigma sub-methodologies: DMAIC and DMADV (Url 8).

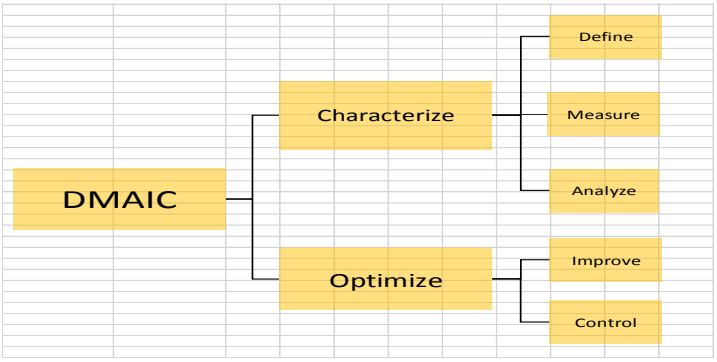
DMADV: Define, Measure, Analyze, Design, Verify (for new product improvement)

DMAIC: Define, Measure, Analyze, Improve, Control (for process improvement). The new approach to six sigma developed by Mikel Harry and Richard Schroeder (2000) and packaged them in the Define, Measure, Anlyze, Improve and Control (DMAIC) methodology. The DMAIC approach consists of management involvement, organizational structure to facilitate the improvement, customer focus, opportunity analysis, extensive training and reward and recognition for successful problem solving. Its benefits include the standardization of the methods, global adaption of the methodology (Gupta, 2004).

DMAIC has five phases including:

These phases are seperated into two groups as shown in Figure 3:

Figure 3: Phases of DMAIC (Url 9)



The six sigma companies are concerned with the reduction of variation and with the reduction of detective products and services. However, they often focus on "defects per million opportunities" (DPMO) or DPMO as metric and their performance standards are:

Table 1: DPMO Table (XB Holding, Black belts Education Documents)

SIGMA	DPMO	COPQ	CAPACITY
6 Sigma	3,40	<10% of sales	World Class
5 Sigma	230	10 to 15% of sales	
4 Sigma	6.200	15 to 20% of sales	Average Industry
3 Sigma	67.000	20 to 30% of sales	
2 Sigma	31.000	30 to 40% of sales	Noncompetitive
1 Sigma	700.000		

It is seen on Six Sigma Process Capability Chart above (in Table 1); as the sigma level increases, DPMO decreases, COPQ (Cost of poor quality) decreases and capability gets stronger.

2.3. Relevant Tolls/Concepts:

They are many tools and concepts used in Define phase as below:

-Project Charter –Brainstorming –Voice of the customer (VoC) – Pareto Analysis –Process mapping –SIPOC (Suppliers, Inputs, Process, Outputs, Customers) -IPO (Input, Porcess, Output)

Business Porcess Mapping

It helps to understand baseline better by using mapping techniques as below:

-Completing verifying and validating hgh-level (not 'should be' or 'could be') business porcess map.

-Completing SIPOC representation, describing the Suppliers, Inputs, Process, Outputs and Customers.

Customers: It requires to determine customers such that:

-Identifying customers and segmenting according to their different needs and requirements.

-Collecting data and displaying to better understand customer8s) critical needs and requirements.

Project Charter: The main items are business case, problem and goal statements, project scope, milestones, roles and responsibilities, project-research plan.

Opportunity Statement: The opportunity statement should adress these questions:

-What is wrong or not working? -When and where do the problems occur?

-What is the impact "pain" on our employees? -How extensive is the problem?

-What is the impact "pain" on our customers? -What is the impact "pain" on our business?

-Does it make strategic sense to adress this issue?

The opportunity statement does not make assumptions that cannot be supported with data, that will not create value, contain or identify a solution. Six Sigma attempts to identify the root cause of the problem, artificially define the scope too narrowly or define the scope too broadly.

2.4. The culture of Six Sigma

A significant feature of six sigma culture is the existence of what could best be described as management by fact. The performance of the process at the start of the project is recorded in a run chart or a control chart. At the completion of the project, the improvement is clear. The charts prominently displayed alongside other diagrams and a clear explanation of what is being mesured and how the chosen improvement actions were arrived at. There is evidence of a readily accepted self-discipline in this style of communication, a discipline which often surprises visitors and alarms suppliers.

2.4.1. Implementation of Six Sigma Strategies

There are three general policies of six sigma improvement. They are six sigma engineering organization, six sigma organization and strategy selection for individuals and research areas.

Six sigma organization uses a methodology of training all individuals, in all scopes, in the fundamental concepts and tools included in the implementation of six sigma. Six sigma engineering organization strategy focuses on training and improvement within the design and manufacturing engineering departments of the business.

The third classification strategy selection refers to the development on strategically-selected employees (Sabders and Hild, 2000). As with other quality programs such as TQM, determining Six sigma varies between organizations. For some companies, applying six sigma methodology means solely using the specific problem-solving stages involved; others use one or more of the strategies (Ingle and Jacobs, 1999).

2.4.2. Roles and Responsibilities:

Many organizations fail to execute programs such as Six sigma because they attempt to implement too many of them. Six sigma requires strong executive sponsorship and support from functional leaders. Incentives must be created so that “Line Managers” consistently. Some of them are that

-devote adequate resources to research teams, seek out potentiall high effect research ideas. Leaders must be identified to become black belts and succession plans need to allow for the repatriation of these black belts into the business.

2.5. The Requirements of Six Sigma

Strong leadership and top management commitment

Good support from top management is imperative in the restructuring of the organisation and achieving the cultural change and motivation of employees towards quality and the Six Sigma strategy to the business.

Cultural change

Employees have to be motivated towards the introduction and development of Six Sigma program through various reward and recognition schemes. There can also be a problem of employees dismissing Six Sigma as the latest fad or hype. To overcome this problem and also to allay the fears that employees may have, there has to be early and effective communication to all employees on the why and how of Six Sigma.

Training

Training is a crucial factor in the successful introduction and development of Six Sigma program. It is important to communicate both the ‘why’ and ‘how’ of Six Sigma as early as possible and provide the opportunity to people to improve their comfort level through training classes. There is usually a hierarchy of expertise, which is identified by the belt system. Within GE, the belt system is fundamentally divided into:

Campions: These are trained in the essentials of the six sigma methodology focusing on selecting the research that is aligned with business goals. Identify strategically important projects for the improvement teams and provide resources, typically receive an orientation to Six Sigma rather than detailed training, promoting and leading the Six Sigma deployment in significant or critical areas of the business Champions must select and mentior six sigma project leaders callled “Belts” and align and integrate the six sigma launch into their organization (Knowles et al., 2005).

Master Black Belt (MBB): This concept is often the key support person to the champions to assist in research defininition and the mentor and lecturer of six sigma green belts and six sigma black belts. It requires extensive research management experience according to the six sigma methodology and statistical tools. fully trained quality leaders responsible for Six

Sigma strategy, training, mentoring, deployment and results generally serve as instructors and internal consultants. Black Belts (BBs) – Full-time Black Belts lead improvement projects and typically receive 4 weeks of training

Six Sigma Belts: A black belt should present team leadership, understand team dydnamics and assign team member roles and responsibilities. This belt has a thourough understanfding of all aspects of the DMAIC model in accordance with six sigma belts. Green belt analyzes and solves quality problems and is involved in quality improvement research scopes. Green Belts are part-time improvement specialists that receive less training since they provide supporting roles on the improvement projects. Yellow belt is typically employee and has a solid basic knowledge of the six sigma methodology. The work of these types reflects excellent understanding of a quality etc. profit relationship (Knowles et al., 2005; Url 11; Url 12; Hoerl, 2001).

3. ADVANTAGES OF IMPLEMENTİNG SIX SIGMA METHODOLOGY

3.1. Manufacturing sector

Six Sigma has been achieve used in different production companies such as General Electric, Boeing, DuPont, Toshiba, Seagate, Allied Signal, Kodak, Honeywell. For example, Motorola saved US$2.2 billion in the first 5 years of adoption. Honeywell saved in excess of US$2 billion since implementing Six-Sigma in 1994. General Electronics initiated Six-Sigma in 1996 and generated more than US$2 billion of revenue in 1999. Black & Decker's savings rose to approximately US$75 million in 2000, which was more The original concept of six sigma than double the savings of the previous year, and accumulated over US$110 million since 1997. For example, Motorola recovered US$2.2 billion in the first 5 years of adoption. Honeywell saved in excess of US$2 billion since applying Six-Sigma in 1994. General Electronics launched Six-Sigma in 1996 and emergent more than US$2 billion of revenue in 1999. Black & Decker's savings rose to nearly US$75 million in 2000, which was more The original concept of six-sigma than double the savings of the previous year, and generated over US$110 million since 1997 (Mayor, 2003; Wang et al.,2004).

3.2. Financial Sector

Typical six sigma projects in financial institutions contain developing accuracy of allocation of cash to decrease bank charges, automatic payments, developing accuracy of reporting, decreasing documentary credits defects, decreasing check collection defects, and decreasing difference in collector performance Bank of America reported as developping in customer satisfaction and reducing in customer problems after applying six sigma

3.3. Healthcare Sector

Six sigma fundamentals and the healthcare sector are very convenient because of the healthcare system of zero tolerance for failures and potential for decreasing medical defects and issues. Some of the successfully implementations of six sigma studies involve developping timely and accurate requirements, following the process of healthcare distribution, and decreasing the inventory of operative equipment and interconnected costs.

3.4. Engineering and Construction Sector

In 2002, Bechtel Corporation, one of the largest engineering and construction companies in the world, provides savings a huge amount with an investment in its six sigma program to determine and obstruct failures and defects in engineering and construction industries from design to manufacturing, production, construction to on-time delivery of payroll of staff member (Knowles et al., 2005; Mayor, 2003; Wang et al.,2004).

In Six Sigma, the important point is on process development in order to increase capability and decrease variation of process. Six Sigma methodology is about solving problems. The main isue is on business problems that causes harmful effects and extra costs to the companies (Oudrhiri and Pellizzetti, 2006). It is not a classification of best applications or methods. Every corporation is different and thus failures and extra expense etc. are the problems they face. Six Sigma refuses pre-defined solutions and determines and analyses problems to the degree of their root causes (Russ et al., 2008). Six Sigma is generally applied in organizations by determining Six Sigma projects. Six Sigma tools are used in implementing Six Sigma research. Some of the Six Sigma tools can be listed as; descriptive statistics, run chart, control chart, probability plot, check sheets, pareto chart, brainstorming, affinity diagram, interrelationship diagram, matrix diagram and prioritization matrices, quality function deployment, single factor analysis of variance, design of experiments and etc. (Rick and Michael, 2002; Siviy et al., 2007). The CMMI's (Capability Maturity Model Integration) process infrastructure offers a foundation for Six Sigma efforts and helps engineering processes of a business relate to its business processes. Moreover, Six Sigma is not a competitor to CMMI or other process fundamental structures and standards.

Integrating CMMI and Six Sigma can be in several ways as given below (Russ et al., 2008).

- Implement CMMI process areas as Six Sigma research ans studies.

- Use Six Sigma as the tactical engine for high capability and high maturity.

- Apply Six Sigma to improve or optimize development strategy of an organization and

steps of processes.

- Integrate CMMI, Six Sigma, and all other development initiatives to support a standard for

the execution of every study throughout its life cycle.

4. SIX SIGMA IMPROVEMENT TECHNOLOGY IN THE SUPPLY CHAIN NETWORK

A risk management strategy need to objective to occur and continue a supply chain that is both resistant and elastic. An elastic supply chain can reply to unimagined sudden input differences. Accomplishment resistant and elastic supply chain will turn on the degree of visibility and control over operations as well as the timely availability of spare operation capacity. A main first stage toward successing this is to take critical path operations under control. Though controlling processes' outputs by control as it was in traditional quality management is not a satisfactory way because of non-conforming factors would 'slip through the net' and, in any case, control is 'after the case'. So processes must be under control before it applied outputs.

Globalisation, change of marketing, outsourcing efficiency supply chains to be more complicated factors. Thus they are developing sensitive to risk of distribution. Because supply chains contain the linked operations of the different entities in a network, it is applicable to follow the methods in which operation risk can be directed. Despite they divide a number of properties such as unchanging processes, having monitoring control over output difference and low inventory levels, the key variation between the two is their skill to respond to differences in input. The integration of six-sigma and SCM follows from the fact that both have been lustily selected as "operation approach". With the six sigma essent basically being a operaation oriented approach, a six sigma metric actualy measures how well the underlying operation is performing. This "operation approach" is also a property of supply chain management. Six sigma is a data driven, continuous improvement technics that look for take operations under control and improve operation capability. In one Six sigma project, Ford analyst measured inventory levels at one of their plants during manufacturing time and followed that levels changed by 20% over monte. With further research they found one of the significant causes of invnetory fluctuation was inefficient and inconsistent unloading of parts at the plant docks Following stages to develop dock utilisation and in this way driving out operation led to annual savings of more than $3.7 million due to inventory reduction, decreased overtime for unplanned materials handling and other savings (Christopher and Rutherford, 2004).

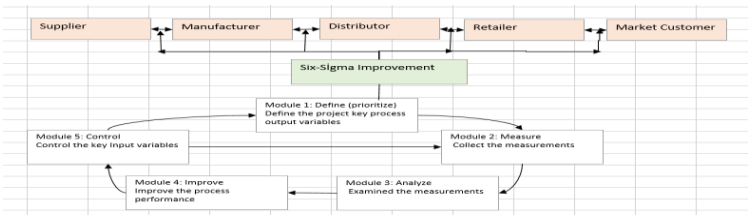
4.1. Applying Six-Sigma to Supplier Development

Supplier development involves a long-term cooperative effort between a buying firm (e.g. a manufacturer) and its suppliers (i.e. its material providers), and is aimed at creating and maintaining a network of competent suppliers. The development activities include supplier selection and monitoring, supplier assistance and training, the provision of incentives for continuous improvements, and supplier organizational integration. The ultimate objectives of supplier development are supplier base reduction, concurrent engineering, reductions in cycle time, reductions in inventories, and increases in customer satisfaction. There should have five steps in the supplier certification process: preliminary evaluation, product design and quality certification, a review of the supply process, performance monitoring, and certification. Raytheon Corporation has applied Six-Sigma to improving its supplier development activities (Breyfogle, 1999).

Six sigma offers more robust supply chain processes. Accross unpredictable events supply chain processes need to be agile and responsive. To accomplish this problem there must be spare process capacity where it is needed most along the critical path (Url 13; Wang et al., 2004).

The key is to approach supply chain activities as processes. Supply chain processes can be improved untill the six sigma level. The framework consists of five integrated modules. These are Define, Measure, Analyse, Improve, and Control steps. The basic framework of the Six-Sigma improvement technology in the supply chain network is given in Figure 4.

Figure 4: The framework of the Six-Sigma model in the supply chain (F. K. Wang et al., 2002)



5. WORK STREAM MAPPING IN PIPE PRODUCTION

The raw materials of pipes reach X Company as steel bands (strip) from the suppliers. All strip passes through an in-line strip wash process which removes oil, dirt and smut from the strip. The strip wash process is essential in meeting with customer and industry requirements. The cleaning process includes a wash tank, rinse tank and strip dryer. These bands are turned into pipes by using rolls. This process forms the strip into a tabular shape. The rolls used in pipe forming process are forming, blade's welding and sizing rolls respectively. First of all, steel bands enter into farming rolls and they are bended. Then they are transferred to blade rolls to get a more bended shape almost approximate to the final shape. After that, they pass through the welding rolls which weld the tube by means of high frequency induction welding and two ends of the tube combine. At last, they go into sizing rolls and attain their last shape. Pipes pass through scarfing process next. Scarfing is the process in which the outside weld bead is removes form the pipe. The necessity of this process is to insure a smooth outer surface of the pipe. At the end of this process, the production is completed. After each production, samples are taken from each pipe batch and they are tested (outside diameter, wall thickness, rockwell, eddy current). Finally, pipelines are packaged and sent to the customer.

Goal Statement:

Define the objective of the research, Y= f(x) and is specific, measurable, attainable relevant and time-bound. Benefits may include: improved response time, lower lead time, reduced warranty claims, improved reliability, lower operating costs. The goal statement should adress these questions:

What is the improvement team seeking to accomplish? How will the improvement team's success be measured? What specific parameters will be mesured, what are the tangible results deliverables (ex. reduce cost, cycle time, etc.)? What are the intangible deliverables/results?

5.1. Layout Optimization for Pipe Production in Factory

The goal is to optimize the factory layout which enables appliers to eliminate non value added activities and thus decrease the unit production cost (Y) to 65$/ton. This involves costs of coil slitting (x1), welding lines (x2), LS finishing lines (X3), packing lines (x4). Amortisation expenses of the investments suggested by this research cannot exceed 0.72$/ton. Therefore, estimated Level 1 benefit out of this research is:

75.000 (tons/year) * (37.66-30-0.72) ($/ton) * 0.7 = 364.350 $/year."

The research scope defines the boundaries of the business opportunity. The goal statement should adress these questions:

-What are the boundaries, the starting and ending steps of a process, of the initiative?

-This is necessary to maintain "boundaries" for the team to remain focused.

-What parts of the business are included?

-What parts of the business are not included?

-What, if anything, is outside the boundaries of the team?

An example can be given for research scope to realize it as given below (XB Holding,2002).

Layout Optimization for Pipe Production in Z Factory (BMB)

Start Boundary: Taking the steel coil into pipe production

End Boundary: Transporting the packed pipes to customer or stock

In Scope: Slitting line

All back pipe and rectangular hollow section units within the welding hall. All back pipe and rectangular hollow steps of finishing lines.

Out of Scope: Layout change for galvanishing line. Layout change for annealing furnace"

This plan identifies the milestones necessary to achieve the objectives and the schedule for completion. A major component of the research plan is an action plan that displays activities, responsibilities and dates by milestone and task.

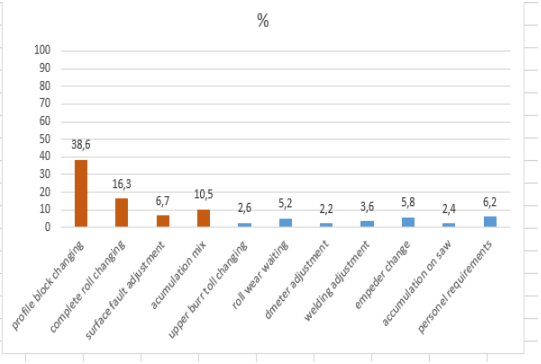
5.2. General Problems in Production Procedure

As a result of the discussions with the black belt three main problems were determined. These are ıdle times, reduction in speed and customer complaints about surface defects. According to the customer complaints, much more meter pipe from

amount of production had surface defects. 215.256-meter length pipes were scraped, taken back. This was caused financial loss. Approximately, 178.396 meter pipes were reproduced. A huge amount compensation was paid. The frequency percentages of Idle time reasons in a year are presented as below in Table 2.

Table 2: Frequency Percentages of Idle Time Reasons in a Year

Reasons of Idle Time	%
profile block changing	38,6
complete roll changing	16,3
surface fault adjustment	6,7
acumulation mix	10,5
upper burr toll changing	2,6
roll wear waiting	5,2
dmeter adjustment	2,2
welding adjustment	3,6
empeder change	5,8
accumulation on saw	2,4
personel requirements	6,2



About data of ıdle times, at the same tiem data about idle times related to the rolls were sorted from all idle time reasons. Idle times related to the rolls are profile block changing, round pipe diameter adjustments, surface defect adjustments, roll amortization etc. By collecting thse types of data, it was aimed to find the most significant reasons of times related to the rolls. The steps of production system were presented in table aa. This table was classified to determine the inputs, processes and outputs of the structure of system. This technique is usable tool to research the capabilities of the system at the starting stage of the study. Setup instructions, operator, roll sets and roll usage plan are components of Input stage. Machine matching, setup, re-maching, roll changing are components of Process stage. Customer complaints, scraps, speed and idle times are components of output stage. This IPO (input-process-output) Table 3 is shown as below.

Table 3: Phases of System Table

Input	Process	Output
Setup instructions	Machine matching	Customer complaints
Operator	Setup	Scraps
Roll sets	Re-maching	Speed
Roll usage plan	Roll changing	Idle times

5.3.Components of System - SIPOC Table

All relevant components of the research should be identified before work begins so the whole system can be understood clearly. Hence, a SIPOC table was organized to answer these questions:

Who are the consumers of the process, what are the requirements of the consumers, and what specifications are placed on the Inputs? When the study does not deal with the real customers or suppliers, machines and warehouse could be thought as these components (see Table 4).

Table 4: Supplier-Customer Framework

	S	I	P	O	C
	Supplier	Input	Process	Output	Customer
Roll remaching process	roll warehouse	Roll used in production	roll remaching	roll sets which have reduced diameters by remaching in CNC	pipe machines
	roll warehouse	Roll used in production	roll remaching	roll sets which have polished surfacess by remaching in turn benches	pipe machines
Roll Usage Process	roll warehouse	rolls ready to use	roll change	machines ready for production	tube machines
	tube machines	setup instructions	roll setups	rolls used in production	roll warehouse

6. PARETO ANALYSIS

A pareto chart can be constructed by segmenting the range of the data into groups. The left-side vertical axis of the pareto chart is labeled frequency- the number of counts for each category, the right-side vertical axis of the pareto chart is labeled with the group names of your response variables (Akın, 1996; Akın and Ozturk, 2005; Url 8). The number of data points is determined by residing within each group and construct the pareto chart, but unlike the bar chart, the pareto chart is ordered in descending frequency magnitude and users define the groups.

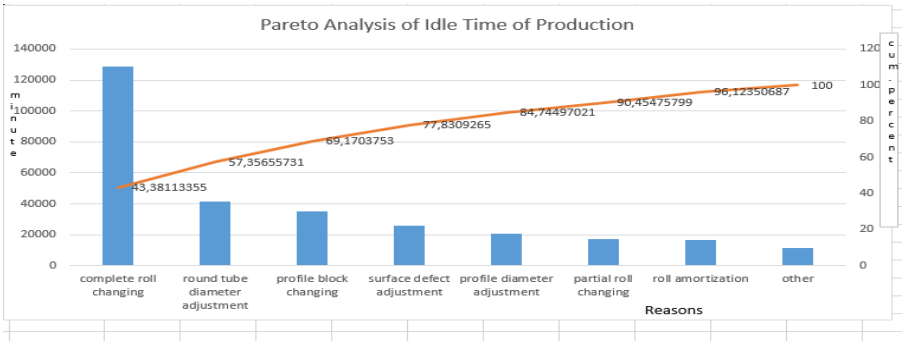
Vilfredo Pareto was a 19th century Italian economist and his researches were about economic problems using mathematics and the technique was named with his name and this method supports to determine and group the detects-error due to percentage importance. He specified that 80 percent of the location in Italy was belonged by 20 percent of the citizens.

This method is a method of analyzing causes of problems to support an effective solution.Generally, this way implements according to 80/20 law. Pareto diagrams are the graphical tool applied in Pareto analysis Pareto analysis is a method which is applied to distinguish reasons from less important ones (Gitlow et al., 2005; Cravaner et al., 1993; Leavengood and Reeb, 2002).

6.1. Analysis for Idle Time Production

Data was collected in measure phase about the idle time reasons of the rolls. In Anlyze phase, pareto diagram is used to see the reasons that affect the idle time most. This diagram can be presented in Figure 5.

Figure 5: Pareto Analysis



Considering pareto Diagram (see figure 5) ; average changing times have been computes and features were arrayed from top to bottom. complete roll changing, round tube diameter adjustment, and profile block changing took the biggest percentage. To see all of the changes, surface detect adjustment has also been examined through the analysis. These four reasons have high risks features. Firstly, These foru reasons must be eliminated from the systems according to the Pareto analysis.

7. GENERAL OVERVİEW OF SIX SIGMA APPLICATION-CASE STUDY IN MANUFACTURING FIRM

The main is to resign Six Sigma quality product or service without suffering from a poor understanding of customer necessaries and resign performance of the suppliers. To release Six Sigma in a supply chain network, management must clearly define potential effects of the enterprise on corporate performance and supply chain groups (Montgomery, 2008). To better know customer necessaries, companies use to keep in mind that errors are decreased if they focus on bring customer necessaries to the Six Sigma level Bob,Galvin, former CEO of M, asked IBM, which is a big firm, to apply Six Sigma Suppliers must also find value in collaborating with the organizatör. One method to detect value to the suppliers is to provide a strategy to occur relationship with the suppliers who commit to apply Six Sigma attempt and aline with the corporate objectives, and decrease the number of suppliers who aren't enthuastic about having great advancement in a short period of time .To release the enterprise, managers must present a commitment document, mutual performance aims, commitment to eductation and progress review schedule (Antony and Banuelas, 2002; Wang, et al.,2004). Since the early 90s, companies have been researching on decreasing number of suppliers. As the companies developed their design and processing processes, dependency on development by supplier developed to success continually demanding business aims Similarly, when applying Six Sigma, M Company asked its suppliers to aply its Six Sigma travel. To obtain the suppliers enthusiastic, quite a few activities were handed to training suppliers for the dramatic development and the objectives it suggested them. After winning the Malcolm Baldridge National Quality Award, M Company felt privileged to distributed its achieve story with its suppliers and customers. Regular supplier informations were held to communicate the M process, its suppliers' expectation and stable sources. To some point, suppliers were stabled to participate in M's Six Sigma process to provide continually developing products and services. To development suppliers' understanding of the Six Sigma process, they were used to share four main factors, which were as follows (Montgomery, 2008; Antony and Banuelas, 2002):

Understanding Six Sigma—An overview of the Six Sigma methodology and terminology.

Process Quality Improvement—An overview of process improvement methods explaining applicability of various tools.

Design for Manufacturability—The process of integrating product and process development for producing defect-free products.

Cycle Time Reduction—Straightening out material flows to practice lean principles.

M University was then established to institutionalize corporate strategies and manage the variable process. It suggested services to uplift M Company's suppliers' capabilities and strengthen the relationship with them. The university also had a consulting group to provide its clients in applying different methods. Moreover training and guiding suppliers, following policy of M Company was varied and procedures were altered to image the expectations of its suppliers of M Company. The new procedures imaged intent and verification to minimize product research and verification facilities, and promote process control and development facilities. Supplier' scorecard would be followed periodically difference success and defining areas for development. Furthermore buying policies, a process for suppliers involvement in theproduct improvement step was applied This allowed M to leverage its suppliers' speciality and sources and conspire with its suppliers in increasing new products This early inclusion of suppliers supplied M engineers to produce designs that its suppliers could manufacture at low fault rates (Antony and Banuelas, 2002; Wang, et al.,2004). Application Six Sigma is supply chain needs a strategic thinking, proper planning and significant commitment to optimize the performance of the supplier group. Besides, including customers allows an organization to specify products trhat are acceptable to customers and spplied by suppliers for preservation the product surplus. There are different companies, such as C, that have included their suppliers in the Six sigma interference to provide regular commitment to the institutional performance development. Six sigma in the supply chain is serious to success higher sigma levels. If a company is application Six sigma, it can't be researching with three-or four sigma suppliers or customers with unrealistic expectations, and hope to success its Six Sigma quality objectives (Nonthaleerak and Hendry, 2006).

8. CONCLUSION

By the help of this research, it is aimed to develop knowledge about six sigma quality system and its tools. Six sigma method is investigated to develop the processes of supply chainSix sigma method requires to be applied forever in a supply chain because of it is a sustainable research area. It is impossible to provide zero failure, defect in real life but it is a continuous process. Firstly, A deep literature survey has been accomplished to understand the concept better. For the case study, the present system was examined and the research scope was specified by considering the operations that needs development. The applications of six sigma methodology to the supply chains was presented. Six sigma methodology was applied to define the importance of quality and satisfaction of customers in the supply chain management and production area. The probable problems of the system were observed and main causes of problems were discovered to eliminate them or reduce the effects of these problem in the systems. The phase of analyses provided to understand which items of problems causes the failures and defects. The research can be extended with other problems and factors of system according to the discovering from aspects of quality. Today, supply chains management includes more and more complex structures because of their inconsistent objectives. Conducting comlex structures are very hard so this frame needs to be developed.

REFERENCES

Akın, B. , 1996. "ISO 9000 Uygulamasında İşletmelerde İstatistik Proses Kontrol Teknikleri", Istanbul: Bilim Teknik Yayınevi, pp. 49-50.

Akın, B., Ozturk, E., 2005. Istatistiksel Proses Kontrol Tekniklerinin Bilgisayar Ortamında Uygulanması, VII.Ulusal Ekonometri ve İstatistik Sempozyumu Bildiriler CD, Istanbul, pp.1-15.

Antony, J., Banuelas, R. 2002, "Key ingredients for the effective implementation of Six Sigma program", Measuring Business Excellence, vol.6, no. 4, ABI/INFORM Global, p.20

Benita M. Beamon, B.M., Tonja M. Ware, T.M. 1998, " A process quality model for the analysis, improvement and control of supply chain systems", International Journal of Physical Distribution & Logistics Management, Vol. 28 No. 9/10, pp. 704-715.

Breyfogle, F.W. 1999, Implementing Six Sigma: Smarter Solutions Using Statistical Methods, John Wiley: New York.

Christopher, M. ,Rutherford, C. (2004) "Creating supply chain resilience through agile six sigma",CriticalEye, June- August 2004, pp.24-28

http://www.valuenetworkmanagementforum.org/vnmf/images/past_events/competitive/publications/Creating_Supply_Chain_Resilience_Through_Agile_Six_Sigma.pdf)

Cravener, T.L., Roush, W.B., Jordan, H., 1993."Pareto Assessment of Quality Control in Poultry Processing Plants", Applied Poultry Science, 2, pp.297-302.

Erol, E., Ferrell, Jr. GW. 2004, "A methodology to support decision making across the supply chain of an industrial distributor a Department of Business Administration", Int. J. Production Economics 89 (2004), pp. 119–129

Gitlow, H.S., Openheim, A.J., Oppenheim, R.,Levine, D.M. (2005)."Quality Management ", Third Edition, McGrawHill, USA, pp. 366-367.

Gunasekaran, A. , Ngai, E.W.T. 2004, "Information systems in supply chain integration and management", European Journal of Operational Research 159 (2004), pp. 269–295.

Gupta, P. 2004 "Six Sigma Business Scorecard, Creating a Comprehensive Corporate Performance Measurement System", The McGraw-Hill Companies.

Halseth, K. 2008 , "Process Modelling & Mapping: The Basics, Regional Coordinator", Forms Management & Production, David Thompson Health Region, Session 64, http://c.ymcdn.com/sites/www.bfma.org/resource/resmgr/articles/08_64.pdf.

Hoerl, R.W. 2001," Six Sigma Black Belts: What do they need to know?", J. Quality Technology, vol. 33 no.4, pp 391–401.

Ingle, S., Jacobs, D. 1999, "Approaches to TQÖ adoption in an Irish context", Ho,S.K.M., TQM and Innovation, Hong Kong Baptist University, Hong Kong.

Knowles, G., whicker, L., Femat, J.H. and Canales, F.D.C. 2005, "A conceptual model for the application of Six Sigma methodologies to supply chain improvement", International Journal of Logistics: Research and Applications Vol. 8, No. 1, March, pp. 51–65.

Leavengood,S. and Reeb,J.,2002. "Part 3:Pareto Analysis and Check Sheets, Performance excellence in the wood products industry" Statistical Process Control, EM8771, January 2002.

Linderman, K., Schroeder, R.G., Zaheer, S., Choo, A.S. 2003, " Six Sigma: a goal-theoretic perspective", Journal of Operations Management 21, pp.193–203.

Mayor, T.2003, "Six Sigma comes to IT: targeting perfection", CIO Magazine, www.cio.com/archive (accessed 24 January 2004).

Montgomery, D.C., William H. Woodall, W.H. 2008, An Overview of Six Sigma, International Statistical Review vol.76, no.3,pp.329–346

Nonthaleerak, P., Hendry, L.C. 2006, "Six Sigma: Literature review and key future research areas". Int. J. Six Sigma Competitive Advantage, 2(2), pp.105–161.

Oudrhiri, R., Pellizzetti, F. 2006, "Using maturity models for Six Sigma and strategy execution", IQPC 7, Annual conference, Amsterdam.

Radouane, O., Fabrizio, P. 2010,"Design for Six Sigma (DFSS): a cultural and paradigm shift from Deterministic to Probabilistic Design", SEPG Europe, Portugal.

Raisinghani, M.S. 2005,"Six Sigma: concepts, tools, and applications" , Industrial Management & Data Systems Vol. 105 No. 4, pp. 491-505, Emerald Group Publishing Limited.

Rick, H., Michael, S. 2002,"Optimize Your Solution: Integrating Six Sigma and CMM/CMMI-Based Process Improvement", 2.Software Technology Conference, 2002.

Russ, R., DanaSperling, D., Rometsch, R., PeterLouis, P. 2008,"Applying Six Sigma in the Field of Software Engineering", Lecture Notes in Computer Science, vol. 5338, Springer.

Sabders, D., Hild, C. 2000 , "A discusiion of strategies for Six Sigma implementation", Quality Engineering (12)3, pp.303-9.

Siviy, J.M., Penn,M.L, Stoddard, R.W.2007,"CMMI & Six Sigma: Partners in Process Improvement", Addison-Wesley.

Sheikh, K. 2003, "Manufacturing resource planning (MRP II) with introduction to ERP, SCM and CRM", McGraw-Hill, New York, pp.538-539

Tennant, G. T. 2001, "Six sigma: spc and tqm in manufacturing and services", The Evolution of Six Sigma, . Gower Publishing, Ltd. p. 6. ISBN 0-566-08374-4. Retrieved 2012-03-19.

Wang, F.-K., Du,T.C., Eldon, Y. L. 2004, "Applying Six-Sigma to Supplier Development", Total Quality Management, Vol. 15, No. 9–10, pp. 1217–1229. http://iceb.nccu.edu.tw/proceedings/APDSI/2002/papers/paper90.pdf

Wing, C.L., Chowb, S., Maduc, C.N., Kueic, C.-H., Yua, P.P. 2005, "A structural equation model of supply chain quality management and organizational performance", Int. J. Production Economics, vol. 96, issue 3, pp. 355-365.

XB Holding, Green Belts Education Documents, 2002

Url 1- Supply Chain Management, Rai Technology University Engineering Minds-Chapter 1, pp.3-11 http://164.100.133.129:81/eCONTENT/Uploads/Supply_Chain_Management.pdf, access time: 12.04.2017

Url 2- What is Supply Chain Management?, Penn State University, Smeal College of Business, https://www.smeal.psu.edu/cscr/pitch/what-is-supply-chain-management , access time: 12.04.2017

Url 3- Six Sigma, http://leansixsigmadefinition.com/glossary/six-sigma/ , http://www.cqeweb.com/chapters/chapter4.pdf, p.12

Url 4- The History of Six Sigma, http://www.tbmcg.co.uk/about-tbm/six-sigma.html

Url 5- Sinikov, C. 2002, " The Six Sigma phenomenaold or new perception of quality?", www.alunet.com

Url 6- Six Sigma Vs. Quality Methodology, Jean Scheid, http://www.brighthubpm.com/six-sigma/34733-six-sigma-vs-quality-methodology/

Url 7- Six Sigma, http://www.brighthubpm.com/six-sigma/p3.html

Url 8-Six Sigma Methodologies, http://www.6sigmastudy.com/Six-Sigma-Methodologies.asp

Url 9- How to Calculate Your DMAIC Process Sigma, http://www.brighthubpm.com/six-sigma/42956-how-to-calculate-your-dmaic-process-sigma/

Url10-Mary K. Pratt-posted by Margaret Rouse, Business process mapping, October 2016. http://searchcio.techtarget.com/definition/process-mapping.

Url 11- How to Become Certified as a Six Sigma Green Belt or Black Belt, http://www.brighthubpm.com/six-sigma/27993-how-to-become-certified-as-a-six-sigma-green-belt-or-black-belt/ Heidi Wiesenfelder, edited by Michele McDonough.

Url 12-Six Sigma Green Belt , https://www.isixsigma.com/events/six-sigma-green-belt-390/

Url 13- Six Sigma in Supply Chain Management, written by Suvo; edited by Marlene Gundlach, http://www.brighthubpm.com/six-sigma/43061-six-sigma-in-supply-chain-management

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

STATISTICAL ANALYSES OF SAME CONTENT TEXTS WRITTEN IN DIFFERENT LANGUAGES

DOI: 10.17261/Pressacademia.2017.580
PAP-WCTIE-V.5-2017(18)-p.124-129

Mentor Hamiti[1], Elissa Mollakuqe[2], Asmir Rahmani[2], Florinda Muhaxhiri[3]
[1]South East European University, m.hamiti@seeu.edu.mk
[2]University of Prizren "Ukshin Hoti", elissamollakuqe@gmail.com
[3]State University of Tetova, florinda.muhaxhiri@gmail.com

ABSTRACT

Language is the basic and most consummate way of communication between people. It can be materialized in two ways: spoken and written. Every society has a spoken language, even the primitive ones. Although only civilized societies have a written language with a defined alphabet. The presence of letters in the context of words determines the meaning, while the determined order of these in words presents a work of art. It is, thus, understandable to raise the question: Which letter is used the most and least in different languages? Or maybe there is similarity on their distribution even though it has to do with languages which use different alphabets? Or in general, which are the differences or what could different languages have in common when they interpret the same content?! The answer to this question remains within the scope of this paper

Keywords: Text, English, Albanian, Turkish, Bosnian

1. INTRODUCTION

Computer language software and their presence on the Internet have become a vital part of communication and modern concepts of the so-called "scientific field". Linguists have taken seriously the provocation of the computer era in the field of linguistics, because computer linguistics is the only way for protecting, enriching and advancing every language in the world. The aim of this research is to present the continuous development of the languages, including the statistical research component. With the help of the original program, written in C# programming language, we set the computer in service of different languages, since text with same content were written in English, Albanian, Bosnian and Turkish languages. Linguists can use the gained results for further linguistics research and analyses (Hamiti, 2015).

2. CLASSIFICATION OF ALPHABET LETTERS AND THE SPECIFICS OF COMPUTER BASED PROCESSING

The organization of the letters for the alphabets in question is structured in separate forms for each language. The four languages belong to different families. English belongs to the family of the Germanic languages and uses 26 Latin letters (Murphy, 2008). Albanian is part of Indo-European family of languages with a total of 36 letters, of which 34 are standard (Latin) (Hamiti, 2005). Among these 9 are considered as consisting of 2 letters (digraphs) and 2 are diacritical marks. Bosnian belongs to the family of Slavic languages consisting of 30 letters, of which 24 are standard letters, while 6 are nonstandard, containing also composite ones (Schumann, 2010). Turkish is a language on its own, consisting of 29 letters of which 22 are standard while 7 are nonstandard and does not contain digraphs (Attaoullah, 2012). Because of the stated differences between the languages that are being treated, the standard existing applications, which in most cases are in English, cannot be used for universal processing of texts written in these languages. Therefore, it is necessary to design an original application that takes into account the specifics of the respective languages.

2.1. English Language, the Specifics of Computer Based Processing

English is described as computer language and notably advances in terms of computer processing compared to many other languages. All letters of the English alphabet are found in all standard computer keyboards, therefore the writing of texts and the text analysis written in this language are much more convenient and easier to realize. The English alphabet letters belong to ASCII standard, which remains as the international global standard for the interpretation of commands and texts

based on communication protocols. In addition, this standard has the support of all programming languages like C, C++, C#, etc. Also it allows the possibility of using any of them, without difficulties in processing the written texts in this language. That is not the case with other languages that require UNICODE support for interpretation and processing (wordandphrase, 2015).

Figure 1: Letters of the English Alphabet

A a	B b	C c	D d	E e	F f	G g	H h	I i	J J	K k	L l	M m
N n	O o	P p	Q q	R r	S s	T t	U u	V v	W W	X x	Y y	Z z

2.2. Albanian Language, the Specifics of Computer Based Processing

Albanian consists of 36 letters, of which 34 can be written easily from the computer keyboard, while two are letters with diacritical signs: "Ë, ë" and "Ç, ç", which means that the process of their writing is more complex. It can be done in one of the following ways: by using keyboard configuration where some additional Albanian language characters are hidden within special keys like parentheses, by installing fonts for text processors, or by using codes for generation of these two symbols. In addition, 9 out of the other 34 letters are treated as compound letters, also known as digraphs, which are formed by linking two Latin characters in one single letter for representing a single phoneme in Albanian language. This in programming is accompanied by another feature, that does not allow each letter to be treated as special character, but takes the meaning of the string, thus additionally complicates the programming and does not allow the use of existing applications for text processing (Academia, 1976)(Hamiti, 2015). Therefore, within the designed application in C#, special algorithms for solving the mentioned problem are used. As an illustration, it is worth mentioning that the MS Word count for Albanian language does not generate the correct result!

Figure 2. Letters of the Albanian Alphabet

A A	B B	C c	Ç ç	D d	DH dh	E e	Ë ë	F f	G g	GJ gj	H h	I i	J j	K K	L l	LL ll	M m
N N	NJ Nj	O o	P p	Q q	R r	RR rr	S s	SH sh	T t	TH th	U u	V v	X x	XH Xh	Y y	Z z	ZH zh

2.3. Bosnian Language, the Specifics of Computer Based Processing

Bosnian has another category of the alphabets. Bosnian uses both alphabets simultaneously: The Latin and the Cyrillic alphabet (Ronelle, 2010). The alphabet contains a total of 30 letters, of which 24 are standard letters, while 5 are considered as nonstandard based on ASCII code, while 2 are composed letters, and one of them is a letter with diacritical mark. Therefore, writing of texts in this language through standard keyboards is very difficult and almost impossible without the installation of special fonts and respective keyboard configuration. But in terms of programming, however, processing of texts is conditioned through the programming languages that enjoy UNICODE support.

Figure 3: Letters of the Bosnian Alphabet

A A	B b	C c	Č č	Ć ć	D d	Dž dž	Đ đ	E e	F f	G g	H h	I i	J j	K K
L L	LJ lJ	M m	N n	Nj nj	O o	P p	R r	S s	Š š	T t	U u	V v	Z z	Ž Ž

2.4. Turkish Language, the Specifics of Computer Based Processing

Turkish has an alphabet suitable for the sound of its own language. The Turkish alphabet consists of 29 letters, which can be categorized into two simple categories, with 22 standard and 7 nonstandard noncomplex letters (Asuman, 2001)(Underhill, 2010). So, in terms of programming, it allows the possibility of using characters, which was not the case with the previous languages, the Albanian and Bosnian.

Figure 4: Letters of the Turkish Alphabet

A A	B b	C c	Ç ç	D d	E e	F f	G g	Ğ ğ	H H	I ı	İ i	J j	K k	
L L	M m	N n	O o	Ö ö	P p	R r	S s	Ş ş	T T	U u	Ü ü	V v	Y y	Z z

3. THE APPLICATION DESIGNED FOR TEXT PROCESSING

Considering the specifics of the analyzed languages, "Microsoft Visual Studio 2013" platform and C# programming language were used within this paper, for designing a specific program, which fulfills the needs for textual processing of English, Albanian, Bosnian and Turkish languages. The source code for this program, used for generating the results, is being fully presented through this paper. The program is called "Analyses of letters".

This program totally answers the needs for analyzing written texts in all four languages, and can be very easily adapted for other languages. It enables reading textual files, sets total number of characters within the textual file (file with .txt extension) (Hamiti, 2015), counts every single letter, punctuation marks and special characters separately, like for example empty line, the use of TAB and jumping on new line, etc. Also, it enables direct calculation of letter appearance frequency within the individual textual file.

Unable to find more professional texts with similar content translated into these four different languages, in the frame of this paper, only 205 pages of text were analyzed. In order to obtain convincing results, we categorized them into ten groups: Basic expressions; greetings and invitations; communications and notifications; merchandise, money and numbers; transportation; food and beverage; guidance; documents (regulations); health and emergencies; and in the tenth group the other cases, containing some specific data such as messages, conversations, etc. In the paper we generated the results of averages of specific groups in order to avoid eventual errors and the possibility of deviation of results for specific content. Therefore, the possibility for some deviation remains open for the obtained values if considering bigger dataset such is the case with the EU languages, where you can find text with the same content in dozens of different languages, like the case with JRC-Acquis dataset. But the Albanian, Turkish and Bosnian languages do not enjoy this privilege, because they are not yet EU members.

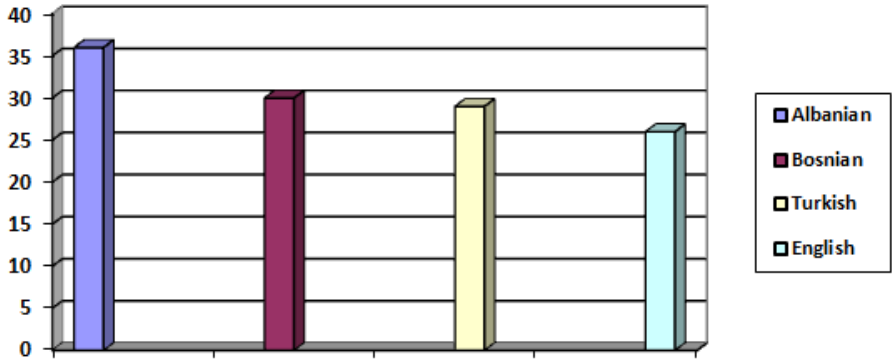
Figure 5: Analysis of letters

4. SAME CONTENT TEXT ANALYSEZ RESULTS

The number of characters, punctuation, standard and nonstandard letters, differ between the four analyzed languages, the English, Albanian, Bosnian and Turkish. The results are quite interesting, on how the same content is presented in four different languages. Figure 6. shows the total number of characters for all four languages, where the first appears to be the Albanian with a total of 703.192 characters, which has 36 letters in the alphabet, followed by the Bosnian language with 665.840 characters, which has 30 letters in the alphabet, where as third appears the Turkish with a total of 625.037

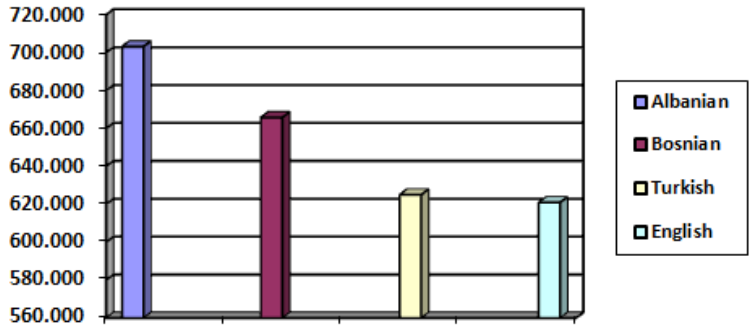
characters, and which has 29 letters in the alphabet, and in the end appears to be the English language with 620.920 characters with an alphabet consisting of 26 letters.

Figure 6: Number of Letters in Alphabets



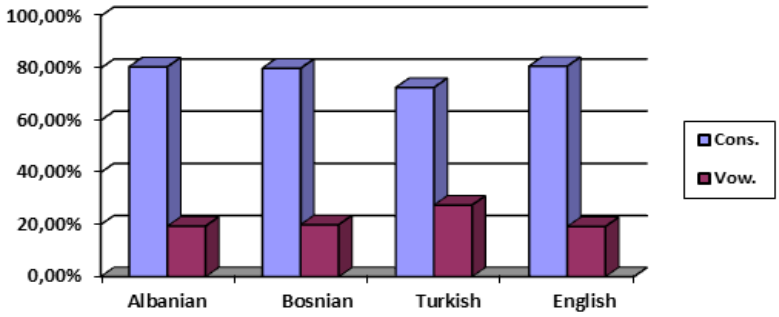
While analyzing these four languages and their distribution of characters, it is important to note that the study also reviews punctuation marks, spaces and other symbols (numbers, special signs, etc.). In this case, the study shows that the distribution of these symbols varies significantly between the four languages. Albanian in the examined text contains 2.44% of punctuation marks, Turkish has 2.62%, Bosnian with the highest percentage of punctuation marks with 2.77%, and English with least, around 2.21%. Furthermore, regarding spaces used the results are as follows: Albanian uses 16.99%, Bosnian luses 16.53%, Turkish uses 16.15%, and English uses approximately 16.02% of space. This helps us to understand that for the same content the Albanian uses more words, with Bosnian as second, followed by Turkish, and English in the end, by what we conclude that this sequence is the same as the one with the number of letters in the alphabets.

Figure 7: Number of Used Characters



During the analysis of these four alphabets, it is noted that in terms of using the vowels and consonants, they are not very compatible. Albanian, English and Bosnian languages have almost the same vowels "a, e, i, o, u", while Turkish has additionally vowels "ı , ö, ü". However, the distribution of these components is similar, wherein: the Albanian in the alphabet has 80.56% consonants and 19.44% vowels, Bosnian has 20% vowels and 80% consonants, Turkish has 27.48% vowels and 72.58% consonants, and the English has 80.77% consonants and 19.23% vowels, which are shown in the figure below.

Figure 8: Dispersion of Vowels and Consonants



Vowels are mostly used in Albanian text with 42.32% and the most used vowels appears to be “e, ë, I”. Second is Turkish with 41.02% vowels in text, where the most used vowels are: “e, a, o, I”. Third is English with 40.52% vowels in text, where most commonly are used "e, i". Fourth is Bosnian with 38.72% vowels in the text, where most commonly used vowels are: “a, i, e”. In this case it is noted that the Albanian, English and Turkish more frequently use vowel “e”, while Bosnian does not. The highest frequency in Albanian and English have letters E and T, with about 48.52%, in Bosnian vowels A and I have the highest frequency, around 25.02%, while E and A have the highest frequency in Turkish with 24.28%. In the four graphs below we show the frequency of letters for the four languages as well as the features provided by the application.

Figure 9: Frequency of Letters in Albanian

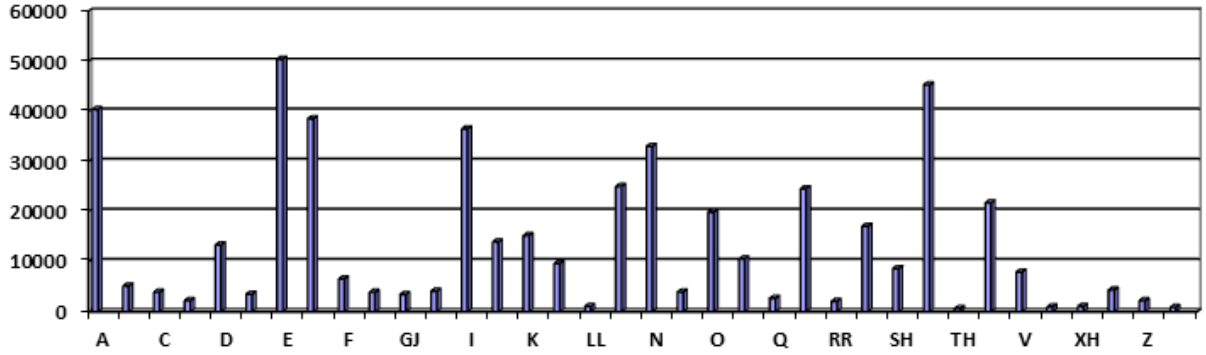


Figure 10: Frequency of Letters in English

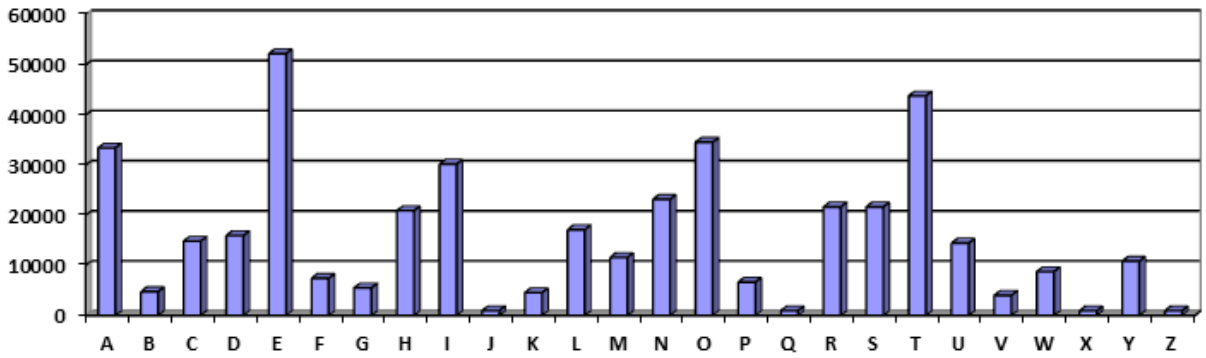


Figure 11: Frequency of Letters in Bosnian

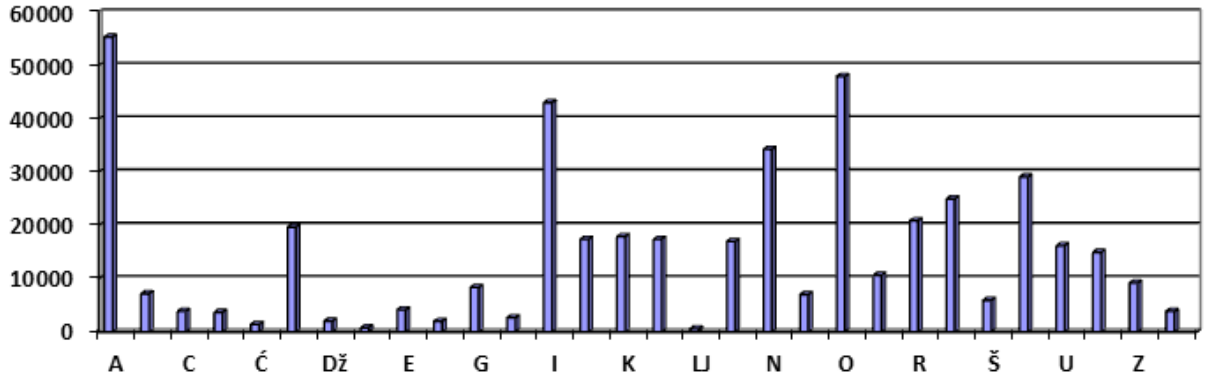
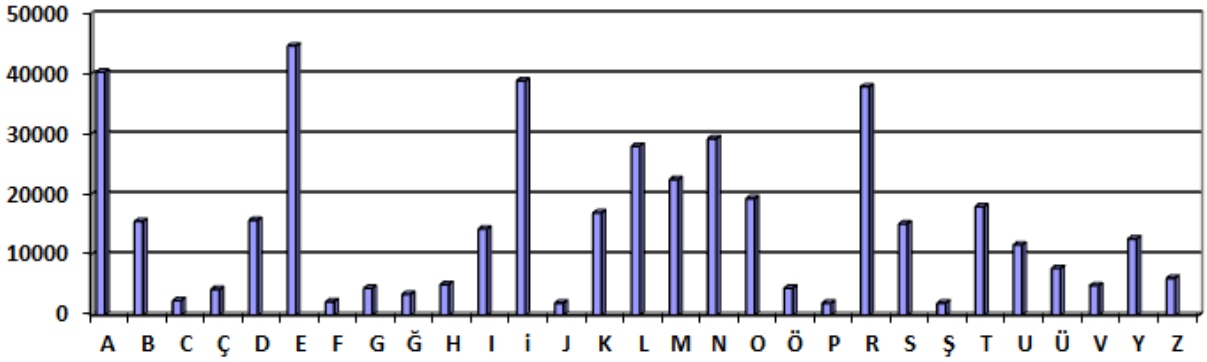


Figure 12: Frequency of Letters in Turkish



5. CONCLUSION

After conducting many statistical analyses on the same text, written in four different languages, Albanian, English, Bosnian and Turkish, we generated the following conclusions:

- Fewer characters were used in the text written in English i.e. the language that has fewer letters in the alphabet, 26 in total, followed by Turkish with 29 letters, Bosnian with 30 letters, and Albanian with 36 letters in the alphabet and that uses most characters.
- Punctuations in Turkish and Bosnian are almost equivalent with about 2.75%, and have direct relationship compared with the total number of characters.
- Spaces used in written Albanian occupy 16.99% of the text, which leaves us to understand the texts in Albanian use more words to interpret the same content, followed by the Bosnian 16.53%, Turkish with 16.15% and finally English with 16.02%.
- The ratio between vowels and consonants in Albanian is equivalent to 80.56% with 19.44%, in Bosnian is 80% to 20%, Turkish has a ratio of 72.58% to 27.42% and English language has a ratio of 80.77% by 19.23 %.
- In Albanian, English and Turkish, vowel “e” is used more frequently, whereas in Bosnian language most frequent vowel is “a”.
- All four languages have a single letter with a very low frequency of usage. In Albanian it is “x”, in Bosnian "đ", in Turkish “o” and in English the letter “z”.
- The application realized in C# meets the requirements for the languages set in the study and it is easy for use and modification from other researchers based on the requirements for specific languages.
- Also the application is offered by the authors as an open source for all concerned in continuing the research in the field of computational linguistics for educational purposes.

REFERENCES

Alexander, R. (2010). Bosnian, Croatian, Serbian, a Textbook: with Exercises and Basic Gramma. Sarajevo.

Attaoullah, F. A. (2010). Beginner's Turkish (Beginner's (Foreign Language)). Izmir.

Dodi, A. (2011). Fonetika dhe fonologjia e gjuhës shqipe. Tiranë: Akademia e shkencave në Shqipëri.

Hamiti, A. (2005). Fonetika dhe fonologjia e gjuhës standarde shqipe. Shkup: Interdiskont.

Hamiti, M. (2015). Text Analyzer.

Letërsisë., I. i. (1976). Fonetika dhe gramatika e gjuhës së sotme letrare shqipe. Tiranë: Akademia e Shkencave e Republikës së Shqipërisë.

Mentor, H. (2015). Analyses of Same Content Texts written in Different Languages.

Murphy, R. (2008.). English Grammar In Use. Third edition. Cambridge University Press.

PHRASE, W. A. (n.d.). Frequency of Letters in English Language . Retrieved 12 23, 2015, from //www.wordandphrase.info/frequencyList.asp

Pollard, A. Ç. (2001). Teach Yourself Turkish1. Istambul.

Schumann, J. (2010). Bosnian For Beginners: A Book In 2 Languages. Sarajevo: Bilingual.

Underhill, R. (2010). Turkish Grammar (Turk dili grameri, dil, Turk dili, Turkce grameri) (English and Turkish Edition).

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

USING MULTI-CRITERIA DECISION MAKING FOR SELECTION PROVINCES FOR FATTENING FARM

DOI: 10.17261/Pressacademia.2017.581
PAP-WCTIE-V.5-2017(19)-p.130-138

Burcu Ozcan[1], Gulsen Akman[2], Cagin Karabicak[3], Yasin Erdogan[4]
Kocaeli University, Engineering Faculty, Industrial Engineering Department , burcu.ozcan@kocaeli.edu.tr
Kocaeli University, Engineering Faculty, Industrial Engineering Department , akmang@kocaeli.edu.tr
Kocaeli University, Engineering Faculty, Industrial Engineering Department , cagink@kocaeli.edu.tr
Kocaeli University, Engineering Faculty, Industrial Engineering Department , yasinerdogan0103@gmail.com

ABSTRACT

Turkey has a high potential in the field of agriculture and livestock production. Turkey has a long history in the production of livestock system, which became the “buffer” sector of economy and the “locomotive” sector of Turkey. Although nowadays pastures are decreased, Turkey has quite large pastures. Besides, the domestic demand for meat in Turkey is quite large and tend to a show considerable growth. Animal husbandry is carrying great importance in terms of protein source and fattening cattle in the red meat production. In this sense, the development of meat production in terms of quality and quantity is also very necessary in terms of the nutritional problem that arises parallel to the growing population. In this study, the causes of increase in meat prices in Turkey were investigated, and the findings showed that, the cause of the price increases made as inadequate animal fattening. As a result, it was decided to establish a new place for the production of cattle-raising. Before starting the construction of the fattening livestock farm, firstly the location of the establishment is considered. In the decision-making process, it is quite difficult to make choices because of the different alternatives in terms of various criteria's. For this purpose, the PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluation) method, which is an easy and straightforward method of multi-criteria decision making techniques which are frequently used in researches in recent years, has been used to select the place of establishment in Turkey. The choice of fattening farm establishment site was evaluated in 81 provinces; cattle feed areas, demand on meat, urban development index value, average minimum temperature, average maximum temperature, precipitation amount and grant support criteria were evaluated with using the PROMETHEE method. The interactions between the criterion and the criteria that influence the process were determined as a result of the survey conducted. As a result of this survey study, criteria weights and criterion values were determined by Analytical Hierarchy Process (AHP) method and solved by PROMETHEE method considering the results. As a conclusion, the most suitable and the less suitable provinces for fattening livestock farm are selected

Keywords: Multiple criteria decision making methods, analytical hierarchy process (AHP) method, PROMETHEE, selection of establishment place for fattening farms in Turkey.

1. INTRODUCTION

In the rapidly developing and changing world, while adequate and balanced nutrition is at the beginning of the problems, it actually contains solutions in itself. Livestock as a branch of agriculture has a strategic importance in terms of revitalizing development in the country's economy. Livestock nowadays has become an industry in developed countries and it become an integral part of the economy. This situation suggests that agriculture and therefore animal husbandry is a sector that needs to be developed at national level.

According to the TUIK (Turkish Statistics Institution) data, the GDP growth rate with fixed prices for the last 11 years increased by 4.9% on average, while the growth rate in the agricultural sector was 2.2%. This difference in rates of change has reduced the share of agriculture in GDP from 12.2% to 9.2%.

In their study, Ocaklı et al. report that in the future, the demand for animal-origin food will not increase significantly but that the demand for meat and dairy products will increase by two times in countries with rapid population growth in the 2020s. In Turkey, the livestock sector has tried to develop according to years but these developments have not been able to

catch up with the growing population and the growing livestock sector in intercountries. The fact that agriculture and animal husbandry are not being carried out in our country in large quantities has dragged us into importing, and causing the animals to be taken from foreign countries.

Turkey thought that red meat prices are very high in 2010. Therefore, Turkey started to import beef, butchery and live cattle and sheep lamb. However, the carcass and meat prices continued to increase in the mentioned period.

Though the exact number is not known, it is estimated that over 30.000 fattening businesses are in our country and more than 90% of these businesses can be considered as small enterprises.

Planning and projecting of animal shelters in our country in order to provide suitable environmental conditions and improvement of existing shelters are obligatory (Ekmekyapar, 1991). In our country where climatic conditions show differences between regions, the preparation of shelter plans suitable for the special conditions of the regions depends on making the right facility place.

Beef farming are made to increase the yield of beef. In order to obtain more meat and better quality meat from younger males and females who do not have breeding features , milk cattle breeders who are left out of the squad, and old cattle, a special nutrition is applied for a certain period of time before slaughter, and this period is called beef farming (Hacıbebekoğlu et al., 2013). Basic methods of fattening are classified as pasture fattening (entansive), before pasture fattening than barn fattening (semi-entansive), and barn fattening (entansive).

Ecological animal husbandry is a production system which number of animal is low, has appropriate feeding and shelter conditions, produce quality product with an appropriate production and marketing methods. In order to design an ecological animal husbandry, firstly selection of facility place should be performed appropriately and correctly. The climate, the location and the environmental factors are very important for the place selection.

People make decisions in many topics during their lives. These decisions have certain criteria. The weights of these criteria vary from person to person. In the selection of the place of beef farming, weights of criteria, the selling, pastures, temperature or IPARD (Rural Development Component) support vary from person to person. A similar situation applies to businesses. For this reason, Multi Criteria Decision Making (MCDM) techniques are used to make scientific selection in evaluating alternatives when people or businesses make decisions.

In this study, it was aimed to determine the most suitable place of beef farming in cities in Turkey by using multi criteria decision making techniques. Firstly, a literature research was carried out. In the first phase of the study, the criteria affecting the place of beef farming were determined. The weights of these criteria were determined by AHP method and the accuracy was checked by consistency analysis. According to these criteria, 81 provinces were evaluated with the help of Visual Promethee package program which is application software of PROMETHEE method. In the last part, the results of the work were evaluated and various suggestions were made.

2. LITERATURE REVIEW

2.1. Studies Related with Animal Husbandry and Beef Farm

With parallel to the importance of the livestock sector in our country, there are a number of studies published in this issue.

Vural and Fidan (2007) reported that the activities types of animal husbandry in Turkey village-type; animal husbandry, agricultural worker activity which feeds one or two cattle's for family consumption, businesses which performs animal husbandry as a subsidiary activity beside herbal production, specialized businesses to produce to market and modern animal.

Çakır and Saner (2005) compared the traditional and ecological animal husbandry systems in Turkey, which have a great potential in terms of animal existence They evaluated necessity and applicability of ecological system in Turkey. Also they evaluated current situation according to different animal production branches.

Kıral (1993) made a general assessment of the technical and economic aspects of beef cattle breeding in Turkey by focusing on the importance and the role of fattening businesses in Turkish livestock farming. In the study, economic analysis of fattening enterprises which have 1-50 heads cattle's , 51-100 heads 1-50 heads cattle's and over 100 heads 1-50 heads cattle's were carried out by emphasizing the socio-economic structure of fattening enterprises (Şerefoğlu, C., 2008).

Ömürbek et al. (2013) have tried to determine the areas where livestock can be done in the province of Isparta by using the AHP method in. In the study, 7 districts in Isparta province were compared according to 5 different criteria (location, environmental factors, labor force, investment costs, laws) and decided to optimum area.

2.1. Multi Criteria Decision Making Methods

Multi-criteria decision-making methods achieve the best compromised solution by evaluating available alternatives based on deterministic criterion values. Genç and Masca (2013) have included in their articles the statements that the decision maker can sort, group or choose among the available alternatives by means of the MCDM methods. When the literature is examined, it is seen that mathematical programming and multi-criteria decision making methods are generally used as numerical methods in the studies about the place of beef farming.

2.1.1. Analytical Hierarchy (AHP) Method

Saaty (1990) describes the Analytical Hierarchy Process (AHP) method, which is widely used in computer aided multi-criteria decision making defined as a method which is based on pair wise comparison of alternatives according to common criterion and provides important support to the decision maker in reaching the conclusion of multi-criteria and multi-choice.

Cengiz and Çelen (2003) used AHP method in rural development, and at the end of the study, they found that AHP is a useful method in rural development studies. Eren (2006) worked on the determination of the best location in the leather industry using the AHP method. Özdağoğlu (2008) used the fuzzy analytical production process to study the facility place, and as a result he determined the appropriate one from the 4 alternative place (Ülke, 2016).

Özel et al. (2014) used the AHP method to select the appropriate area for the new forestation work to be carried out in the Bartın Havza. Bakan (2013) used the AHP method to determine the appropriate districts among 14 district alternatives to establish the university according to the criteria determined.

2.1.2. PROMETHEE

The PROMETHEE method is a multi-criteria prioritization method developed by Jean-Pierre Brans in 1982. The PROMETHEE method has been developed because of the difficulties in implementation of existing prioritization methods in the literature, and used in many studies (Dağdeviren and Eraslan, 2008). Some of the PROMETHEE applications in the literature are summarized below.

Maragou and Tsakiri (2005) proposed that PROMETHEE can be used to simplify the selection process alternative practices of reducing the damage caused by flood problems. Dağdeviren and Eraslan (2008) used PROMETHEE in supplier selection problem of a firm and calculate priority of alternative suppliers.

Athawale and Chakraborty (2010) have solved the problem of more efficient plant place selection using the PROMETHEE II method, and found out how much this selection has an effect in production organizations.

Kutay and Tektüfekçi (2013) used the PROMETHEE-GAIA approach which is one of computer-aided multi-criteria decision-making processes in order to determine the managerial accounting decisions according to the degree of importance. Soba (2012) applied PROMETHEE method, using criteria- prices, fuel, maximum speed, safety, horsepower and performance - for six different panelvans in the same class.

Genç and Masca (2013) used TOPSIS and PROMETHEE methods to obtain ranking of European Union countries and Turkey according to some economic criteria and compared their results. As a result, it is seen that the ranking values obtained as a result of the PROMETHEE and TOPSIS methods are very similar to each other

3. DATA AND METHODOLOGY

Many different methods can be used together to choose the place of beef farming. AHP, ANP, Data Envelopment Method, ELECTRE, TOPSIS, PROMETHEE and integration of these methods are generally used. Selection criteria and method selection are the most important features of facility place selection. When the place of beef farming is selected, AHP and PROMETHEE methods which be able to show the difference in the comparison stage and including mathematical solution will be used. By using these methods, a different point of view has been tried to be achieved.

3.1. Analytic Hierarchy Process (AHP)

The AHP method is based on the naturalness of the human brain in view of events (Kokangül and Susuz, 2009). AHP method is based on pair wise comparison of alternatives according to the criteria. The AHP provides decision support in solving multi-criteria and multi-alternative problems (Çiftçioğlu, 2013). The AHP was firstly introduced by Myers and Alpert in 1968. The AHP method developed by Saaty in 1977 consists of five basic steps (Ömürbek et al., 2014). In the first step, the problem is defined, the main objective is determined in the hierarchical structure. In the second step, hierarchy of criteria and alternatives is created. In the third step; pair wise comparison matrices are created. In the fourth step, weight vector is

found. In the last step, the consistency rate is calculated. In case of inconsistency, pair wise comparisons are controlled, and process is repeated until they are consistent.

3.2. PROMETHEE

The PROMETHEE method, which was introduced in 1982, is a multi-criteria sorting method. PROMETHEE is a method that allows for the sorting of alternatives taking into account conflicting criteria. This method starts with the evaluation table. Alternatives in this table are evaluated according to different criteria. There are two kinds of information needed for the implementation of PROMETHEE. These are function preferences of the decision maker and the relative importance of the considered criteria for comparing the contribution of alternatives in each criterion. The PROMETHEE method consists of 7 steps (Dağdeviren and Erarslan, 2008); In the first step; The data matrix A = (a, b, c, ...) is constructed for the alternatives evaluated by the weight of w = (w1, w2, ..., wk) and the criterion c = (f1, f2, ..., fk). In the second step, the preference function is defined for each criterion. In the third step, the common preference functions for alternative pairs are determined on the basis of preference functions. In the fourth step, the preference index for the alternative is calculated based on the common preference functions. In the fifth step; Positive (Φ +) and negative (Φ-) superiorities are determined for each alternative. In the sixth step; Partial priorities are determined by PROMETHEE I. Partial priorities explain the relationships of alternatives to each other. These relations are stated the preference situations of the alternatives and the determination of alternatives that are indifferent from each other. In the seventh step, full priorities for alternatives are calculated using PROMETHEE II and the calculated values are evaluated in the same plane as all alternatives and the exact order is determined.

3.3. Methodology

In this study, the selection of the facility place of beef farming among the provinces in Turkey was realized by using AHP and PROMETHEE methods. Criteria weights were determined by the AHP method and the ordering of the provinces was performed by PROMETHEE method.

As a result of the study, AHP and PROMETHEE methods were used to determine the ideal provinces and various suggestions were developed. Microsoft Excel and Visual Promethee package programs are used for calculations of AHP and PROMETHEE methods.

In the first phase of the study, the basic criteria required for the establishment of the beef farming were determined with interviews with the staff of the Directorate of Agriculture and in order to obtain expert opinions before the implementation. These criteria are, percentage of feed crop area, percentage of beef demand, urban development index, average minimum temperature, average maximum temperature, annual precipitation amount, grant supports. The criteria are shown in Table 1 below.

Table 1: Affecting criteria of selection of facility place

	Criteria
H1	Percentage of feed crop area
H2	Percentage of beef demand
H3	Urban development index
H4	Average minimum temperature
H5	Average maximum temperature
H6	Annual precipitation amount
H7	IPARD grant supports

In the second stage of study, As a result of the interviews among the employees, the relations between the criteria were determined, and the analytical averages of the values were determined, and the percentages were calculated, and the AHP weight matrix was formed. Weight percentages of the criteria were found as a result of AHP analysis. The accuracy of the weighted matrix was checked by consistency analysis. The results are shown in Table 2.

Table 2: AHP Weighted Matrix

	H1	H2	H3	H4	H5	H6	H7	W
H1	0,26	0,24	0,26	0,29	0,29	0,29	0,23	0.26
H2	0,49	0,46	0,35	0,41	0,41	0,41	0,45	0.42
H3	0,04	0,06	0,04	0,02	0,02	0,02	0,05	0.03
H4	0,05	0,07	0,09	0,06	0,06	0,06	0,05	0.06
H5	0,05	0,07	0,09	0,06	0,06	0,06	0,05	0.06

H6	0,05	0,07	0,09	0,06	0,06	0,06	0,05	0.06
H7	0,09	0,09	0,17	0,12	0,12	0,12	0,1	0.11
When ratio is 0,01295 <0,10 result is consistent								

In the third stage of the study, the appropriate preference function was chosen to implement the PROMETHEE method, and these values are shown in Table 3. After the preference function is determined, Visual Promethee program is used to implement the method. ("1" is not included in IPARD grant grants, "0" is set.) (1 - number of cattle / population) * 100.

Table 3: PROMETHEE Preference Functions and Data

CRITERIA	**H1**	**H2**	**H3**	**H4**	**H5**	**H6**	**H7**
Units	%	%	Index	-°C	+°C	mm	IPARD
Preferences							
Min/Max	Max	Max	Max	Max	Min	Max	Max
Weights (W))	0.26	0.42	0.03	0.06	0.06	0.06	0.11
Preference Functions	Type I (Usual)	Type I (Usual)	Type I (Usual)	Type II (U-Shape)	Type I (Usual)	Type I (Usual)	Type III (V-Shape)
Q	n/a	n/a	n/a	-10,00	+30,00	n/a	n/a
P	n/a	n/a	n/a	n/a	n/a	n/a	1,00
Provinces							
ADANA	2,36	90,21	0,47	5,5	34,6	654,4	0
ADIYAMAN	0,63	85,50	0,38	1,3	37,7	695,1	0
AFYONKARAHİSAR	18,85	55,54	0,63	-3,4	29,5	427,3	1
AĞRI	32,70	37,77	0,30	-15,9	30	519,6	1
AMASYA	16,06	46,60	0,59	-0,9	31,3	459,4	1
ANKARA	28,43	93,10	0,62	-3	30,3	366,2	1
ANTALYA	16,05	93,16	0,55	6	34,2	1074,6	0
ARTVİN	53,15	65,88	0,63	-0,2	26,2	701,8	0
AYDIN	2,57	67,40	0,54	4,3	36,1	644	1
BALIKESİR	6,63	58,76	0,63	1,3	31,3	576,8	1
BİLECİK	19,06	83,37	0,63	-0,3	28,6	454,5	0
BİNGÖL	27,10	54,17	0,44	-6,1	34,5	944,1	0
BİTLİS	10,23	83,13	0,40	-6,4	30,7	1221,2	0
BOLU	17,87	52,75	0,66	-3,2	28	553,9	0
BURDUR	7,86	21,62	0,55	-0,8	32,2	426,9	1
BURSA	15,58	93,44	0,60	1,7	31	695,1	1
ÇANAKKALE	9,70	60,70	0,59	3,2	30,7	628,8	1
ÇANKIRI	35,96	30,00	0,60	-4	31,1	410,6	1
ÇORUM	34,85	59,11	0,50	-4,2	29,4	431,5	1
DENİZLİ	5,92	75,41	0,59	2,3	34,4	566,1	1
DİYARBAKIR	2,94	79,02	0,35	-2,3	38,5	490,1	1
EDİRNE	0,22	62,15	0,57	-0,6	31,7	602,4	0
ELAZIĞ	19,52	73,41	0,48	-4	34,2	410,2	1
ERZİNCAN	23,13	59,56	0,57	-7	31,9	376,8	1
ERZURUM	30,47	14,38	0,53	-14,3	27,3	406	1
ESKİŞEHİR	37,57	83,65	0,62	-3,5	29,3	370	0
GAZİANTEP	1,77	92,05	0,47	-0,8	35,3	552,7	0
GİRESUN	37,96	80,20	0,58	4,4	26,5	1266	1
GÜMÜŞHANE	32,24	52,17	0,53	-5,7	28,7	461,3	0
HAKKARİ	5,40	86,50	0,33	-8,1	30,9	781,3	0
HATAY	0,77	92,30	0,44	4,7	32	1128,9	1
ISPARTA	16,12	65,50	0,67	-1,8	30,5	545,4	1
MERSİN	11,22	94,16	0,48	6,3	31,5	588,4	1
İSTANBUL	0,25	99,49	0,65	3,1	26,8	813,2	0
İZMİR	2,41	85,74	0,60	5,9	33,2	690,3	0
KARS	0,92	-71,02	0,38	-16	26,2	499,3	1
KASTAMONU	20,62	40,77	0,57	-4,3	28	485,2	1

KAYSERİ	37,21	78,23	0,55	-6,8	30,8	388,6	0
KIRKLARELİ	0,83	58,82	0,58	0	30,5	570,2	0
KIRŞEHİR	40,54	31,74	0,59	-4,1	29,7	379,9	0
KOCAELİ	15,50	93,72	0,59	3,3	29,6	809,3	0
KONYA	30,09	64,10	0,62	-3,9	30,2	318,7	1
KÜTAHYA	14,19	68,37	0,65	-3,2	28,5	560,3	1
MALATYA	44,26	76,35	0,52	-3	34	382,6	1
MANİSA	4,50	83,97	0,58	3,1	34,9	730,1	1
KAHRAMANMARAŞ	12,52	84,38	0,53	1,3	35,9	732,1	1
MARDİN	2,38	89,16	0,29	0,5	34,9	666,4	1
MUĞLA	17,33	77,34	0,54	1,6	33,6	1169,4	0
MUŞ	15,88	23,20	0,28	-11,1	33	762,7	1
NEVŞEHİR	24,41	73,47	0,56	-3,8	28,4	415,2	1
NİĞDE	32,65	58,95	0,53	-4,5	29,4	336,4	0
ORDU	38,88	83,05	0,51	3,8	27,3	1035,1	1
RİZE	0,00	93,18	0,63	3,5	26,6	2245,3	0
SAKARYA	0,16	83,58	0,67	2,9	29,4	834,6	0
SAMSUN	8,14	76,93	0,57	3,9	27	706,3	1
SİİRT	15,13	93,70	0,42	-0,6	37,1	716,3	0
SİNOP	39,49	52,79	0,63	4,2	26,2	676	0
SİVAS	38,46	55,73	0,56	-7	28,5	432,3	1
TEKİRDAĞ	0,02	85,29	0,58	2,1	28,1	589,1	0
TOKAT	13,10	59,30	0,56	-1,7	29,6	432,4	1
TRABZON	0,00	83,62	0,60	4,4	26,8	810,2	1
TUNCELİ	51,16	64,35	0,45	-5,7	35,2	792,7	0
ŞANLIURFA	14,92	87,04	0,35	2,2	38,7	458,4	1
UŞAK	0,71	59,30	0,65	-1,2	30,6	544,5	1
VAN	32,96	85,17	0,37	-7,7	28,2	387,4	1
YOZGAT	27,11	44,98	0,49	-5,2	26,5	578,7	1
ZONGULDAK	3,08	88,36	0,55	3,4	25,3	1216,8	0
AKSARAY	34,99	49,29	0,51	-3,6	30,5	345,7	1
BAYBURT	21,89	-7,92	0,54	-10,8	27,4	438,3	0
KARAMAN	8,10	73,68	0,59	-3,8	31,1	329,4	1
KIRIKKALE	41,38	77,35	0,63	-2,9	31	382,6	0
BATMAN	0,94	90,82	0,32	-1,5	39,4	488,8	0
ŞIRNAK	4,89	92,90	0,32	-1,1	33,3	683,7	0
BARTIN	0,00	73,83	0,56	0,3	28,2	1040,5	0
ARDAHAN	28,45	-208,96	0,35	-17	24,4	554,1	1
IĞDIR	6,37	45,53	0,36	-8,1	33,2	256	0
YALOVA	37,88	95,33	0,62	3,3	28,7	748,7	0
KARABÜK	43,05	81,73	0,63	-0,5	32,5	489,8	0
KİLİS	9,50	94,28	0,40	2,2	36,2	494,6	0
OSMANİYE	0,85	87,34	0,41	3,4	34,2	834,8	0
DÜZCE	0,00	85,41	0,53	0,4	29	822,1	0

In the fourth stage of study, when selection problem of facility place for beef farming is solved with the help of Visual Promethee package program, it was seen that Yalova is the first, Kars was the last. The results are shown in Table 4.

Table 4: Sequencing Results Calculated by PROMETHEE

Provinces	Phi	Phi+	Phi-
YALOVA	0,6341	0,7604	0,1263
BURSA	0,4783	0,6800	0,2018
MERSİN	0,4560	0,6696	0,2136
ANKARA	0,4420	0,6623	0,2203
GİRESUN	0,4355	0,6590	0,2235

ORDU	0,4350	0,6591	0,2241
KOCAELİ	0,4069	0,6464	0,2395
ANTALYA	0,4061	0,6464	0,2403
SİİRT	0,3168	0,6023	0,2855
KİLİS	0,2764	0,5819	0,3055
KARABÜK	0,2701	0,5776	0,3075
VAN	0,2636	0,5736	0,3100
HATAY	0,2528	0,5680	0,3153
ŞANLIURFA	0,2499	0,5664	0,3165
İSTANBUL	0,2483	0,5676	0,3194
MALATYA	0,2324	0,5576	0,3253
KAHRAMANMARAŞ	0,2199	0,5510	0,3311
ESKİŞEHİR	0,2151	0,5509	0,3358
ARTVİN	0,1756	0,5304	0,3548
KIRIKKALE	0,1744	0,5294	0,3550
MARDİN	0,1689	0,5263	0,3574
ZONGULDAK	0,1656	0,5261	0,3605
RİZE	0,1649	0,5201	0,3553
MANİSA	0,1530	0,5178	0,3648
ŞIRNAK	0,1359	0,5116	0,3758
ADANA	0,1136	0,5005	0,3869
BİLECİK	0,1121	0,4986	0,3865
İZMİR	0,1011	0,4939	0,3928
KAYSERİ	0,0796	0,4831	0,4035
SAMSUN	0,0760	0,4793	0,4033
MUĞLA	0,0750	0,4810	0,4060
TUNCELİ	0,0720	0,4799	0,4079
OSMANİYE	0,0685	0,4781	0,4096
GAZİANTEP	0,0546	0,4710	0,4164
NEVŞEHİR	0,0445	0,4635	0,4190
TRABZON	0,0347	0,4538	0,4190
HAKKARİ	0,0160	0,4519	0,4359
BATMAN	0,0069	0,4471	0,4403
ISPARTA	0,0063	0,4448	0,4385
SİNOP	-0,0014	0,4419	0,4433
KONYA	-0,0025	0,4400	0,4425
DENİZLİ	-0,0128	0,4345	0,4473
BİTLİS	-0,0136	0,4369	0,4505
ELAZIĞ	-0,0145	0,4344	0,4489
KÜTAHYA	-0,0199	0,4315	0,4514
SİVAS	-0,0335	0,4245	0,4580
ÇORUM	-0,0366	0,4235	0,4601
ADIYAMAN	-0,0479	0,4194	0,4673
SAKARYA	-0,0651	0,4111	0,4763
DÜZCE	-0,0688	0,4039	0,4726
TEKİRDAĞ	-0,0724	0,4071	0,4795
DİYARBAKIR	-0,0726	0,4051	0,4778
ÇANAKKALE	-0,0768	0,4025	0,4793

AYDIN	-0,0986	0,3921	0,4908
KARAMAN	-0,1173	0,3823	0,4995
ERZİNCAN	-0,1210	0,3808	0,5018
AKSARAY	-0,1520	0,3656	0,5176
AFYONKARAHİSAR	-0,1535	0,3638	0,5173
TOKAT	-0,1573	0,3600	0,5173
ÇANKIRI	-0,1785	0,3520	0,5305
BALIKESİR	-0,1970	0,3420	0,5390
NİĞDE	-0,1990	0,3436	0,5426
YOZGAT	-0,2009	0,3414	0,5423
BİNGÖL	-0,2126	0,3374	0,5500
AMASYA	-0,2253	0,3283	0,5535
AĞRI	-0,2289	0,3274	0,5563
KIRŞEHİR	-0,2431	0,3214	0,5645
KASTAMONU	-0,2480	0,3173	0,5653
GÜMÜŞHANE	-0,2610	0,3126	0,5736
BARTIN	-0,2641	0,3064	0,5705
BOLU	-0,2673	0,3103	0,5775
UŞAK	-0,2851	0,2963	0,5814
ERZURUM	-0,3076	0,2873	0,5949
ARDAHAN	-0,3371	0,2729	0,6100
MUŞ	-0,3519	0,2659	0,6178
EDİRNE	-0,3864	0,2501	0,6365
KIRKLARELİ	-0,4244	0,2311	0,6555
BURDUR	-0,4365	0,2230	0,6595
BAYBURT	-0,4620	0,2125	0,6745
IĞDIR	-0,5860	0,1509	0,7369
KARS	-0,6233	0,1300	0,7533

4. CONCLUSION

Manufacturers in the livestock sector live problems such as inefficiency, inadequate use of technology, unconsciousness in union or cooperative style organization, and disconnection from the market as it is in many other sectors.

The productivity and profit increase which are main objectives of the animal husbandry will increase depending on the correct selection of the facility place of the enterprises. According to traditional methods, there are many problems in control and reduction of costs due to high investment, operating costs and labor demands in agricultural holdings in cattle feeding which are carried out in closed system stalls (Toker et al., 2010).

In this study, AHP and Promethee methods of multi criteria decision making methods were used for choosing beef farming place. For selection; seven criteria were defined as feed plant area percentages, percentage of demand of meat, city development index, average minimum temperature, average maximum temperature, annual rainfall quantity, grant supports. It is aimed that to select the correct criteria for the beef farming place, to find the weights of these selected criteria, and to determine the preference functions of the found weights were performed with Visual Promethee program.

As a result of the evaluations, it was determined that preference of Yalova province is suitable for the selection of the place. The province of Kars should be the last one to be preferred.

When the top ten provinces in the list are evaluated, it is seen that they are the ones in Marmara, Mediterranean, Black Sea, Central Anatolia and Southeastern Anatolia regions in Turkey. These five regions are emerging as priority regions for the establishment of a new district.

REFERENCES

Athawale, V.M. ve Chakraborty, S. 2010. "Facility Location Selection Using PROMETHEE II Method", Proceedings of the 2010 International Conference on Industrial Engineering and Operations Management Dhaka, Bangladesh,pp: 59-63.

Bakan, H., 2013, "Analitik Hiyerarşi Yöntemiyle Üniversite Kurulması Uygun Olan İlçelerin Belirlenmesi" Ekonomi ve Yönetim Araştırmaları Dergisi, 2(2), s:43-63.

Çiftçioğlu, B., 2013, "İnşaat Sektöründe AHP Yöntemi İle Alt Yüklenici Seçimi: Bir Konut Projesinde Uygulama", Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, İstanbul.

Çukur, F., Saner G., 2005, "Konvansiyonel ve Ekolojik Hayvancılık Sistemlerinin Sürdürülebilirliği ve Türkiye Üzerine Bir Değerlendirme", Adnan Menderes Üniversitesi, Ziraat Fakültesi Dergisi, 2(1), s:39-44.

Dağdeviren, M. ve Eraslan, E. 2008, "PROMETHEE Sıralama Yöntemiyle Tedarikçi Seçimi", Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 23(1): 69-75.

Ekmekyapar, T., 1991, Hayvan Barınaklarında Çevre Koşullarının Düzenlenmesi, Atatürk Üniversitesi Ziraat Fakültesi Yayınları, Erzurum.

Genç, T., Masca, M. 2013, "Topsis ve Promethee Yöntemleri ile Elde Edilen Üstünlük Sıralamalarının Bir Uygulama Üzerinden Karşılaştırılması", Afyon Kocatepe Üniversitesi İİBF Dergisi, 15(2): 539-566.

Hacıbebekoğlu, A., Hacıbebekoğlu M., Çetinel, S., Ülver, M. 2013, Büyükbaş Hayvan Besiciliği Yatırım Fizibilitesi, Karacadağ Kalkınma Ajansı Destek Programı.

http://arastirma.tarim.gov.tr/gaputaem/Belgeler/tar%C4%B1msal%20veriler/gaputaem%20gncel/HAYVANCILIK%20RAPORU%202015.pdf

http://www.tepge.gov.tr/dosyalar/yayinlar/7b4e10c046074a319ca607fb3c7aae15.pdf

http://www.tkdk.gov.tr/ProjeIslemleri/DesteklenenIller

http://www.tuik.gov.tr/HbGetir.do?id=21871&tb_id=3

https://www.mgm.gov.tr/veridegerlendirme/il-ve-ilceler-istatistik.aspx

Kokangül,A., Susuz, Z., 2009, "Integrated analytical hierarchy process and mathematical programming to supplier selection problem with quantity discount", Applied Mathematical Modelling, cilt: 33(3), pp. 1417-1429.

Kutay, N., Tektüfekçi, F., 2012, "Yönetsel Muhasebe Kararlarının Verilmesinde Bir Araç Olarak Promethee Sıralama Yöntemi ve Bir Uygulama", Verimlilik Dergisi, 3: 83-96.

Maragoudakı, R. ve Tsakiris, G., 2005, "Flood Mitigation Planning Using PROMETHEE", European Water, 9(10): 51-58.

Ocaklı, B., Ertosun, Z., Demiröz, M., Horasan B., Peker, Ş., 2017, "Türkiye'de Hayvancılığın Durumu ve Et İthalatı", http://www.zmo.org.tr/resimler/ekler/cb16f6f3938e162_ek.pdf, accessing date: 15th April 2017

Ömürbek, N., Karaatlı, M, Eren, H., Şanlı,B., 2014, "AHP Temelli PROMETHEE Sıralama Yöntemi İle Hafif Ticari Araç Seçimi", Süleyman Demirel Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi, cilt: 19, sayı: 4, pp. 47-64.

Ömürbek, N., Üstündağ, S., Helvacıoğlu, Ö., C., 2013, "Kurulus Yeri Seçiminde Analitik Hiyerarşi Süreci (AHP) Kullanımı: Isparta Bölgesi'nde Bir Uygulama", Yönetim Bilimleri dergisi, 11(21), s:101-116.

Özel, H. B., Karayılmazlar, S., Demirci, A., 2014, 'Bartın Havzasında Analitik Hiyerarşi Prosesi (AHP) Yöntemiyle Akdeniz Çam Türleri (Pinus brutia Ten. ve Pinus pinea L.) Kullanılarak Yapılacak Ağaçlandırma Çalışmaları İçin Yer Seçimi' , II. Ulusal Akdeniz Orman Ve Çevre Sempozyumu, Isparta.

Saaty, l.T., 1990, "How To Make A Decision: The Analytic Hierarchy Process", European Journal of Operational Research, 48: 9-26.

Soba, M., 2012, "Promethee Yöntemi Kullanarak En Uygun Panelvan Otomobil Seçimi ve Bir Uygulama", Journal of Yaşar University, 28(7): 4708-4721.

Şerefoğlu, Coşkun, 2008, SAPARD_IPARD programlarının Analiz Edilmesi-Besi İşletmelerinin IPARD Programındaki Yeri Üzerine Bir Araştırma, Atatürk Üniversitesi, Fen Bilimleri Enstitüsü Doktora Tezi, Erzurum.

Toker, M.T., Akdeniz, R. C., Ayhan, V., Boyar, S., Bayhan, A. K., Ünal, H. B., Yılmaz, H.İ., 2010, "Açık Besi Sığırcılığı İşletmelerinin Tasarımı", Türkiye Ziraat Mühendisliği VII. Teknik Kongresi. TMMOB Ziraat Mühendisleri Odası (ZMO), Cilt:2, S.759-765, Ankara.

Ülke, K., 2016, AHP Yöntemiyle Çorum İlinde AVM Kuruluş Yeri Seçimi, Hitit Üniversitesi Sosyal Bilimler Enstitüsü Yüksek Lisans Tezi, Çorum.

Vural, H., Fidan, H., 2007, "Türkiye'de Hayvansal Üretim ve Hayvancılık İşletmelerinin Özellikleri", Tarım Ekonomisi Dergisi ; 13(2), s: 49 - 59

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

DETECTING PHISHING WEBSITES USING SUPPORT VECTOR MACHINE ALGORITHM

DOI: 10.17261/Pressacademia.2017.582
PAP-WCTIE-V.5-2017(20)-p.139-142

Dogukan Aksu[1], Abdullah Abdulwakil[2], M. Ali Aydın[3]
[1]Istanbul Kültür University, Engineering Faculty, Computer Engineering Department, 34156, Bakırköy, Istanbul, Turkey. d.aksu@iku.edu.tr
[2]İstanbul University, Engineering Faculty, Computer Engineering Department, 34320 Avcılar, Istanbul, Turkey. abdullah19.right@gmail.com
[3]İstanbul University, Engineering Faculty, Computer Engineering Department, 34320 Avcılar, Istanbul, Turkey. aydinali@istanbul.edu.tr

ABSTRACT

Cybersecurity is one of the most important areas which aims to protect computers or computer systems, networks, programs and data from an attack such as; financial systems, biometric security systems, military systems, personal information security etc. Nowadays, there are a lot of rule-based phishing detection systems which are created to help people who can't understand which URL is real and which one is fake URL address. This paper proposes a method with supervised machine learning that classifies the URLs to legitimate and phishing. By using support vector machine (SVM) classification, a machine-learning algorithm, with an MATLAB-based computer program to give a warning message to the users about the reliability of the web page. In this paper, phishing detection system is implemented with SVM to avoid the internet users from becoming a victim of phishers to do not lose financial and personal information.

Keywords: Cyber security, phishing, machine learning, support vector machine, matlab

1. INTRODUCTION

In this fast growing modern technology driven world, the internet is one of the most important technology not only for individual users but also for organizations that have an online business. Most the organizations have an online business such as sales of product and services (Liu & Ye, 2001). This may put internet users to different types of risks which may result in loss of private information, identity theft, and financial losses (Abdelhamid, Ayesh, & Thabtah, 2014). Phishing is a method to deceive end-users to visit fraudulent web pages which have an almost same look and feel of the trusted web pages with the aim to steal personal and financial information. The information gained from the user such as usernames, passwords, social security numbers, and bank account details will give the attacker the ability, to imitate the victim and make financial transactions and put the victim in financial and emotional losses.

According to Anti-Phishing Working Group (APWG) report, In 2016 the total number of phishing attacks were 1,220,523 which show 65% increase over 2015 (Anti-Phishing Working Group, 2017). These websites can trap a lot of users and mislead them to expose their sensitive information, mainly the users who have less awareness about the safe usage of internet and do not inspect the web pages before exposing their personal information.

In this paper, we focused on web pages which have phishing URL address. Phishing is one of the major security issues in the modern technological world where all the information is in electronic format. To illustrate, if the user may not understand a phishing web page, the user may victimized of fraudulent act of phishing and phisher can have access to the bank account of the user and can withdraw money or the best case scenario is, if the phisher could not withdraw money at least he or she will have access to some valuable information of the victim.

Remaining parts of the paper is organized into the following sections. Section 2 explains a brief survey of the literature review. In Section 3, we present the features, SVM machine learning algorithm, and methodology. Section 4 highlights the finding and discussion. Finally, we conclude our paper in Section 5.

2. LITERATURE REVIEW

Numerous researchers have researched to eliminate phishing and guard online users against deceiving, but very few of these researchers are focusing on web page phishing detection. This section presents an outline of phishing detection methods. Broadly phishing detection is categorized into two, the first one is user education based and the second one is software based. Furthermore, the software based phishing detection is categorized into blacklist (heuristic), machine learning, and visual similarity. Machine learning methods utilize the features that are both common to phishing web pages and legitimate web pages such as length of the URL, Number of dots in the URL, At "@" symbol in URL, etc. In heuristic approach, the web page is considered as phishing if it matches the rules defined by the heuristic (Xiang, Hong, Rose, & Cranor, 2011). The drawbacks of phishing detection based on visual similarity and blacklist are, that they usually do not detect new phishing attack web pages(zero hour attack) (Jain & Gupta, 2016).

(Zhang, Hong, & Cranor, 2007) introduced CANTINA which is content-based approach and to reduce false positive rate some heuristics are also introduced.

3. DATA AND METHODOLOGY

Firstly, we create a dataset which consists of Phishing URLs have been taken from PhishTank (Phishtank) and for Non-Phishing URLs, the browsing history has been used.

Secondly, we defined the most known features in literature that are extracted from the dataset. These features are described as follows.

1- Long URL: Phishers are using long URLs to hide mistrustful parts in the address bar.

2- Dots: A legitimate URL can contain at most 5 dots. If a web page contains more than 5 dots it may be considered as phishing URL. For example, the following URL is considered as phishing because it contents more than five dots http://forccis.com.br/audit.com.verification.filling.information.ub/w2-form/ (Phishtank)

3- IP Address: IP addresses can be used instead of Fully Qualified Domain Name (FQDN). For example http://81.31.25.201/ok/rmf/pessoafisica.php (Phishtank). For security reasons legitimate websites do not use this method so if an IP Address exists in a web page link, it is declared as phishing. This features can be demonstrated as an important phishing detection feature because the online availability of phishing websites are few days.

4- SSL Connection: In general, web pages that have SSL Connection(HTTPS) can be considered non-phishing.

5- At (@) Symbol: Phishers may use @ symbol to give the impression of a legitimate internet address. Using @ symbol in the URL ignores whatever is there before @. This allows the phishers to write a legitimate URL before @ and hide the fraudulent part of the URL, for example, https://mail.google.com/mail/@http://www.phish1.com. Therefore if URL contains @ symbol it is considered as phishing.

6- Dash (-) symbol: The dash symbol is not commonly used in legitimate URLs. Phishers are adding suffix or prefix separated by dash (-) symbol to misguide the user about the originality of the web page, for example, http://appleid-icloud.in (Phishtank) is considered as phishing web page.

Innumerable real-world learning tasks have solved successfully by using support vector machine (SVM) that is machine learning algorithm developed by Vapnik (Cortes & Vapnik, 1995). In this study, binary classification has been used because we have had two different types of URLs phishing or non-phishing. Proposed algorithm in this paper is presented in **Figure1**.

Figure 1: The Design of Phishing Detection System

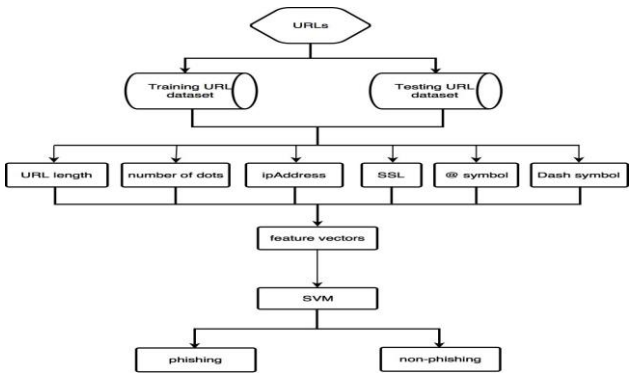


Figure1 represents the design of our system which is demonstrated as follows.

1) Phishing URLs have been taken from PhishTank (Phishtank) and for Non-Phishing URLs, the browsing history has been used.
2) In feature selection, six features which are URL length, the number of dots, ipAddress, SSL connection, at symbol(@), dash symbol(-) have been selected.
3) The selected features extracted from the dataset.
4) In training, SVM is trained via above-extracted features.
5) URLs which are not included in the training dataset are given to SVM classifier for testing.
6) SVM classifier returns either an URL is phishing or non-phishing. The confusion matrix for phishing detection can be seen in Table 1 (Fang, Koceja, Zhan, Dozier, & Dipankar, 2012)

Table 1: The Confusion Matrix for Phishing Detection

	DP	DNP
AP	TP	FN
ANP	FP	TN

- AP: Actual Phishing
- ANP: Actual Non-Phishing
- DP: Detected as Phishing
- DNP: Detected as Non-Phishing
- TP: True Positive
- TN: True Negative
- FP: False Positive
- FN: False Negative

4. FINDINGS AND DISCUSSIONS

The algorithm presented in the study is implemented using MATLAB programming language. The output of the system is given in Table 2.

Table 2: Output of the System

	DP	DNP
AP	22	3
ANP	2	73

Results in Table2 shows that the percentage of detected as phishing to the actual phishing is 88 % while the percentage of detected as non-phishing to the actual phishing is 12 %. Similarly, the percentage of detected as phishing to the actual non-phishing is 0.02 % while the percentage of detected as non-phishing to the actual non-phishing is 97 %.

Table 3: Performance Measurement

Measure	**Formula**	**Result(%)**
Accuracy	(TP+TN) / (TP+FP+FN+TN)	95
Recall	TP/(TP+FN)	88
Precision	TP/(TP+FP)	91,66
F1 score	2TP/(2TP+FP+FN)	89,79

Accuracy, recall, precision and f1 score are used for performance measurement (Shouval, Bondi, Mishan, Shimoni, Unger, & Nagler, 2014). The performance measurement of our system is shown in **Table3.** As it can be seen that accuracy, recall, precision and f1 score rates are 95 %, 88 %, 91.66 % and 89.79 % respectively. In this paper, by decreasing the number of features to 6 we have been able to achieve approximate accuracy to the study of (Akanbi, Amiri, & Fezaldehkordi, 2015)which has been obtained using 9 features. Despite the fact that we have decreased the number of features, our system produces promising results.

5. CONCLUSION

In this paper, the method that provides a detection system to prevent the users from being a phishing victim is presented. The first dataset has been created using Phishtank for phishing URLs and non-phishing URLs have been taken from browser history. Specified features are extracted from the training dataset and used for classification of URLs to phishing and non-

phishing. By using SVM classification algorithm, phishing URLs are detected. This method can detect phishing URLs but sometimes when the URL contains a feature which is not available in our features, the system gives the wrong result. Furthermore, features can be changed and affect the result according to URLs, therefore features must be defined dynamically. In future work, dynamically selecting the best features can be handled using a deep learning algorithm to get more accurate results.

REFERENCES

Abdelhamid, N., Ayesh, A., & Thabtah, F. (2014). Phishing detection based Associative Classification data mining. *Expert Systems with Applications*, 5948-5959.

Akanbi, O. A., Amiri, I. S., & Fezaldehkordi, E. (2015). *A Machine Learning Approach to Phishing Detection and Defense.* ELSEVIER.

Anti-Phishing Working Group, J. (2017, Feb. 23). *Phishing Activity Trends Report, 4th Quarter 2016.* Retrieved March 10, 2017, from APWG: https://docs.apwg.org/reports/apwg_trends_report_q4_2016.pdf

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning. 20(3): 273-297.

Fang, X., Koceja, N., Zhan, J., Dozier, G., & Dipankar, D. (2012). An Artificial Immune System for Phishing Detection. *IEEE World Congress on Computational Intelligence*.

Jain, A. K., & Gupta, B. B. (2016). Comparative Analysis of Features Based Machine Learning Approaches for Phishing Detection. *International Conference on Computing for Sustainable Global Development (INDIACom)*, (pp. 2125-2130).

Liu, J., & Ye, Y. (2001). Introduction to e-commerce agents: marketplace solutions, security issues, and supply and demand. *In E-commerce agents, marketplace solutions, security issues, and supply and demand*, 1-6.

Phishtank. (n.d.). Retrieved February 9, 2017, from OpenDNS: http://www.phishtank.com

Shouval, R., Bondi, O., Mishan, H., Shimoni, A., Unger, R., & Nagler, A. (2014). Application of machine learning algorithms for clinical predictive modeling: a data-mining approach. *Bone Marrow Transplantation*, 49, 332–337.

Xiang, G., Hong, J., Rose, C. P., & Cranor, L. (2011). Cantina+: A feature rich machine learning framework for detecting phishing web sites. *ACM Transactions on Information and System Security (TTSSEC)*, 14(2): p. 21.

Zhang, Y., Hong, J., & Cranor, L. (2007). Cantina: a content-based approach to detecting phishing web sites. *Proceedings of the 16th international conference on World Wide Web.*

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

DEVELOPMENT OF A HYBRID MODEL TO QUANTIFY THE SINGLE AND DOUBLE STRAND BREAKS FOLLOWING THE IRON 57 DE-EXCITATION

DOI: 10.17261/Pressacademia.2017.583
PAP-WCTIE-V.5-2017(21)-p.143-152

H.Oudira [1], D. Lotfi [1], A.Saifi [2]
[1]Faculty of Technology, Department of Electronics University Mouhamed Boudiaf of M'Sila , Algeria. Oudira_houcine@yahoo.fr
[2]Faculty of Technology, Department of Electronics University Mouhamed Boudiaf of M'Sila , Algeria. lotfi_djouane@yahoo.fr
[3]Faculty of Engineering Department of Electronics University of Constantine 1 , Algeria. Saifi19@yahoo.fr

ABSTRACT

The therapeutic utility of certain Auger emitters such as iodine-125 depends on their position within the cell nucleus. Or diagnostically, and to maintain as low as possible cell damage, it is preferable to have radionuclide localized outside the cell or at least the core. One solution to this problem is to consider markers capable of conveying anticancer drugs to the tumor site regardless of their location within the human body. The objective of this study is to simulate the impact of a complex such as bleomycin on single and double strand breaks in the DNA molecule. Indeed, this simulation consists of the following transactions: Construction of BLM -Fe- DNA complex, simulation of the electron's transport from the metastable state excitation of Fe 57 by the Monte Carlo method, treatment of chemical reactions in the considered environment by the diffusion equation. For physical, physico-chemical and finally chemical steps, the geometry of the complex is considered as a sphere of 50 nm centered on the binding site , and the mathematical method used is called step by step based on Monte Carlo codes.

1. INTRODUCTION

The toxicity of Auger emitters, such as iodine-125 is dependent on their position within the cell nucleus [1]. Thus, to reach the tumor cells, targeting agents that connect selectively to the double helix of the DNA molecule is used [2]. This is the case of bleomycin associated with a metal and an oxidant will be able to catalyze DNA strand breaks [3].

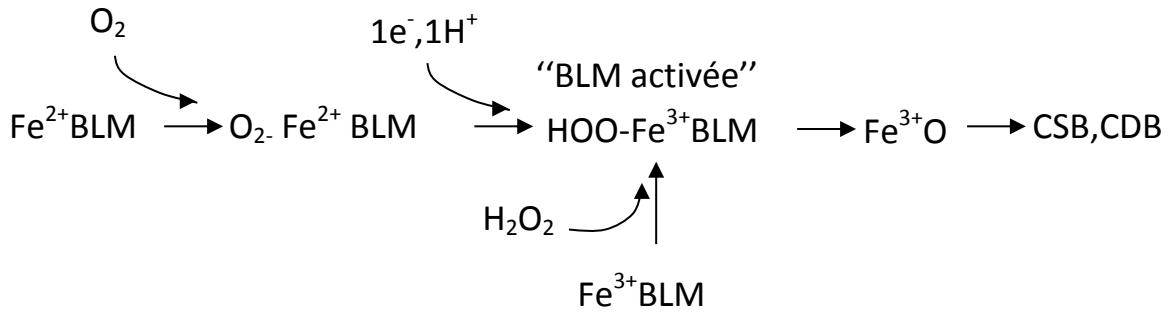
Indeed, the complex (HOO-Fe (III) BLM) said "activated bleomycin" is responsible for strand breaks induced on sites guanine-cytosine and / or guanine-thymine. In addition, it is postulated that this complex is the only active compound in vivo [4], which justifies the appropriateness of this study to compare the results obtained with those deduced from the radio-induced damage simulation by the disintegration of iodine 125 situated in chromosomal fiber [1].

The rest of this paper is organized as follows. In section 2 and 3 we introduce the problem position and the mathematical models used respectively. In section 4 presents the data used to solve the problem. The application, results and discussion are the aim of the sections 5 and 6. Finally, in the last section we present our conclusion and potential future work.

2. PROBLEM POSITION

To simulate the primary physical processes relative to a passage of an ionizing radiation through a biological medium; the Monte Carlo method of the type step by step which aims to reproduce the randomness of interactions [5] is used. Thus, software for tracing the full path target of an incident electron and secondary created through a DNA are developed. This work is firstly to be considered as a sequence of events leading to breakage of the strand as shown in Figure 1, and also to develop codes based on the excitation of the iron 57 incorporated in the DNA molecule. Indeed, this radionuclide emits by de-excitations a series of low-energy electrons, causing damage near to the site of decay. To reduce the computation time associated with the stochastic approach, we simulate the indirect effect of de-excitations through the diffusion equation [6-7].

Figure 1: Sequence of Events Leading to Strand Breaks



3. MATIMATICAL MODELS

3.1. Stochastic Model of Monte Carlo

The Monte Carlo method of the type step by step is adapted perfectly to the study of the physical and physic-chemical evaluation phases of the energy deposited in the target volume, to the extent that we consider the path λ that an electron can traverse in a heterogeneous milieu and the probability of interaction with a given entity.

$$\lambda = -\overline{\lambda(E)}.\log R \qquad (1)$$

$\overline{\lambda(E)}$ is the range and is given by the relation:

$$\overline{\lambda(E)} = \frac{1}{\sum_{ij} N_i \sigma_{ij}(E)} \qquad (2)$$

Where: R is an arbitrary number equally distributed between 0 and 1.

N_i is the number of atoms or molecules of the type *i* by unit of volume.

$\sigma_{ij}(E)$ is the effective section of interaction of the type *j* on the atom or the molecule of the type *i* for an incidental particle of energy *E.*

$$P_i = \sigma_i(E).N_i.\overline{\lambda(E)} \qquad (3)$$

P_i is the probability for that an incidental particle interacts with an atom or a molecule of the *i* component, knowing that there is interaction.

If there are k_i types of different interactions from the incidental particle with the *i* component, the *j* type of interaction having an effective section total $\sigma_{ij}(E)$ by atom or by molecule, the probability so that the interaction is of the type *j*, knowing that the interaction takes place with the *i* component, is thus :

$$P_{ij} = \frac{\sigma_{ij}(E)}{\sum_{l=1}^{k} \sigma_{il}(E)} \qquad (4)$$

During simulation, the point of interaction being fixed, we will have an interaction with the *i* component if *i* verifies:

$$\sum_{l=1}^{i-1} P_l < R_1 \leq \sum_{l=1}^{i} P_l \qquad (5)$$

and the interaction will be of the type j on the *i* component if *j* verifies:

$$\sum_{l=1}^{j-1} P_{il} < R_2 \leq \sum_{l=1}^{j} P_{il} \qquad (6)$$

R_1 and R_2 are arbitrary numbers equally distributed between 0 and 1.

At the end of the physical stage (10^{-15} Second) the irradiated molecules are in an excited or ionized state [7] whose fate is drawn by probabilities indicated above.

at 10^{-12} second, the electronic trace produced, in its neighbor, the chemical species such as :

e_{aq}^{-} ,H, OH, H2, H2 O2, and H_3O^+.

With regard to, the cross sections of biological interest molecules [8], we adopted the correction introduced by Laverne and Pimblott [9].

$$\sigma_{DNA} = \frac{A_{DNA}}{A_{water}} . \sigma_{water} \qquad (7)$$

Where A_{DNA} and A_{water} are respectively the molar masses of the DNA and water.

Each atom of the DNA is differentiated from the other atoms by introduction of an empirical grandeur known as ray of Van Der Waals (RVDW). The total cross section σ_i of an atom of VDW RVDW's ray is written:

$$\sigma_i(E) = \frac{R_{VDW}^i}{\sum_{Base} R_{VDW}^i} \frac{A_{Base}}{A_{water}} \sigma_{xater} \qquad (8)$$

Thus, we exploited the rays of VDW presented by Burkert [10] for the simulation of the electrons transport in the biological environment.

3.2. Diffusion Equation

To reduce the calculating time machine, relating to the study of the chemical phase, we chose a deterministic approach based on the resolution of the diffusion equation [11].

$$\frac{\partial C_i}{\partial t} = D_i \nabla^2 C_i + S \qquad (9)$$

- *Ci* is the concentration of the species of the type *i* (mole dm^{-3})

- $D_i \nabla^2 C_i$ is the product of the Laplacian of C_i by the constant of diffusion D_i ($cm^2\ s^{-1}$) of the species of the type *i*;

- *S* is the term source (mole $dm^{-3}\ s^{-1}$), considered as the algebraic sum of terms representing, following the chemical reaction taken in consideration, the impoverishment or the production of the species of the type *i*:

$$S = \left(\sum_w k_{iw} C_w + \sum_l \sum_j k_{lj} C_l C_j \right) - \left(k_i C_i + \sum_j k_{ij} C_i C_j \right) \qquad (10)$$

$k_{iw} C_w$, $k_{lj} C_l C_j$ represent respectively the production of species *i* following the dissociation of w, and the reaction of the species *l* with the species *j*. $k_i C_i$ and $k_{lj} C_l C_j$ represent respectively the impoverishment of species *i* following its dissociation, and its reaction with the particles of the type *j*; k_{iw} k_i, k_{lj} and k_{ij} indicate the constants of reaction's speed; k_{iw} and k_i are expressed in s^{-1}, while k_{lj} and k_{ij} are in dm3 $mole^{-1}\ s^{-1}$.

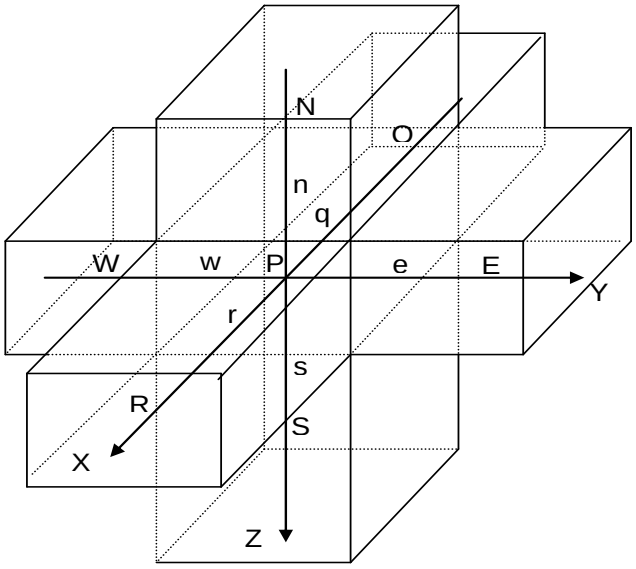
3.2.1 Discretizing the Diffusion Equation

To solve the equation (9), it is necessary to discretize it, that is to say, to replace it by simple algebraic equations expressing the same physical information that can be solved numerically. For this, converts the equation (9) in an integral equation by applying the weighted residuals method:

$$\iiint \int \frac{\partial C_i}{\partial t} W(X,Y,Z,t)dXdYdZdt = \iiint \int \left(D_i \nabla^2 C_i + F\right) W(X,Y,Z,t)dXdYdZdt \quad (11)$$

The choice of the test function W (X, Y, Z, t), and integration of equation (11) cause algebraic equations such that the concentration C_i in a point P (C_{iP} noted), the center of a elementary volume within the computational domain, is connected to those of the points E, W, R, Q, S, and N center of neighboring elementary volumes considered (Figure 2).

Figure 2: Immediate Neighboring of an Elementary Volume Inside the Three-Dimensional Grid



To access this formulation we have used the continuous finite element method [12], [13] and the method of centered differences [10]. Thus, the explicit method can replace the equation (11) by the following algebraic equation:

$$A_P C_{iP} = A_E C_{iE} + A_W C_{iW} + A_R C_{iR} + A_Q C_{iQ} + A_S C_{iS} + A_N C_{iN} + K \quad (12)$$

3.2.2. Calculation of Equation (12) Coefficients

For a point P situated inside the three-dimensional grid, integrating (11) on the following ranges:

$$\left[t\,;t+\Delta t\right], \left[X_P - \frac{\Delta X}{2}\,;\,X_P + \frac{\Delta X}{2}\right], \left[Y_P - \frac{\Delta Y}{2}\,;\,Y_P + \frac{\Delta Y}{2}\right], \left[Z_P - \frac{\Delta Z}{2}\,;\,Z_P + \frac{\Delta Z}{2}\right]$$

provides the coefficients A_P, A_E, A_W, A_R, A_Q, A_S, A_N and K relating to the equation (12) and whose values:

$$A_E = D_i \frac{\Delta X \Delta Z}{\Delta Y} = A_W$$

$$A_R = D_i \frac{\Delta Y \Delta Z}{\Delta X} = A_Q$$

$$A_S = D_i \frac{\Delta X \Delta Y}{\Delta Z} = A_N$$

$$A_P = \frac{\Delta X \Delta Y \Delta Z}{\Delta t} + A_E + A_W + A_R + A_Q + A_S + A_N$$

$$k = \frac{\Delta X \Delta Y \Delta Z}{\Delta t} C_{iP}^{0} + F_{P}^{0} \Delta X \Delta Y \Delta Z$$

ΔX , ΔY, ΔZ are spaces step in the three directions of the three-dimensional grid.

Δt is the time after which the changes in the concentration of a given radiolytic species are evaluated. The boundary conditions on the extreme points of the domain used to associate their coefficients whose value varies according to the position of the elementary considered volume [10]. Following integration of equation (11) on all the elementary volumes of space concentrations, we obtain another matrix system by block, whose resolution by the method of generalized Thomas matrix calculation (called TDMA) is immediate [14]. Also, we give a detailed attention to the calculation of the space step Δx whose value is reported with the temporal step Δt to ensure the stability of the solutions of the equation of diffusion, to respect the electronic balance of the milieu as much as that are possible. This relation uses the constant of diffusion D of the fastest particle $\left(H_{aq}^{+}\right)$.

$$\left(\Delta x\right)^2 \geq \sqrt{4D}.\Delta t \qquad (13)$$

4. DATA OF THE PROBLEM TO BE TREATED

4.1. BLM-Fe-DNA Complex

It is with regard to the DNA molecule, an octamer histonnes which may be considered as a kind of bobbin around which have two supercoils of a left DNA's supercoiled. The whole forms a disc diameter of about 110 Å, 57 Å of thickness, and average step (step between two successive turns 27-28Å) [8] and [9]. Bleomycin was built of bioplymer module molecular modeling software MSI, and iron was added and conformation of bleomycin modified to fulfill with the distances and angles listed in Table 1. The entire complex was then minimized by applying the algorithm of maximum slope and the conjugates gradients to keep geometry similar to that taken as a reference [10]. The iron atom is then excited to cause a jet of Auger electrons and increase the effectiveness of bleomycin in terms of single and double strand breaks (CDB and CSB) of the DNA molecule. It's called the Mossbauer Effect and Nuclear Resonance Gama (RGN).

Table 1: Parameters Used for the Minimization of BLM-Fe-DNA Complex

Total atoms Number	16302
Van Der Waals Truncation (VDW)	15Å
Minimization algorithm	Greater slope (500 steps) Conjugate Gradient (2000 steps)
convergence criterion	RMS=24 kcal $mol^{-1}Å^{-1}$
Initial potential energy	6 10^6 kcal mol^{-1}
Final potential energy	62400 kcal mol^{-1}
Time calculation machin	1 Day

4.2. Total and Differential Effective Sections.

The total and differential effective sections are given in the form of curves which can be seen in figures 3 and 4. Indeed, the effective sections are calculated here according to the energy of the electrons emitted by 57mFe de-excitation. This energy range of 10eV to 100 eV.

Figure 3: Total atomic cross sections 10 and 1000 eV

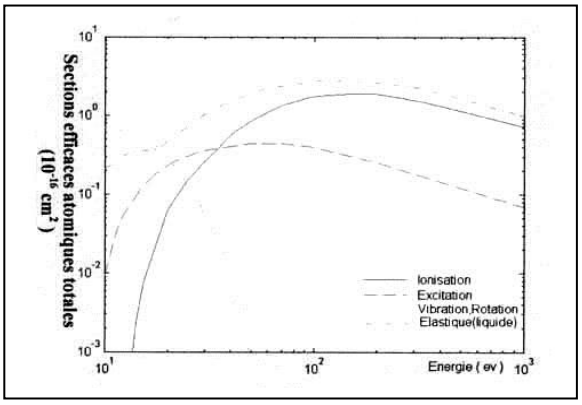
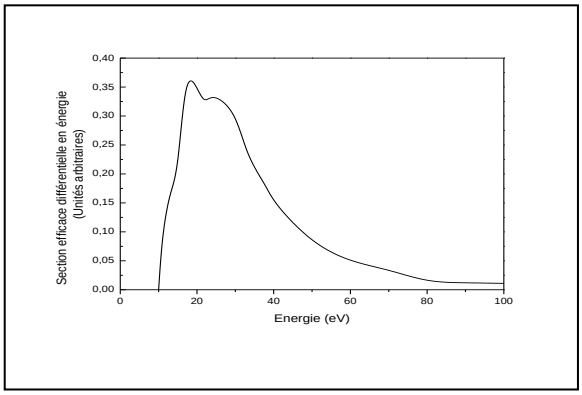


Figure 4. Echontionnage of the differentia cross between section l as a function of energy



4.3. The 57mFe De-excitation Spectra

The 57mFe de-excitation spectra are those obtained by Charlton [11] and Pomplun [12]. For both cases we have retained the part of the spectrum where the energy of the emitted electrons is between 1 eV and 100eV.

4.4. Coefficients of Diffusion and Chemical Reactions.

The created radiolitic species diffuse and react with molecules of biological interest. The values of the coefficients of diffusion and reaction rate constants are those given by Vrigneaud [10].

5. PUTTING IN COMPUTER FORM

The simulation of the diffusion and reaction processes of the radiolytic species until a time noted tmax is given by the flowchart in [1]. Indeed, a main program based upon three subroutines of space discretization, calculation of the concentrations and the radiochemical yields, and the system compression is developed. It is useful to note that the adoption of a logarithmic temporal step Δt is in adequacy, on the one hand with the taking in consideration of the whole of the chemical reactions of the radiolytic species at the beginning of the chemical phase, and on the other hand with the evolution of the species having the greatest constant of diffusion (H_{aq}^{+}) at the end of the this same stage. Thus, to the injection of a temporal step Δt corresponds the calculation of the equation coefficients (12) to the level of all elementary volumes constituent the three-dimensional grid and consequently the concentrations of all the species; this enables us to deduce the yields by using the equation (14). In our case, N_i was defined by the following relation:

$$N_i = \left[\sum_k \sum_l \sum_m C_i(k,l,m,t) \right] . (\Delta x)^3 \qquad (14)$$

When the computer system reaches its limits of managing different files, it uses the compression subroutine that recalculates the appropriate spatial step Δx and deduces Δt that need to inject. It is true that this procedure leads to an inaccuracy in the spatial location of the particles in question, but has the advantage of being able to continue our calculations until a time tmax

6. RESULT AND DISCUSSION

The choice of the spatial step Δx depends on the value of the temporal step Δt and therefore by accuracy and stability of the solutions of the diffusion equation (12).

The spatial discretization subroutine calculates the optimum step space from the Cartesian coordinates of the various radiolytic species from the physical and physico-chemical phases and deduces the initial concentrations at time 10^{-12} s. To illustrate the importance of the finesse in solving the diffusion equation, we present in Figures 5, 6 and 7 radiochemical yields of major radiolytic species (OH, e_{aq}^{-} , H) and for Δx = 10 Å , Δt = {10^{-10} sec; 10^{-11} s; 10^{-12} s}.

When Δt = 10^{-10} s, the relation (13) does not hold, and thus the yields are oscillating curves without any possible meaning to be interpreted. For Δt = 10^{-11} s, the relation (13) holds, however the yields are fairly flat. The temporal step is too large, which results in the failure to take into account a number of reactions. Finally, when Δt = 10^{-12} s, the yields obtained are comparable to those obtained from other models (Vrigneaud 2000). To optimize machine time calculation, we adopted a logarithmic time step in line with the process of diffusion and reaction of radiolytic species. In figure 8, it can be seen, for e_{aq}^{-} a clear concordance between our results (they have been derived based on the spectra Pomplun et al (1987) or Charlton and Humm (1985)) with those of Vriggneaud [10]. It is clear that the spectrum of Pomplun led to the formation of more of e_{aq}^{-} because of the electrons number of very low energy that it contains.

The yields obtained from the deterministic method are comparable to those deduced from the stochastic model taking into account the statistical fluctuation (Figure 9). Besides this similarity between results (yields) from the two approaches is confirmed when calculating simple and double strand breaks in the DNA molecule. Indeed, Figures 10 and 11 show the same tendency to increase when the energy of the electrons from the de-excitation of 57mFe increases.

Figure 5: Radiochemical yields of OH, H and e_{aq}^{-} according to the time for an electron incident 700 eV in liquid water (Δt = 10^{-10} s).

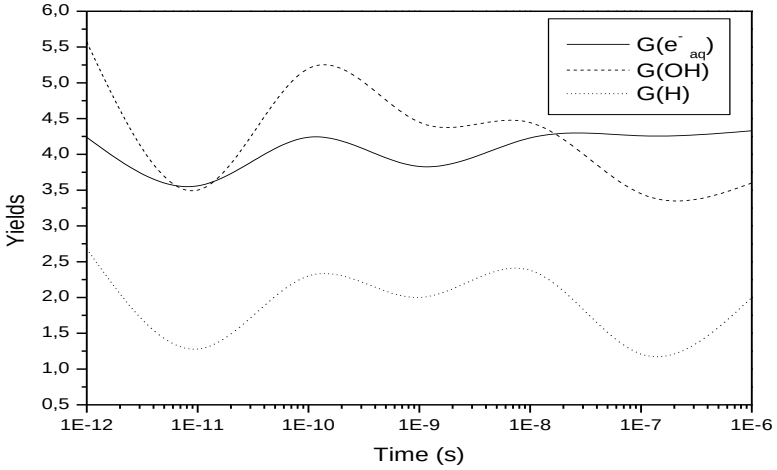


Figure 6: Radiochemical yields of OH, H and e_{aq}^{-} according to the time for an electron incident 700 eV in liquid water (Δt = 10-11 s).

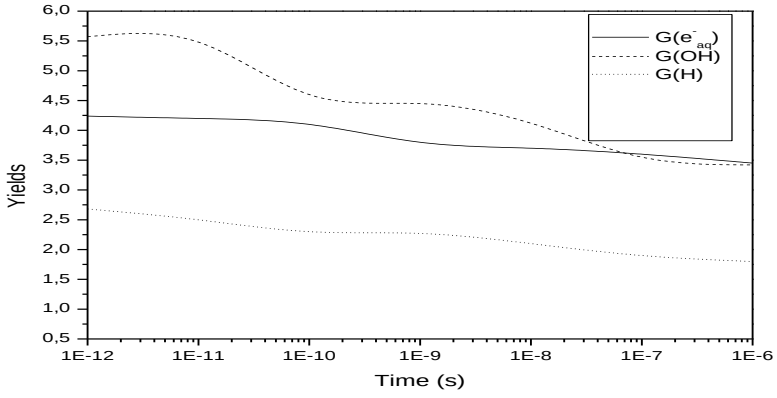


Figure 7: Radiochemical yields of OH, H and e_{aq}^{-} according to the time for an incident electron of 700 eV in liquid water (Δt = 10-12 s).

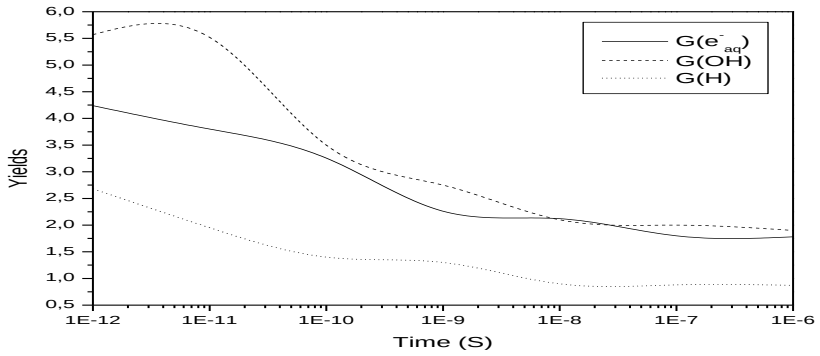


Figure 8: Comparison of yields e_{aq}^{-} with other models

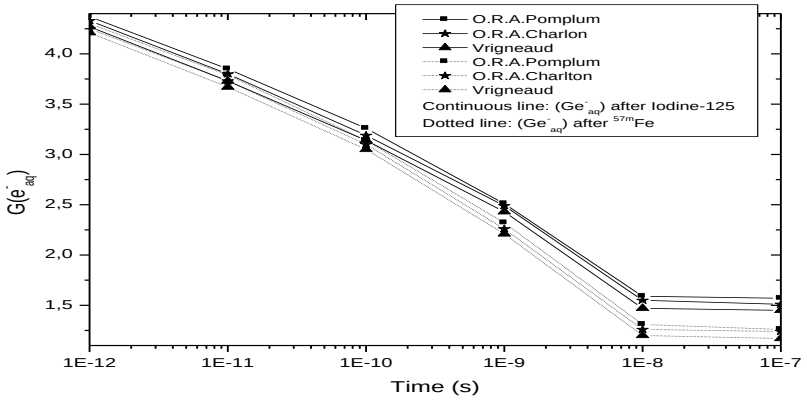


Figure 9: Variation of G (e_{aq}^{-}) as a function of time

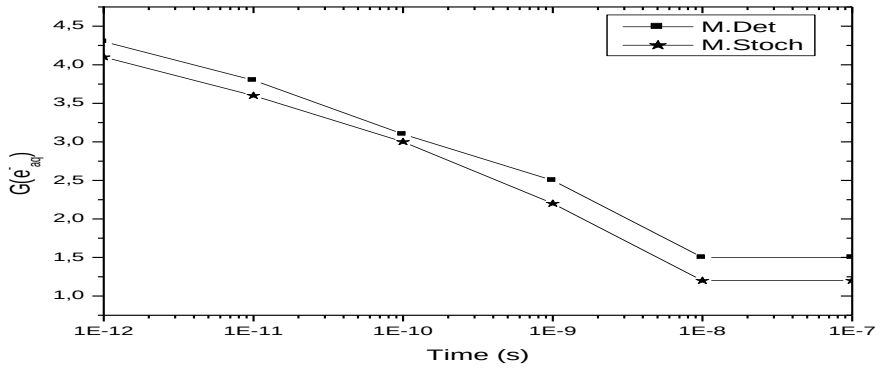


Figure 10: Yield of simple-stranded breaks according to the energy of emitted electrons by the excitation of 57mFe

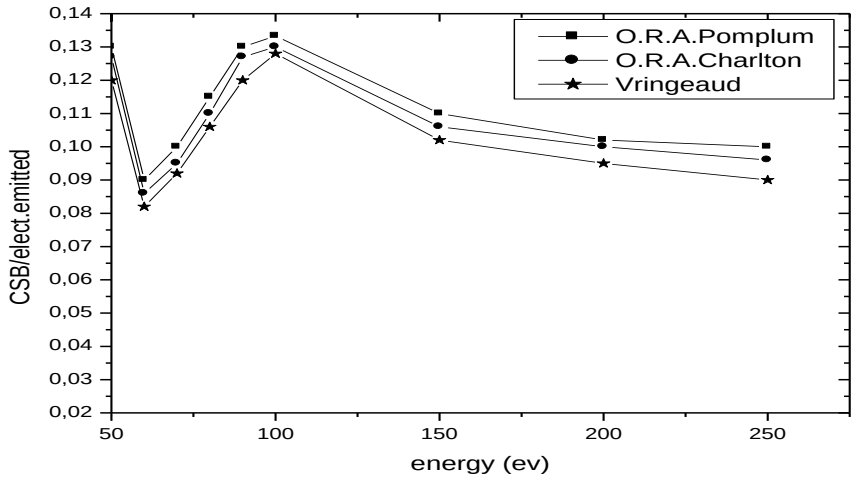
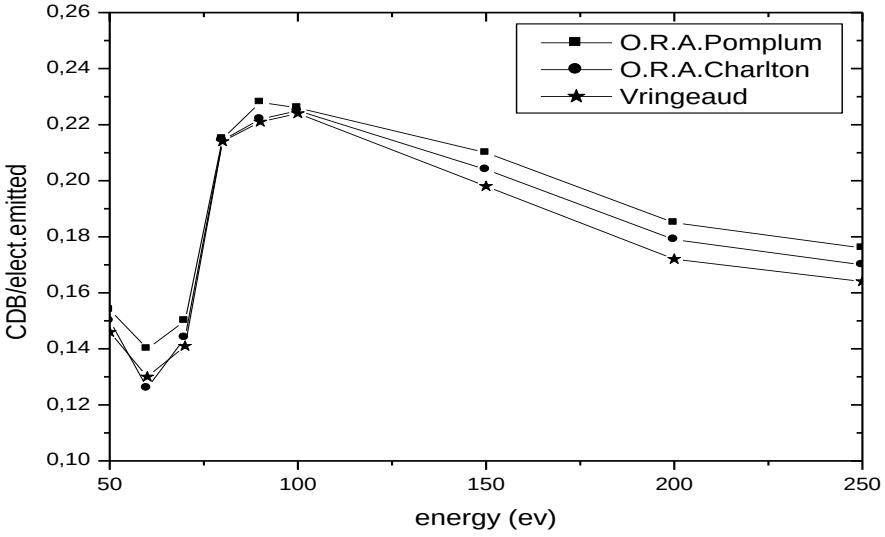


Figure 11: Yield of double strand breaks as according to the energy of emitted electrons by the excitation of 57mFe



7. CONCLUSION

Through the data necessary for the simulation of damage radiation induced by electrons after the de-excitation of the iron 57 to the DNA molecule through the diffusion equation, we can notice that:

Our results are comparable to those obtained by the Monte Carlo method of type step by step. Furthermore, we show the inefficiency of analytical methods in the time interval [10^{-12}s, 10^{-6}s] where the concentrations of radiolytic species are very unstable because of the numerous chemical reactions between radicals and subunits of DNA molecule. The developed software programs are powerful enough to the extent where we can integrate multiple parameters that can

affect the performance of breakages caused to the DNA molecule. Indeed, it may include several 57mFe sources for irradiating one or more cells.

REFERENCES

[1]- Oudira H., Saifi A., "Quantification of the simple and double strand breaks following the disintegration of Iodine 125 in situ in chromosomal fiber", *Radioprotection*, vol. 45, pp. 389-408, 2010.

[2]- Panyutin I.G., Neuman R.D., "Sequence-specific DNA breaks produced by triplex directed decay of iodine 125", *Acta oncologica*, vol. 35(7), pp. 817-823, 1991.

[3]- Burger R.M., Peisach J. et Horuvitz S.B., "Activated bleomycin: Atransient complex of drug, iron and oxygen that degrades DNA", *Biol.chem*, vol. 256(22), pp. 11636-11644, 1981.

[4]- Absalon M.J., Stuble.J. et Kozarich J.W., "Sequence specific double-strand cleavage of DNA by Fe-Bleomycin.1. The detection of sequence specific double strand breaks using hairpin oligonucleotides", *Biochemestry*, vol. 52, pp. 2065-2075, 1995.

[5]- Terrissol M., Pomplum E., "Computer simulation of DNA incorporated I^{125} auger cascades and of the associated radiation chemistry", *radiat, Prot. Dos*., vol. 52, pp. 177-181, 1994.

[6]- Djamai D., Oudira H., Saifi A., "Application d'un modèle hybride à l'étude des dommages radio-induits par un faisceau d'électrons sur la molécule d'ADN dans son environnement", *Radioprotection*, vol. 43, pp. 357-387, 2008.

[7]- Oudira H., Djamai D., Saifi A., "Application d'un modèle déterministe à l'étude de l'influence des molécules radioprotectrices sur les rendements des cassures simple et double brin de la molécule d'ADN", *Radioprotection*, vol. 43, pp. 389-408, 2008.

[8]- W. Saenger., "Principles of nucleic acid structure", *Ed. Springer-Verlag, New-York*, 1984.

[9]- Widom J., Towae d., "a unified model of chromatin folding", *Annu , Rev Biohys. Chem*., vol. 81, pp. 365-395, 1989.

[10]-Vrigneaud J.M., Ph.D, "Développement d'un modèle biophysique pour l'évaluation des dommages radio-induits dans la fibre chromosomique", (Université, CPAT de Toulouse, 2000).

[11]- D.E. Charlton and J.L. Humm, "A method of calculating initial DNA strand breakage following the decay of incorporated I^{125}", *Int.J.Radiat.Biol*, vol. 53, pp. 353-365, 1985.

[12]- Pomplum E., Booz.J. et Charlon.D.E., "A Monte Carlo simulation of Auger cascades", *Rad.Res*, vol. 111, pp. 533-552, 1987.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

APPLYING FUZZY LOGIC THEORY TO PERFORMANCE MANAGEMENT

DOI: 10.17261/Pressacademia.2017.584
PAP-WCTIE-V.5-2017(22)-p.153-161

Cemal Ardil[1]
[1]Okan University, cemalardil@mail.com

ABSTRACT

Organizations usually apply classical methods of employee performance evaluation. In this classical system, employee performance depends on work results, and it is evaluated only as success or failure in job behaviors. The non-classical performance evaluation methods such as fuzzy logic may mainly be used to many forms of decision-making including artificial intelligence systems. This research proposes a new employee performance evaluation method based on fuzzy logic systems. The process of performance measurement for evaluating the effectiveness, efficiency, and productivity of employees encompasses data collection, data design, and data analysis stages and it involves a level of uncertainty associated with performance measures. In evaluating employee performance, it usually involves granting numerical values or linguistic labels to employee performance in the organization. The scores accorded by the appraisers are only approximations, which are then, used to represent each employee's achievement by reasoning incorporated in the computational methods. In this paper, the fuzzy logic theory approach is used to represent the uncertainty caused by performance measures during its design, use and analysis stages. This research seeks to describe and execute the fuzzy logic theory approach for decision making in the employee performance appraisal process. Finally, reasoning based on fuzzy logic theory provides an alternative way in dealing with imprecise data, which is often reflected in the way humans think and make judgments in real life.

Keywords: Performance management, performance appraisal, performance tracking, evaluation parameters, performance measurement.

1. INTRODUCTION

The evaluation of employees in the organization is the appraisal of the extent to which the employees have met their professional obligations with regard to organizational goals and objectives. The process of performance appraisal is a structured interaction between superiors and their subordinates in the organization. The work performance of the subordinates is usually considered and discussed by the process of embracing a periodic and systematic interview. As an instrument creating competitive advantages amongst employees, performance appraisal system is mostly used by organization management to evaluate the management of the effectiveness, efficiency, and productivity of employees within the organization. In the modern organizations, each employee is evaluated through a performance appraisal system that is designed to be a systematic annual process which involves evaluating employee's set targets, perceived behavior evaluation, and work achievement during the annual evaluation. This evaluation may be considered by the organization management in making decisions with respect to appointments, salary and any other purposes where an employee's performance may be a relevant consideration. Therefore, systematic performance appraisal and ranking of employees applying for evaluation and promotion is important in strategic human resource management. Performing performance appraisal of employees with skills and expertise in relation to a specific position is an important task for managing an organization's human resource information system. Superiors within organizations are concerned with performance appraisal, recognition decisions, and evaluations that they have to make on their subordinates. Nevertheless, subordinates in organizations increasingly realize the importance of performance appraisal system, which is affecting their rewards and future career path in the business domain. Global knowledge-based economy reminds all organizations of the importance of maintaining their highly skilled and talented knowledge employees (Moon et al., 2007).

Organization management essentially needs to discover, evaluate and promote qualified employees because valuable human skill and expertise are the main sources of competitive advantage for organizations. Thus, the creation of

performance evaluation criteria is an important requirement towards performance appraisal in organizations. The performance evaluation criteria also ensure fairness, objectivity, and transparency since appraisers determine the metrics of performance evaluation and the weighting among the metrics before aggregating the appraisal scores to determine the ranking of each subordinate in the organization. Superiors must objectively maintain a certain level of professional distance so that they can assign tasks, review and evaluate employee performance without bias in the workplace. The performance appraisal mainly focuses on the integration and achievement of employee targets, behaviors, and performance at work in comparison to the goals of the organization. Performance appraisal procedure is periodically conducted within an organization to track and examine the work performance of subordinates so as to identify the strengths and weaknesses as well as opportunities for engagement and improvement among employees. As an employee development and coaching instrument for organization management, the performance appraisal results are then used to determine if a particular employee should be eventually terminated or reinforced to give an essential evaluation of an employee's readiness for promotion and serving as the foundation for giving merit bonus. Thus, the employee ranking is determined by applying the fuzzy set operations and membership functions. In order to make the decision for the promotion rank of subordinates in the organization, it is proposed a fuzzy evaluation and ranking procedure in conjunction with integrated performance appraisal and promotion ranking system for organization management (Shaout and Al-Shammari, 1998).

The fuzzy logic theory is based on the notion of relatively graded membership, as inspired by the processes of human perception and cognition. Following the awareness and understanding of the conceptual framework of fuzzy sets, and fuzzy logic theory, researchers from diverse fields, have increasingly applied the fuzzy logic theory to the performance appraisal process. Fuzzy logic theory presents proper ways of managing multiple variables that have imprecise data to apply fuzzy logic concept reasoning, which reflects the way of human thinking and inference mechanism of action. So far various studies have been largely conducted to improve the application of fuzzy logic theory in areas of performance appraisal and performance measurement. The fuzzy logic theory is highly suitable and applicable basis for developing knowledge-based systems in performance appraisal for tasks such as the selection of employees, the evaluation of various training methods, the team ranking, and the real-time monitoring of employees data. Performance appraisals are mainly used for judgmental and developmental purposes in order to make good administrative decisions (Schweiger and Sumners, 1994).

Performance measures are meant to provide complete and precise information about an entity's performance. Organizations usually rely on the appraisers' skills and expertise, when they are expected to supervise and make an accurate judgment of how well their subordinates have performed over a period of time. However, the process of performance appraisal is often flawed by the appraisers own biases and information-processing problems. Therefore, the evaluation process may be involved with information of uncertainty and subjectivity. The process of employee performance evaluation is usually involved with awarding numerical values or linguistic labels to employee performance. In most evaluations, these values and labels are used to represent the employee's achievement by reasoning using computational methods. However, in the evaluation using the fuzzy logic method, the performance of the appraiser usually involves the measurement of ability, competence, job behaviors, and skills, which are fuzzy concepts that may be captured during the performance appraisal process (Shaout and Yousif, 2014).

Therefore, in this paper, it is proposed a new fuzzy evaluation method using fuzzy rule approach in the multicriteria analysis, an application of fuzzy logic theory to decision making process in the domain of performance appraisal system. Consequently, the fuzzy logic theory approach can be implemented to handle the uncertainty information involved in employee performance evaluation in the organization. This paper is basically divided into four sections. First, the introduction describes the conceptual topic of this research paper. Second, some related works along with the approach and methods in this study were explained. The third section discusses the results of this study and the final section concludes the study.

2. RELATED WORK AND BACKGROUND

2.1. Performance Appraisal and Performance Management

A performance appraisal, performance review, performance evaluation, or employee appraisal is an evaluation method by which the work performance of an employee is grounded and evaluated. Performance evaluation method and evaluation model have extensively been applied in many fields such as decision analysis, supporting system, and system engineering as well as in performance management. Performance appraisal is a formal performance management information system that ensures the evaluation of the quality of an employee's performance in an organization. There are a vast number of applicable performance appraisal systems in organizations. Organizations use a variety of methods for evaluating employee performance. There are so many types of performance management methods, understanding how each of them works assists determine the proper one to use within an organization. Each type of performance management methods has benefits and drawbacks; however, an assessment of the workforce, management style, and business environment makes the decision easier for evaluators (Shaout and Trivedi, 2013).

Performance appraisal provides the means to evaluate an employee's current and past performance relative to the employee's performance standards in the organization. It is a systematic process which involves creating work standards, evaluate employees' actual performance relative to those work standards, and giving feedback to employees so as to motivate them to improve the job performance or to eliminate performance deficiency. Some potential aims of performance appraisal system might include identifying particular job behaviors, abilities or skills. Various appraisal methods including graphic rating scale, behavioral checklist, management-by-objectives, forced distribution, multi-rater feedback, performance ranking, combined manager-employee appraisal, work planning and review, trait scales, critical incident, narrative, and criteria-based and peer review have been proposed to evaluate the performance of an employee in organization (Venclová et al., 2013).

As a consequence, with all the available applicable methods, it is crucially essential to understand that different organization might use different methods for assessing employees' performance. Since all the aforementioned methods have their own advantages and disadvantages, most organizations might mix and match different methods for their own performance appraisal system that can realize their organizational needs. Performance appraisal system has already become one of the most valuable management instrument by which organization management use to achieve collective goals and objectives. In order to ensure that the results of the performance appraisals are useful and reasonable to the supervisor when evaluating their subordinates, the performance appraisal system should consistently produce reliable and valid results for the management of an organization.

Table 1: Comparison between Performance Appraisal and Performance Management

Performance Appraisal	Performance Management
top down assessment	joint process through dialogue
performed annually	continuous reviews
use of ratings is very common	use of ratings is less common
linked to traits and characteristics	linked to quantifiable objectives,values, and behaviors
monolithic systems	flexible systems
often linked to pay	is not directly linked to pay
rigid structure / system	supple / flexible process
operational	strategic
usually housed in the human resource department	conducted by managers and supervisors
individualistic	holistic / collective
quantitative approach	combines qaulitative and qualitative approaches
retrospective for corrections	prospective, future-oriented for growth
not linked to business needs	linked to business needs
often linked to compensation	not usually linked with compensation
bureaucratic-complex paperworks / documents	less concerned with documentation

The contemporary organizations are experiencing a transformation for coping with the changing needs of the environment and excelling in the business for managing change proactively. The performance appraisal system is no longer adequate the needs of the changing environment as it was an employee evaluation process in which the appraisers were impelled to make subjective decisions and judgments about job behaviors, and the performance of the employees against the predetermined job standards. Performance appraisal system exercises control and monitors the activities of the employees through disciplinary actions and management of rewards and promotions . As a result of the globalization of competitive business, and internationalization of human resources, the organizations have changed their focus from performance appraisals to performance management that focuses on observed job behaviors, and concrete results from the previously established strategic goals and objectives.

2.2. Fuzzy Logic Theory and Performance Management

The fuzzy logic theory method has largely been applied to several performance appraisal systems. The fuzzy logic theory based multi-criteria assessment study in the group decision-making of promotion screening recommended that the methodology is a good model for a transparent and fair multi-criteria performance evaluation in military organizations (Moon et al., 2007). In many circumstances, the fuzzy set theory is successfully used to solve multiple criteria problems; appraiser tends to use vaguely defined qualitative criteria in evaluating the performance of subordinates (Jing, R.C et al., 2007). When applying fuzzy set theory to fuzzy group decision support system, it assists the decision maker to make better decisions and recommendations under different circumstances and alternatives. The multifactorial evaluation model is proposed for the application of the fuzzy logic theory to a decision-making process in information, decision, and control systems. The reviewed works on fuzzy approach support the fuzzy logic theory as a conceptual framework for use in the development of the performance appraisal system because fuzzy logic theory allows performance appraisal system to be developed by using fuzzy variables and relationships (Moon et al., 2007).

3. RESEARCH METHODOLOGY

The decision making is a process of problem-solving that involves pursuing of goals under constraints, while the outcome is a decision that results in an action (Chan, D.C.K. et al., 2002). This difficult process is basically involved with incomplete and imprecise information subjectivity, and linguistics factors, which tend to be present to some degree (Gokmen et al, 2010). As knowledge involved in employee appraisal evaluation is approximate, and fuzzy logic theory is successfully used for approximate reasoning in such circumstances, its application becomes significant to manage the uncertainty in the evaluation system. In fuzzy logic applications, fuzzy reasoning resembles human decision making with an ability to generate precise solutions from certain or approximate information (Berenji and Khedkar, 1998). The advantage of fuzziness dealing with imprecision, vagueness, and uncertainty of human expressions fits ideally into decision-making systems modeled in the fuzzy logic theory (Garibaldi and Ifeachor, 1999). Fuzzy logic theory approach has extensively been used to evaluate many types of performances and comprehensive evaluations such as environmental evaluation, weather forecast, teaching evaluation, risk management, power generation, car retrieval system, marketing, finance, manufacturing, consumers, government and so on (Zeng and Feng, 2014). In accordance, this research is also concerned with the fuzzy logic theory which mainly aims to handle the uncertainty information and human-like reasoning and approached the domain problem in performance management.

3.1. Fuzzy Logic Theory

The theory of fuzzy sets, whose members are vague objects, was introduced to model uncertainty in subjective information and analysis of complex systems (Zadeh, 1965). Crisp set of input data is generally received in terms of linguistic judgments and beliefs in natural language, which is then converted to the form of fuzzy sets in order to provide a base for logical and mathematical reasoning in information, decision, and control systems (Zadeh, 1975). A fuzzy set is represented by a membership function defined on the universe of discourse, where the fuzzy variables are defined.

Let X be a non-empty set, and fuzzy set A in X, the universal set, is characterized by μ_A ; its membership function described as

$$\mu_A : X \rightarrow [0,1] \tag{1}$$

where μ_A is interpreted as the degree of membership of element x in the fuzzy set A for each x ∈ X. [0,1] denotes the interval of real numbers from 0 to 1. Thus, the value 0 is used to represent complete non-membership, and the value 1 is used to represent membership fully, and values in between are used to represent intermediate membership degrees. The fuzzy set, A, is usually denoted by the set of pairs

$$A= \{ x, \mu_A(x)), x \in U \} \tag{2}$$

When *U* is a finite set $\{x_1,...,x_n\}$, the fuzzy set on *U* may also be represented as:

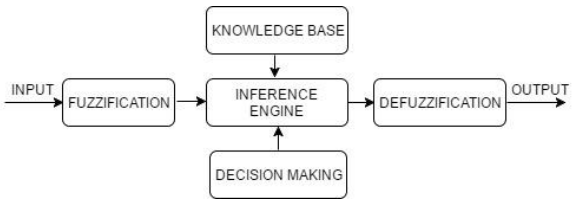
$$A = \sum_{i=1}^{n} x_i / \mu_A(x_i) \tag{3}$$

When *U* is an infinite set, the fuzzy set may be represented as

$$A= \int_x x / \mu_A(x) .$$

A fuzzy logic system with fuzzification, rule evaluation, and defuzzification parts as shown in Figure 1 can be defined as the nonlinear mapping of a crisp set of input data set to a scalar output data.

Figure 1: Fuzzy System Diagram



1) Fuzzification (using membership function to graphically describe situation in environment, crisp set of input data are gathered and converted to a fuzzy data set using fuzzy linguistic variables, fuzzy linguistic terms and membership functions)

2) Rule evaluation (application of fuzzy rules, an inference is made based on a set of rules)

3) Defuzzification (obtaining output data or actual results, the resulting fuzzy output is mapped to a crisp output using the membership functions)

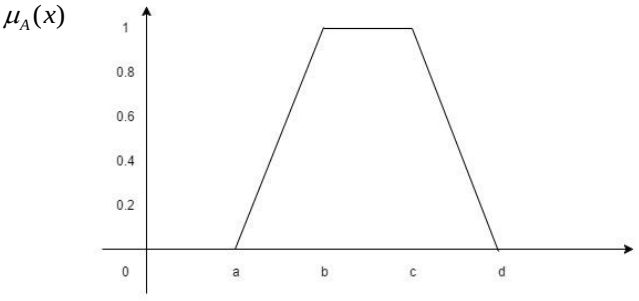
Fuzzy set theory, which generally involves three phases of fuzzification, fuzzy inference, and defuzzification, provides a means to represent uncertainties, and deals with problems in a vague environment of perception of things and human thinking. The fuzzy logic theory has successfully been applied in various fields with keywords “performance,” “evaluation,” “severity,” “possibility,” “importance,” and “satisfaction,” but their meaning varies with the situations (Yadav and Singh, 2011).

A fuzzy set contains elements that have varying degrees of membership in the set. This fuzzy logic theory transforms linguistic variables into fuzzy sets to replace the crisp sets. The values of linguistic variables are not just numbers, but verbal variables or sentences in a natural or artificial language. The fuzzy concept is very useful in complicated situations to be appropriately described by quantitative expressions.

The membership states all information contained in a fuzzy set. Membership functions of fuzzy sets must be precisely defined in respect of function type and function parameters. Both the parameters and shape of the membership functions strongly influence the accuracy. The well-known membership functions are namely triangular, trapezoidal, and Gaussian type. In this study, the trapezoidal membership function is preferred as shown in Figure 2. This function is very often used in practice for its simplicity. A small amount of data is needed to define the membership function. In addition, this type membership is very effective, especially for class interval-based works. Figure 2 shows the main parts of a trapezoidal membership function. The function specified by four parameters given by ‘‘A = trapezoid (x, a, b, c, d)’’:

Figure 2: Trapezoidal Fuzzy Membership Function

$$\mu_A(x) = f(x;a,b,c,d) = \begin{cases} 0, x \le a \\ \dfrac{x-a}{b-a}, a \le x \le b \\ 1, b \le x \le c \\ \dfrac{d-x}{d-c}, c \le x \le d \\ 0, d \le x \end{cases}$$



or,by $\mu_A(x) = f(x;a,b,c,d) = \max(\min(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}), 0)$

3.2. Fuzzy Multifactorial Evaluation

The fuzzy multifactorial evaluation method is used to quantify employee performance function in an organization. The method of trapezoidal distributed function is considered as a bridge in the mathematical method. Firstly, based on trapezoidal distributed function, the membership function of every grade (high, good, medium and low) is established about every target according to each target value of measurement of employee performance. Secondly, the number of

the weight of every target is considered based on work data and experience. Finally, the result is obtained based on the principle of maximal membership degree of multifactorial evaluation. The multifactorial evaluation method is explained with the targets of the performance function.

A multifactorial evaluation system requires three elements: A set of principal influence factors, $U = \{u_1, u_2, ...u_m\}$, A set of verbal grades, $E = \{e_1, e_2, ..., e_p\}$ and for every object $u \in U$, there is a single-factor evaluation matrix $R^{(u)} = (r_{jk}(u))_{mxp}$. With the preceding three elements, for a given $u \in U$, its evaluation result $D^{(u)} \in F(E)$ can be derived. The multifactorial evaluation model suggests that a mapping consists of factor weights (W) and relation matrix (R) which combines with an aggregation function as follows (Tutmez et al., 2007):

$$\zeta : U \to F(E) \tag{4}$$

$$u \to \zeta(u) \text{ and } D^{(u)} = f(W, R^{(u)}) \tag{5}$$

where ζ is the decision criterion, which is used to evaluate alternatives. Thus, the decision maker may choose a decision function that best reflects the goals of the decision. Aggregation operators which require different transformations for the judgments are used for evaluating different types of decision behaviors. The simultaneous satisfaction of all the decision factors can be modeled by using *t*-norm operators. *t*-Norms are fuzzy set versions of intersection operation on sets. In this sense, they are used for the conjunctive type of aggregation. Therefore, the *t*-norm based operator is considered for aggregation. The transformation function which leads to weighted minimum (and maximum) operators that can be applied in the setting of the possibility theory (Dubois and Prade, 1986). The weighted minimum is given by

$$D^{w}(\mu_1, \mu_2, ...\mu_m) = \Lambda_{i=1}^{m} \left[(1 - w_i) \vee \mu_i \right] \tag{6}$$

The set of principal influence factors on employee performance function can basically be described as follows:

$$U=(u1 , u2 , u3, u4) \tag{7}$$

where,the principal influence factors ui (i=1 ,2, …4) may be fuzzy or non-fuzzy in nature, but the mathematical relation between ui and U is only given by ui$\in$U.

The four targets of performance function are selected as (Pmax100, Pmax80, Pmax60, Pmax40). The grade of evaluation of the performance function is targeted into four classes, and the standard of the grade is established based on work data of performance function. Thus, based on the needs of performance considerations, the rank of performance function is classified into four sorts as high, good, medium and low. Evaluation criteria should be based on the actual situation, and evaluation criteria are divided into four grade levels. Then, let V_I = high degree, V_2 = good degree, V_3 = medium degree, V_4 = low degree. Therefore, fuzzy sets of the rank of performance function for the evaluation set is given as follows

$$V = \{V_I, V_2, V_3, V_4\} \tag{8}$$

Evidently, the final evaluated results vj (j=1,2,3,4) are mainly obtained from fuzzy multifactorial evaluation approach. Indeed, fuzzy multifactorial evaluation process aims to obtain thoroughly assessed crisp results based on comprehensive consideration of principal influence factors ui.

The principal influence factor set is formed in vector U={u_I, u_2, u_3, u_4}. The level of influence of each factor for the evaluated results is not identical, for the purpose of representing the degree of influence of each factor, the appropriate weight should be given to each factor wi (i=1 ,2, … 4), thus, set of weight W is established. The distribution of a number of weights for all the factors, such as U_i is, respectively, shown in vector W.

$$W= (w1 , w2 , w3, w4) \tag{9}$$

where W is fuzzy set on U. Evidently, W is a part set of fuzzy sets of U, indicated as follows

$$W = \frac{w_1}{u_1} + \frac{w_2}{u_2} + \frac{w_3}{u_3} + \frac{w_4}{u_4} \tag{10}$$

Therefore, W is determined with the four principal influence factors ui, as follows

$$W=[0.20,0.35,0.3,0.15], W(1\text{-weights})=[0.80, 0.65, 0.70, 0.85] \tag{11}$$

The key to treating the problem with fuzzy multifactorial evaluation is building the appropriate subordinate functions of the influence factors.

Fuzzy multifactorial evaluation is mainly considered by the method of compound operation denoted as follows

D= W **O** R = (d1 ,d2 ,d3, d4) (12)

where R is a fuzzy relation matrix between the set U and V, and it decides a fuzzy reflection. W is the set of weight of factors, and the original image of fuzzy reflection, D is reflected image of fuzzy reflection, or, assessed result.

In fuzzy relation matrix R, Ri =(ri1, ri2, ri3, ri4) is single-factor evaluation of ui, and part set of fuzzy sets of V. d_j is evaluation target as follows

$$d_j = \sum_{i=1}^{n=4} a_i r_{ij} \tag{13}$$

d_j represents the calculation of fuzzy matrix. The proposed model not only considers the effect of the factors but reserves the information of single-factor evaluation of ui.

The single-factor evaluation-relationship matrix R based on membership function between factor set and evaluation set is given by.

$R^{(u)}$ = {(U_i, V_j, U_{rij}) / i == l, ..., 4; j=l, ...4.}

The operation is performed by $D^{(u)} = f(W, R^{(u)})$

$$D^{(u)} = f(W, R^{(u)}) \tag{14}$$

$$= [W1, W2, W3, W4]\mathbf{O}\begin{bmatrix} R11 & R12 & R13 & R14 \\ R21 & R22 & R23 & R24 \\ R31 & R32 & R33 & R34 \\ R41 & R42 & R43 & R44 \end{bmatrix} = (D1, D2, D3, D4)$$

$$D_j = \sum_{i=1}^{n} W_i R_{ij} \quad (j = 1, ..., 4) \tag{15}$$

This is the final result of the fuzzy multifactorial evaluation.

3.3. Applying Fuzzy Multifactorial Evaluation

A normal standard of every target is established (Pmax100, Pmax80, Pmax60, Pmax40). The high range is 80%-100%; the good degree is 60%-80%; the medium degree is 60%-40%; the low degree is under 40%. Therefore, the range of each grade is shown V_1[100, 80], V_2 [80, 60], V_3 [60, 40], V_4[40, 00]. The membership function is established between every factor and each class. The trapezoidal distributive function is adopted for computation. If f_{ij} (1) (i = 1, . . . , 4; j = 1, ..., 4) shows respectively trapezoidal membership function, which is factor U_i relative to the grade V_j.

The fij (x) is lined up in part as follows but others fij (x) are omitted for convenience.

$$fij(x) = \begin{cases} 1, & (80 <= x <= 100) \\ x/80, & (80 > x > 0) \end{cases} \qquad fij(x) = \begin{cases} \frac{1}{40}(100 - x), & (80 <= x <= 100) \\ 1, & (80 <= x <= 60) \\ 1/60(x), & (60 > x > 0) \end{cases}$$

$$fij(x) = \begin{cases} (100 - x)/60, & (60 <= x <= 100) \\ 1, & (60 >= x >= 40) \\ x/40, & (40 > x > 0) \end{cases} \qquad fij(x) = \begin{cases} (100 - x)/80, & (40 <= x <= 100) \\ 1, & 40 > x > 0) \end{cases}$$

For the above trapezoidal membership function, the value of membership function of each target can be reckoned respectively, and it is shown in relation matrix R. The four performance measurement results of an employee Pmax are 50%, 80%, 65% and 35% respectively. The relationship matrix R_I is formed accordingly.

Let the Pmax100, Pmax80, Pmax60, Pmax40 are subordinated to x respectively. Then every value of f_{ij} (Pmax 100) is computed as follows:

f_{11} (Pmax100) = 0.625 , f_{12} (Pmax100)= 0.833 , f_{13} (Pmax100) = 1, f_{14}(Pmax100)= 0.625 ... then R is obtained:

In the last stage of the evaluation, the evaluation vector is found by using the Dubois–Prade decision operator as follows:

then $D = f(W,R)$ = [0.80, 0.65, 0.70, 0.85] **O** $\begin{bmatrix} 0.625 & 0.833 & 1.000 & 0.625 \\ 1.000 & 0.500 & 0.333 & 0.250 \\ 0.812 & 1.000 & 0.583 & 0.437 \\ 0.437 & 0.583 & 0.875 & 1.000 \end{bmatrix}$

d1(u)= (w1 v r11(u)) ^ (w2 v r21(u)) ^ (w3 v r31(u)) ^ (w4 v r41(u))

d1(u)= 0.80 similarly,

d2(u)= 0.65 d3(u) = 0.65 and d4(u)= 0.65

Finally, the evaluation vector is obtained:

$D = f(W,R)$ = (d1, d2, d3, d4) = (0.80, 0.65, 0.65, 0.65)

Therefore, referring to the verbal grades V= {V_1=High[100-80], V_2=Good[80-60], V_3=Medium[60-40], V_4=Low[40-00], }, the obtained employee performance function should be rated as "Good". Finally, the method of fuzzy analysis of multiple-stage is used to get the result of evaluation for performance function of the employees based on the center principle of fuzzy gravity.

4. CONCLUSION

In this paper, the method of multifactorial evaluation is applied to the evaluation of employee performance function. The method of the trapezoidal distributive function is considered as a bridge in the mathematical method approach. The trapezoidal membership function is established between each evaluation target and evaluation object. The method of multifactorial evaluation is used to conduct performance management based on the assessment of employee performance function. Performance classification of employees can reasonably be executed using multifactorial fuzzy approach. With the fuzzy decision-making system, the employee performance function can be evaluated, and information can reasonably be obtained by data processing and different decision functions. The classification of employee performance function is very useful in the evaluation of scheduling, budgeting, and planning of the organizations. It can also be evaluated in employee selection and project scheduling as a decision parameter. The fuzzy multifactorial evaluation approach can reasonably be used in performance measurement.

REFERENCES

Adnan Shaout and Jaldip Trivedi (2013). "Performance Appraisal System – Using a Multistage Fuzzy Architecture".International Journal of Computer and Information Technology, Volume 02– Issue 03, pp.405-411, May 2013.

Adnan, S. and Minwir, A. (1998). "Fuzzy Logic Modeling for Performance Appraisal Systems – A Framework for Empirical Evaluation", Expert Systems with Applications, Vol. 14, No. 3, p. 323-328, 1998.

Adnan Shaout and Mohamed Khalid Yousif (2014). "Employee Performance Appraisal System Using Fuzzy Logic". International Journal of Computer Science and Information Technology, Vol 6, No 4, pp 1-19, August 2014.

B. Tutmez, S. Kahraman, O. Gunaydin.(2007)."Multifactorial Fuzzy Approach to the Sawability Classification of Building Stones". Construction and Building Materials 21, (2007), 1672–1679.

Chan, D.C.K., Yung, K.L., Ip, A.W.H. (2002). "An Application of Fuzzy Sets to Process Performance Evaluation", Integrated Manufacturing Systems, Vol. 13 Issue: 4, pp.237-246, 2002.

Dubois D, Prade H. (1986)."Weighted Minimum and Maximum Operations in Fuzzy Set Theory". Information Sciences, 39(2), pp. 205–210.

Garibaldi, J.M. and Ifeachor, E.C. (1999). "Application of simulated Annealing Fuzzy Model Tuning to Umbilical Cord Acid-base Interpretation", IEEE Transactions on Fuzzy Systems, Vol.7, No.1, pp. 72-84, Feb 1999.

G. Gokmen, T. C. Akinci, M. Tektas, N. Onat. G. Kocyigit, N. Tektas,(2010)."Evaluation of Student Performance in Laboratory Applications Using Fuzzy Logic", Procedia Social and Behavioral Science, 2(2010), pp 902-909.

Hamid R. Berenji, Pratap S. Khedkar(1998)."Using Fuzzy Logic for Performance Evaluation in Reinforcement Learning".International Journal of Approximate Reasoning 18, pp.131-144,1998.

Jing, R.C, Cheng, C. H. and Chen, L. S. (2007). "A Fuzzy-Based Military Officer Performance Appraisal System", Applied Soft Computing, Vol. 7, Issue. 3, p. 936-945, 2007.

L. A. Zadeh, (1965). "Fuzzy Set," Information Control, Vol. 8, No. 3, pp. 338-353, 1965.

L. A. Zadeh, (1975). “The Concept of a Linguistic Variable and its Application to Approximate Reasoning,” Information Science, Vol. 8, pp. 199-249, 1975.

Moon, C., Lee, J., Jeong, C., Lee, J., Park, S. and Lim, S. (2007). “An Implementation Case for the Performance Appraisal and Promotion Ranking”, in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Montréal, Canada, 7-10 October 2007.

Ramjeet Singh Yadav, Vijendra Pratap Singh (2011)."Modeling Academic Performance Evaluation Using Soft Computing Techniques: A Fuzzy Logic Approach".International Journal on Computer Science and Engineering,Vol. 3, No.2,Feb 2011,pp.676-686.

Schweiger, I. and Sumners, G.E. (1994), "Optimizing the Value of Performance Appraisals", Managerial Auditing Journal, Vol. 9 Issue: 8, pp.3-7, 1994.

Venclová Kateřina, Šalková Andrea, Koláčková Gabriela (2013).”Identification of Employee Performance Appraisal Methods in Agricultural Organizations”. Journal of Competitiveness.Vol. 5, Issue 2, pp. 20-36, June 2013.

Wenyi Zeng and Shuang Feng (2014)."An Improved Comprehensive Evaluation Model and its Application". International Journal of Computational Intelligence Systems, Vol. 7, No. 4, 706-714, 2014.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

DECISION MAKING TECHNIQUES FOR ELECTRONIC COMMUNICATION: AN EXAMPLE FOR TURKEY

DOI: 10.17261/Pressacademia.2017.585
PAP-WCTIE-V.5-2017(23)-p.162-168

Semra Erpolat Tasabat[1], **Olgun Aydin**[2]
[1] Mimar Sinan Fine Art University. semra.erpolat@msgsu.edu.tr
[2]Mimar Sinan Fine Art University. olgun.aydin@ogr.msgsu.edu.tr

ABSTRACT

Communication is the way for people exchanging information with each other by using various tools. Electronic communication or E-communication is the process of sending, receiving and processing information or messages electronically. Electronic communication that is closely related to the development levels of countries, has made considerable progress especially in terms technology, innovation and entrepreneur. In this study, it is investigated whether electronic communication in Turkey is used effectively using the data published by Information and Communication Technologies Authority. For this purpose, different decision making techniques in data mining have been used. Obtained results are discussed.

Keywords: Multi criteria decision making, electre, topsis, electronic communication, data mining.

1. INTRODUCTION

Decision making, an inseparable part of life, is the process of choosing one of the available alternatives in order to achieve the specified goals and objectives. Life and working conditions that are rapidly changing and evolving at the same time are increasingly challenging people, institutions or businesses to consistently make "good" and "successful" decisions. Increasing the number of alternatives and criterions in complex decision problems, the contradiction of the criteria, made the selection process very difficult. Undoubtedly, the only way for institutions to gain superiority to their competitors in increasingly competitive conditions is to make quick and correct decisions.

Multi-Criteria Decision Making (MCDM), which is increasingly receiving more attention in recent years, incorporates approaches and methods that try to achieve a possible "best fit" solution that meets more than one criterion that conflicts with a decision situation.

One of the sectors where it is important to make the right decisions is electronic communications. With increasing technological progress and innovations, countries with an increasingly prevalent electronic communication are important elements to shape the future.

In the study, ELECTRE III and TOPSIS were taken up from MCDM methods to evaluate the activities of 81 provinces of Turkey in electronic communication sector. In the application of the mentioned methods, the weights of the determined criteria are determined according to the variable importance scores obtained in the neural network result.

2. LITERATURE REVIEW

From the literature view, it seems that in recent years, the efforts of electronic communication decision-making techniques have started to increase.

Feldman et. al have developed recommendations on when and how to use electronic communication to strengthen decision-making in their work, and explain various decision contexts (Feldman, et. al, 1996). Büyüközkan et. al, have developed a multi criteria group decision making Approach in smart phone selection using intuitive fuzzy TOPSIS (Büyüközkan, 2016).

Ertugrul et. al. have used the concept of fuzzy logic based on individual knowledge and experience in selecting the mobile line that provides the most economical, best and most appropriate solutions for the business needs in their work (Ertugrul et. al, 2014).

Petrovic et. al have offered a decision support tool for step-by-step benchmarking in their work. In a benchmarking exercise, a method is proposed for selecting a preferred development pathway comprising a series of intermediate test sequences(Petrovic et. al, 2014). Kaya et. al. have proposed an e-banking website quality assessment method based on an integrated fuzzy AHP-ELECTRE approach.

The fuzzy set theory has developed to deal with the problems arising from the uncertainty, uncertainty and the nature of human judgment (Kaya et. al ,2011). Chen et. al. investigated the relationship between the player and the struggling player in wireless communication (Chen et. al., 2010).

Asghari et. al. explored five different business models in the field of mobile payments. The MCDM has compared the evaluation method with its applications. Since one of the most effective parameters for selecting a suitable business model is the structure of banks / operators of each country, the proposed business model has been localized based on the framework of operators in Iranian banks (Asghari et. al, 201)

2.1. Multi-Criteria Decision Making Methods

It is possible for businesses to adapt to rapidly changing environmental conditions and make effective decisions in parallel with this change by using scientific methods that can evaluate a large number of qualitative and quantitative factors together in the decision-making process (Taha, 1997).

There are different methods used in the literature for the solution of MCDM problems, none of which provide a complete superiority over others (Dağdeviren, 2007).The most widely used methods of MCDM are ELECTRE, TOPSIS, Promethee, Vikor, Oreste, Mappac and Weighted Sum Approach (WSA). These methods have always been developed according to historical order in order to overcome the theoretical deficiencies of the previous method.

Because of being the most used methods in the literature, the activities in Turkey in the field of electronic communication in Turkey have been investigated by means of ELECTRE and TOPSIS methods in line with the criteria determined by Information and Communication Technologies Authortity (ICTA). In the following section methods are briefly mentioned.

2.1.1. ELECTRE Methods

ELECTRE (Elimination et Choix Traduisant La Realite - Elimination and Choice Translating Reality) is one of the MCDM methods originally developed by Benayoun, Roy et al. In 1966 (Roy 1968, 1991). The method incorporating a systematic analysis of the relationships between all possible pairs of different alternatives is based on the scores on each set of common criteria in each alternative. In the literature, there are several variants of the ELECTRE method in the form of ELECTRE I, II, III, Iv, IS, TRI. All methods are based on the same fundamental concepts, but are both operationally and according to the type of decision problem. Specifically, ELECTRE I is designed for selection problems, ELECTRE TRI for assignment problems and ELECTRE II, III and IV for ranking problems. ELECTRE II is an old version; ELECTRE III is used when it is possible to quantify the relative importance of the criterion and when the quantification is not possible (Hokkanen, J. and P. Salminen 1997). In this study, ELECTRE III method based on pseudo-criteria and fuzzy binary outranking relations is used.

It is possible to summarize the functioning of the method in the form of the decision matrix formation, the formation of the standard decision matrix, the formation of the weighted standard decision matrix, the determination of compliance and nonconformity sets, the formation of compliance and nonconformity matrices, the formation of matrices of superiority and nonconformity superiority matrices.

2.1.2. TOPSIS Method

TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) was developed by Yoon and Hwang as an alternative to the ELECTRE method in 1980 (Hwang, 1981).

The TOPSIS method consists of the generation of the decision matrix, the creation of the standard decision matrix, the creation of the weighted standard decision matrix, the creation of ideal and negative ideal solutions, the calculation of the separation measures, and the calculation of the ideal solution relative localization (Chen, 2000).

2.2. Variable Importance and Feature Selection

Data mining problems may involve hundreds, or even thousands, of fields that can potentially be used as predictors. As a result, a great deal of time and effort may be spent examining which fields or variables to include in the model. To narrow down the choices, the Feature Selection algorithm can be used to identify the fields that are most important for a given

analysis (Clementine Users Guide)

Typically, it can be considered to focus on the variables that give the most importance to modeling efforts and to at least ignore or ignore those that are insignificant. Variable precaution assists in making this variable by specifying the relative importance of each variable in model prediction. Since the values are relative, the sum of the values for all variables on the display is "1.0". Variable importance does not relate to model accuracy. It just relates to the importance of each variable in making a prediction, not whether or not the prediction is accurate. On the other hand "feature selection" ranks each predictor based on the strength of its relationship to the specified target, independent of other predictors. Thus feature selection will be more conservative in screening predictors. But in modeling, interactions and correlations are also taken into consideration (Clementine Users Guide). The importance levels obtained as a result of variable importance or feature selection can be used for analysis by considering them as variable weights.The weight determination process, which has an important place in the MCDM methods, is one of the most important steps of the decision analysis. Different techniques have been used in the literature to weight variables. In this study we used variable importance weights obtained from neural network analysis to weight the criteria.

2.3. Electronic Communication

Communication is needed for decision making, co-ordination, control, and planning. Communication, which is an integral part of our life, has become increasingly easier, faster and at the same time more understandable with developing technology. Communication using electronic media known as electronic communication. Worldwide communication has been facilitated by the electronic transmission of data which connects individuals, regardless of geographic location, almost instantly (http://www.meammarketing.com/advantages-and-disadvantages-of-electronic- communication/). One of the important research topics of today is how to manage the electronic communication, which is a demonstration of the level of development of the countries, better and more efficiently.

3. DATA AND METHODOLOGY

In the study, the variables used in the electronic communication sector to investigate which of the 81 provinces of Turkey are active in the electronic communication sector are discussed by Information Technologies and the publication of the year 2016 by the communication institution. The following table contains these variables and explanations.

Table 1: Explanations

Variable Code	Variable Name	Explanations
V1	Population	Data shows population statistics of Turkish Statistical Agency.
V2	Number of Fixed Telephony Access Lines	It includes number of active fixed analogue telephony access lines, voice channels equivalent of ISDN lines , fixed wireless subscriptions, number of payphones and number of VoIP subscriptions.
V3	Santral Capacity of Fixed Telephony	It shows total capacity of fixed telephone lines that can be handled. It shows number of xDSL broadband internet subscriptions.
V4	Number of Payphones	It shows number of active payphones.
V5	Number of Mobile Telephony Subscriptions – Total	It shows total number of mobile telephony subscriptions.
V6	Number of Mobile Telephony Subscriptions - 2G	It shows total number of 2G mobile telephony subscriptions.
V7	Number of Mobile Telephony Subscriptions - 3G*	It shows total number of 3G mobile telephony subscriptions.
V8	Number of Broadband Subscriptions – Total	It shows total number of broadband internet subscriptions.
V9	Number of Fixed Broadband Subscriptions – Total	It shows total number of fixed broadband internet subscriptions.
V10	Fiber	It shows number of fiber broadband internet subscriptions.
V11	xDSL	It shows number of xDSL broadband internet subscriptions.
V12	Cable	It shows number of Cable TV broadband internet

		subscriptions.
V13	Other	It shows number of broadband internet subscriptions via other means (Frame Relay, Metrı Ethernet, ATM, BPL).
V14	Number of Mobile Broadband Subscriptions **	It shows number of mobile broadband internet subscriptions via 3G or other appropriate mobile networks.
V15	Mobile Broadband Dedicated	It shows number of mobile broadband internet subscriptions via dedicated data cards.
V16	Standard Mobile Broadband	It shows number of mobile broadband internet subscriptions via mobile phones.
V17	Number of Cable TV Subscriptions	It shows number of cable TV subscriptions.
V18	The length of fiber***	It shows length of fiber rolled out for transmission and access.

* 3G Mobile telephone services have been provided since 2009.

** Mobile broadband ratio given here is updated. Currently, it includes the subscribers who have used packages less than one month period, who have used packages longer than one month period and who have accessed internet without any packages. Before, it was just consisted of the subscribers who had packages longer than one month period.

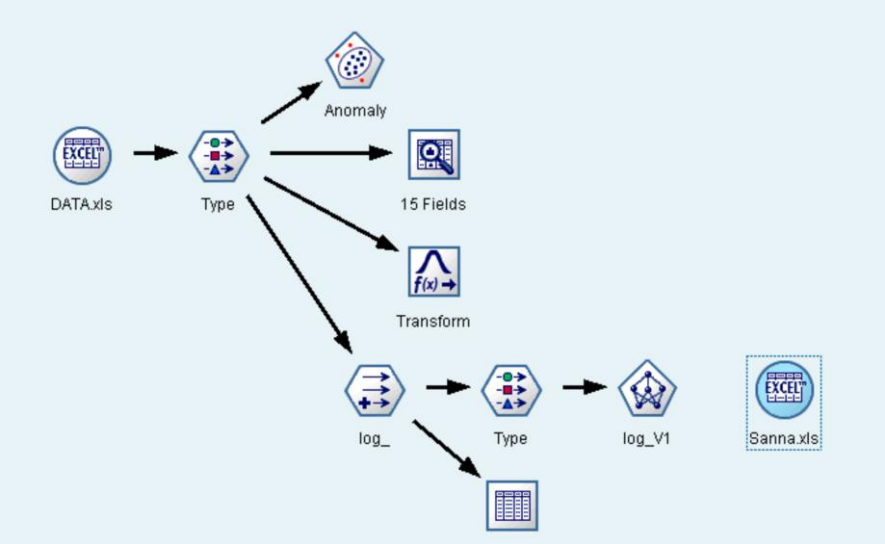
*** The length of fiber data about 2011, 2012, 2013, 2014 and 2015 include data obtained by alternative infrastructure operators.

In the study, Cities were taken as an alternative and the variables determined by ICTA as criteria. Variables V5, V9, V14 are not included in the analyzes since the sum of V5 variables V6 and V7, the sum of V9 variables V10, V11, V12 and V13, is the sum of V14 variables V15 and V16. Thus, a total of 15 criteria were used to measure the activities of the illicit people in electronic communication. The application part of the study consists of three phases; data analysis, variable weighting and efficiency study. The data mining program, called Clementine, was used in the data mining to perform data analysis and weighting. The SANNA plug-in of Excel is used to perform the event analysis.

In the data analysis phase, first, the existence of nudge, outliers and null values were investigated to check whether there were erroneous data entries. Following the necessary regulations, the distributions of the variables were examined and the normal test was performed. A logarithmic transformation is applied to the variables to provide a normality assumption.

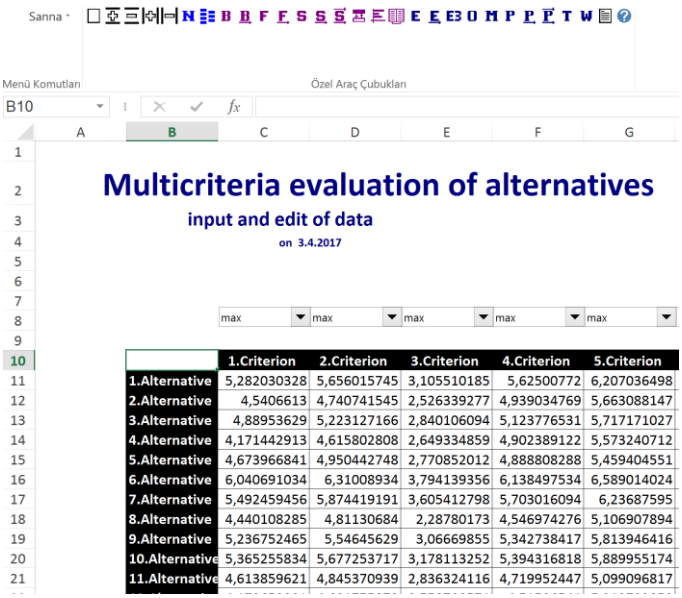
A neural network was used to determine the variable weights of the variables that were organized at the end of the data analysis phase during the variable weighting phase. The purpose of the neural network applied here is only to determine the weights. In the study, "population (V1)" is considered as variable target and other variables as input. Since all of the target and input variables are continuous, the variable importance analysis is based on the statistical based on the correlation coefficient. Figure 1 shows a screenshot of the data analysis and variable weighting steps.

Figure 1: Clementine Screen Shot



In the stage of efficiency analysis, the efficiency of 81 provinces was determined by means of ELECTRE III and TOPSIS methods to use the weights obtained during variable weighting. In Figure 2, there is a section from the screenshot of the data part of the program where the analysis is performed.

Figure 2: Sanna Plug-in Screen Shot

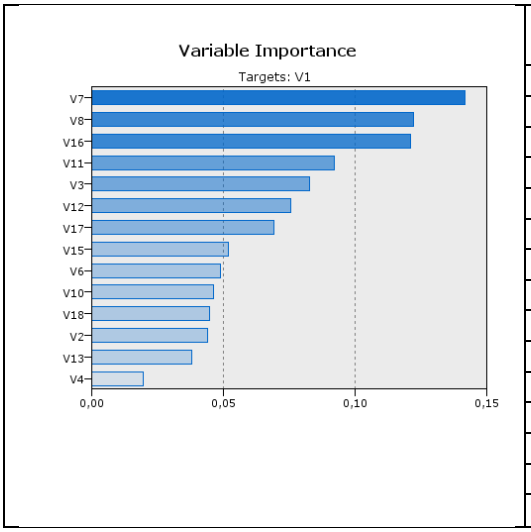


	1.Criterion	2.Criterion	3.Criterion	4.Criterion	5.Criterion
1.Alternative	5,282030328	5,656015745	3,105510185	5,62500772	6,207036498
2.Alternative	4,5406613	4,740741545	2,526339277	4,939034769	5,663088147
3.Alternative	4,88953629	5,223127166	2,840106094	5,123776531	5,717171027
4.Alternative	4,171442913	4,615802808	2,649334859	4,902389122	5,573240712
5.Alternative	4,673966841	4,950442748	2,770852012	4,888808288	5,459404551
6.Alternative	6,040691034	6,31008934	3,794139356	6,138497534	6,589014024
7.Alternative	5,492459456	5,874419191	3,605412798	5,703016094	6,23687595
8.Alternative	4,440108285	4,81130684	2,28780173	4,546974276	5,106907894
9.Alternative	5,236752465	5,54645629	3,06669855	5,342738417	5,813946416
10.Alternative	5,365255834	5,677253717	3,178113252	5,394316818	5,889955174
11.Alternative	4,613859621	4,845370939	2,836324116	4,719952447	5,099096817

4. FINDINGS AND DISCUSSIONS

The significance levels of the variables studied in the study of Turkey's effectiveness in the electronic communication sector are determined by means of the neural network model which gives the best estimation level with the Estimated Accuracy rate 99,539%. In the model in which the multiple method is used, there are two hidden layers containing 19 and 10 neurons, respectively. The variable importance values reached as a result of Neurol network analysis are as shown in Figure 3.

Figure 3: Results of Variable Importance



Variable Code	Weight
V7	0,1418
V8	0,1223
V16	0,1213
V11	0,0921
V3	0,083
V12	0,0756
V17	0,0693
V15	0,0521
V6	0,0491
V10	0,0463
V18	0,045
V2	0,0443
V13	0,0381
V4	0,0197
Total	1,000

According to the results of the analysis, the three most important criteria in electronic communication were determined as Number of Mobile Telephony Subscriptions - 3G (V7), Number of Broadband Subscriptions - Total (V8) and Standard Mobile Broadband (V16).

As a result of considering the variable importance levels as criterial weights, the results of the applied methods of ELECTRE III and TOPSIS are as follows respectively.

Table 2: Order of Alternatives According to ELECTRE III and TOPSIS Methods

RANK	ELECTRE III	TOPSIS	RANK	ELECTRE III	TOPSIS	RANK	ELECTRE III	TOPSIS
1	İSTANBUL	İSTANBUL	28	MALATYA	ŞANLIURFA	55	KİLİS	RİZE
2	ANKARA	ANKARA	29	KAHRAMANMARAŞ	TRABZON	56	NEVŞEHİR	YOZGAT
3	İZMİR	İZMİR	30	AFYONKARAHİSAR	SAKARYA	57	NİĞDE	AĞRI
4	BURSA	BURSA	31	SİVAS	KAHRAMANMARAŞ	58	AĞRI	NEVŞEHİR
5	ANTALYA	ANTALYA	32	KÜTAHYA	MALATYA	59	KIRIKKALE	ŞIRNAK
6	ADANA	ADANA	33	ELAZIĞ	AFYONKARAHİSAR	60	YALOVA	NİĞDE
7	KONYA	KONYA	34	ÇANAKKALE	VAN	61	BURDUR	KARABÜK
8	KOCAELİ	KOCAELİ	35	ÇORUM	ORDU	62	BATMAN	KIRIKKALE
9	GAZİANTEP	GAZİANTEP	36	IĞDIR	SİVAS	63	KARAMAN	BURDUR
10	İÇEL	İÇEL	37	ORDU	ÇANAKKALE	64	KARS	BİLECİK
11	MANİSA	MANİSA	38	VAN	DÜZCE	65	BİLECİK	SİİRT
12	KAYSERİ	DENİZLİ	39	MARDİN	KÜTAHYA	66	KIRŞEHİR	KARS
13	SAMSUN	KAYSERİ	40	GİRESUN	MARDİN	67	SİNOP	KIRŞEHİR
14	DENİZLİ	ESKİŞEHİR	41	ISPARTA	ELAZIĞ	68	BİTLİS	BARTIN
15	BALIKESİR	TEKİRDAĞ	42	AMASYA	ADIYAMAN	69	SİİRT	BİTLİS
16	ESKİŞEHİR	BALIKESİR	43	KASTAMONU	ÇORUM	70	ŞIRNAK	ARTVİN
17	TEKİRDAĞ	SAMSUN	44	DÜZCE	TOKAT	71	BİNGÖL	BİNGÖL
18	ERZURUM	ZONGULDAK	45	RİZE	AMASYA	72	ARTVİN	SİNOP
19	ZONGULDAK	ERZURUM	46	OSMANİYE	KIRKLARELİ	73	ÇANKIRI	MUŞ
20	EDİRNE	EDİRNE	47	YOZGAT	AKSARAY	74	MUŞ	ÇANKIRI
21	MUĞLA	YALOVA	48	UŞAK	ISPARTA	75	GÜMÜŞHANE	HAKKARİ
22	AYDIN	ERZİNCAN	49	BOLU	GİRESUN	76	HAKKARİ	GÜMÜŞHANE
23	SAKARYA	KARAMAN	50	ERZİNCAN	KASTAMONU	77	ARDAHAN	IĞDIR
24	HATAY	MUĞLA	51	BAYBURT	OSMANİYE	78	BARTIN	KİLİS
25	TRABZON	HATAY	52	TOKAT	BOLU	79	KARABÜK	ARDAHAN
26	DİYARBAKIR	DİYARBAKIR	53	KIRKLARELİ	BATMAN	80	AKSARAY	BAYBURT
27	ŞANLIURFA	AYDIN	54	ADIYAMAN	UŞAK	81	TUNCELİ	TUNCELİ

According to Table 2 ELECTRE III and TOPSIS methods, it is seen that the first 11 provinces which are most effective in electronic communication sector are the same. These are Istanbul, Ankara, Izmir, Bursa, Antalya, Adana, Konya, Kocaeli, Gaziantep, İçel and Manisa respectively. In both methods, it is seen that the province with the lowest value in the order of activity is Tunceli.

5. CONCLUSION

In the field of electronic communication, the data in Turkey provided by ICTA were utilized for ordering cities according to their activities in electronic communication.

In the study, the ELECTRE III and TOPSIS methods used in ordering from the MCDM techniques were utilized. In terms of weighting the criteria, a different approach has been taken to use the variable importance values obtained as a result of neural network analysis.

According to the results obtained, the criteria of 3G (V7), Standard Mobile Broadband (V13) and Number of Broadband Subscriptions - Total (V7) In the order of the results of the ELECTRE III and TOPSIS analyzes, it is seen that the first five and the last two orders are the same for both methods. According to this, the most effective illicit items in electronic communication were respectively Istanbul, Ankara, Izmir, Bursa, Antalya and the least efficient ones were Bayburt and Kilis.

REFERENCES

Chen, C. T. (2000). Extensions of the TOPSIS for Group Decision-Making under Fuzzy Environment. Fuzzy Sets and Systems, Vol: 114.

Clementine User Guide, https://web.kku.ac.th/wichuda/DMining/ClementineUsersGuide_11.1.pdf

Hokkanen, J. and P. Salminen (1997), ELECTRE III and IV decision aids in an environmental problem, Journal of Multi-criteria Decision Analysis, 6, 215-226.

Hwang, C. L. ve Yoon, K., Multiple Attribute Decision Making: A State of the Art Survey, Springer-Verlag, New York 1981.

Roy, B. (1968), Classement et choix en présence de critères multiples (la méthode ELECTRE), RIRO, 8, 57-75.

Roy, B. (1991), The outranking approach and the foundation of ELECTRE methods, Theory and Decision, 31, 49-73.

Roy, B., M. Présent and D. Silhol (1986), A programming method for determining which Paris Metro stations should be renovated, European Journal of Operational Research, 24, 318- 334.

http://www.meammarketing.com/advantages-and-disadvantages-of-electronic-communication/

Feldman, M. S.; Sarbaugh-Thompson, M. Information Infrastructure and Policy, v5 n1 p1-14 Jul 1996

Büyüközkan, Gülçin, and Sezin Güleryüz. "Multi Criteria Group Decision Making Approach for Smart Phone Selection Using Intuitionistic Fuzzy TOPSIS." International Journal of Computational Intelligence Systems 9.4 (2016): 709-725.

Ertugrul, Irfan, and Tayfun Oztas. "Business mobile-line selection in Turkey by using fuzzy TOPSIS, one of the multi-criteria decision methods." Procedia Computer Science 31 (2014): 40-47.

Petrović, Marijana, et al. "An ELECTRE-based decision aid tool for stepwise benchmarking: An application over EU Digital Agenda targets." Decision Support Systems 59 (2014): 230-241.

Kaya, Tolga, and Cengiz Kahraman. "A fuzzy approach to e-banking website quality assessment based on an integrated AHP-ELECTRE method." Technological and Economic Development of Economy 17.2 (2011): 313-334.

CHEN, Ya-ding, Shao-qian LI, and Yu-fan CHENG. "Effectiveness Evaluation of Integrated Anti-Jamming for Wireless Communication System." Journal of University of Electronic Science and Technology of China 39.3 (2010): 196-199.

Asghari, Fatemeh, et al. "A fuzzy ELECTRE approach for evaluating mobile payment business models." Management of e-Commerce and e-Government (ICMeCG), 2010 Fourth International Conference on. IEEE, 2010.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

REVIEW OF RISKS ASSOCIATED WITH RENOVATION & MODERNIZATION OF THERMAL POWER PLANTS IN INDIA

DOI: 10.17261/Pressacademia.2017.586
PAP-WCTIE-V.5-2017(24)-p.169-175

Purnima Bajpai[1], K. Chandrashekhar Iyer[2]
[1]The NorthCap University, Gurugram,Department of Civil Engineering, Indian Institute of Technology, Delhi. Purnima.bajpai.88@gmail.com
[2]Department of Civil Engineering, Indian Institute of Technology, Delhi. kciyer@civil.iitd.ac.in

ABSTRACT

Renovation and Modernization (R&M) refers to a set of activities intended to improve the performance and reliability of the existing production facilities, and is seen as a cost-effective alternative to increase the operational life of plants. Through R&M, obsolete generation equipment undergoes technological upgrades to boost capacity and reduce harmful emissions. It consists of a five-step process of identification, assessment, planning, execution and closure in order to plan, implement and monitor the entire procedure. The R&M cost per Megawatt (MW) varies, depending on the age of plant, operational history, fuel type, demand of the equipment to be replaced, the cost of generation and technology. The study aims at of Renovation and Modernization R&M of existing thermal power plants, which as an alternative can help the Country not only to increase the existing capacity and efficiency but will further help to reduce the heavy dependence on coal. Hence it is important to understand the major risks associated with the implementation of R & M projects and mitigate them proactively to enhance the effectiveness of this approach.

Keywords: Risks, effectiveness, efficiency ,renovation and modernization

1. INTRODUCTION

The main objective of Renovation & Modernization (R & M) of plants is to make the existing units well equipped with modified technology, equipment's and systems, and further with an aim to improve and enhance their performance and efficiency purely in terms of the output generated , its reliability, adherence to the original design values and reduction in maintenance cost. Planning and implementing Renovation and Modernisation (R&M) projects is often affected by occurrence of adverse events that can derail the objectives of the project. Identifying and mitigating project risks are crucial to successful management of R&M projects. The major thermal power plants set up during the late nineties are facing the problems of declining efficiency. A power plant is said to be inefficient if the existing inputs are not utilized in an optimum manner and as a result of which its generation becomes lower than its maximum possible generation. With the high capital expenditure on new capacities, poor financial health of the utilities and emerging fuel constraints it is essential to maximize generation from the existing power stations by restoring their rated capacity and reduce the scarce fuel more efficiently. This calls for Renovation & Modernisation (R&M) of existing old power plants, which is one of the most cost effective option to achieve additional generation from existing old units at low cost in short time period.

In India, the thumb-rule for adding one Mega Watt of fresh generating capacity entails a capital expenditure of around $1 million. However, an equivalent capacity can be achieved by investing almost one third of that amount on renovation and modernisation (R&M). The cost of R&M programmes are far less in comparison to new plant constructions and can be completed in a much shorter time compared to the gestation period of a new thermal power projects. R&M is an efficiency improvement tool and can improve power generation by 30 per cent, environmental impact by 47 per cent and efficiency by 23 per cent.

2. LITERATURE REVIEW

Based on the literature the key challenges impeding the market development mainly pertaining to the R& M division of power plants can be categorized into the following categories: Market challenges, Finance related challenges, Regulatory challenges, Operation and Maintenance (O & M) issues, Contractual challenges, Institutional challenges, Implementation challenges. The major challenges for successful implementation of R&M projects for different plants identified through the literature are as follows: Problems in defining the precise scope of work and defining the exact parameters that needs to be taken up, Uncertainty/deterioration in the phases between conducting the reduced life assessment studies as well as Condition assessment studies, award of the work and implementation of the projects, Limited number of agencies for ready to take up R&M projects.

R&M activity is to be fit together with the planned shutdown in order to ensure technology upgradation or extended life of the plant. On the other hand there are major encumbrances associated with the execution of the R & M works; there are various uncertainties because of the difficulty in the estimation of the scope and type of work which is needed until these generating units are opened up after rehabilitation. Taking the power sector as the basis, some of the study some major challenges faced by the thermal power renovation and modernization shall be discussed in detail.

2.1 Market Challenges

Even though there is significant R&M potential in India, the commercial opportunities within the market are limited. The actual achievement with respect to the planned R&M during the 10th and 11th Plan has been only a mere 9% and 17% respectively. The utilities in India usually follow two drastically different models (nomination and tendering) to select the suppliers for undertaking the requisite R&M works. The award of projects which is done generally on nomination basis to the existing O & M Contractors restrict the competition in the market and leads to a major barrier for various other suppliers. An analysis of the R&M market in India indicates that out of the total R&M capacity as per plan of 6501 MW which is almost equal to 31% (2030 MW) of the total capacity was awarded on nomination basis. It can also be seen from records that out of the total tendered capacity of 4471 MW, almost 47%, 2094 MW of the proposed tenders were dropped and not actually executed. There has been a tremendous increase in the global demand for power equipment's and supplies due to the increase demand of power and increase in installed capacity. It has been seen that at the global level most of the manufacturing facilities are booked for the next 16~18 months also to make matters worse in India due to lack of competition and management there is major reluctance to step into the so called naive market for R&M in India.

2.2 O&M Related Challenges

The kind of capital and time that is invested in the regular Operation & Maintenance activities of a power plant directly can be seen as a measure of the efficiency at which the plant is running. Most of the O& M practises followed by the State owned generation companies in India are weak and not upto the mark. Most of the state owned generating companies do not adhere to the schedule of periodic capital overhaul and annual maintenance leading to deterioration in the condition and performance of the plant. Poor O&M practises can reduce the expected efficiency levels of a successful R&M project before the specified extended life of the plant.

Poor O&M practices impacts the long term performance of plant and leads to its continual deterioration. It is important to have a long term proactively generated procedure for implementation of proper O & M Practices and procedures .There needs to be periodic review of the operating procedures from the beginning of the project and also identify the loop holes so as prepare an effective O&M manuals including preventive, capital and breakdown maintenance procedure / guidelines should be formulated.

2.3 Funding Related Challenges

Currently, the public sector through loans or grants is funding most of the R&M projects this is being done through the International Financial Institutions In general the financial demand is likely to be met by the normative structure i.e. 70% through the loan borrowing from commercial banks or other financial institutions and the remaining 30% through equity invested by the state. Through expert discussions it has also been highlighted that the non-availability of funds especially with the SEB'S to take up R&M works is one of the biggest challenges in the Country. As per RBI data the power sector alone utilizes approximately 9.23% of the Gross bank credit as per data published by RBI on 31st March 2013. Another major challenge with the R & M projects lies in the poor financial health of the State utilities resulting in the limited debt servicing ability of the utilities. Due to lack of proper planning the available finance is consumed in fire fighting and inappropriately planned schemes rather than being utilized in the useful schemes with long payback and short-term costs, such as R&M projects. When the R&M is financed through the Govt schemes generally the lowest capital expenditure option is given first

priority. The Capital Expenditure for R&M needs to be managed in such a way as to reduce the short-term increase in power cost from the generating units.

2.4 Regulatory Challenges

The existing regulatory framework offers limited incentives to the generating companies against improving the efficiency of the generating unit and for successful implementation of R&M projects. The benefits obtained on improving the efficiency of power plants were to be fully transferred to the consumer during the tariff revision for that plant based on the Annual tariff setting procedure of the Government of India. Another major challenge lies in the fact that an upfront commitment for the capital costs required and the possible plant performance may be difficult because of the inadequate methodologies adopted for the RLA studies, which cases there intermediate issues and hampers the pace of work. There needs to be a proper analysis comparing the financial cost based tariffs as against the economic pricing of additional power that can be made available by taking up R & M works.The question still remains that who is the deciding Authority for such issues, the lack of a system makes it even more difficult for investment decisions and decision making procedure becomes extremely slow and non-reliable.

2.5 Contractual Challenges

The way the Contract is drafted plays a significant role in the kind of involvement of different stakeholders in a project; it also can have an encouraging or a discouraging impact on the players. It has been observed the risk sharing mechanism is missing in the general Contracts which are being followed by the Generation Companies. During discussions and expert sessions with vendors and consultants, many issues with regards to the Contract were highlighted .The foremost important issue highlighted was the weakly defined scope and many open ended statements in the commercial contract which deliberately shift the risk towards the bidders/Contractors.

It normally takes about 3-4 years between the technical studies (RLA studies, DPR preparation etc) and the commencement of actual R&M work. This clubbed with inadequate/incomplete technical information provided to the bidders prior to bid restricts the executing party to realistically, model and predict the condition of the existing equipment's before they are opened and inspected which ultimately increase the risk profile of the project.

2.6 Institutional Related Challenges

Limited training of utility professionals in the area of planning and execution of R&M projects, absence of dedicated cell/department at the company level, deployment of the best personnel in the field of new generation capacity and frequent transfers are some of the reasons which have contributed to the limited skills and expertise of the generating company to plan and implement R&M projects. Interactions with the various stakeholders including suppliers have revealed that after the projects are awarded, the entire risk and responsibility for completion of the project is passed on to the suppliers with limited support by utilities during the implementation process.

One of the major implementation challenges faced by the utilities is that they are unable to schedule timely shutdown for executing R&M due to grid conditions. Significant energy and peak deficit scenario in most of the states coupled with lack of planning with regard to procurement of power from other sources inhibits shutdown of state owned units for executing R&M works. In certain cases this is driven by socio-political consideration that results in delay in obtaining the shutdown.

2.7 Entry Barriers

The area of R & M is still on the developing stages and at times the barriers provided at the entry level can have a major influence on the development of the potential market. There need to be appropriate levels of barriers and restrictions as these are essential to protect the profitability of existing supplier base as free entry and exit of firm in Industry would affect the profit of firms to a minimal levels and hence, making the industry unattractive. The possible mitigation measures for the above mentioned risks based on the review can be summarized as mentioned in the table below.

Table 1: Summary of the Possible Mitigation Measures

S No	Risk Factor	Possible Mitigation Methods
1.	Supply Risks (Bassanini,2011)	Supply undertaking (Sponsors/Govt. agency), Collaterals (Until reaches upto proven level), Reserves prove-up/Assurance, Reserves depletion protection (Accelerated repayment), Reserve weighting
2.	Market Risks (Smith1999)	Take or pay contracts, Advanced Sales contract, Buy Back Contracts, Throughput agreements, Minimum Quantity Contracts, Market preference Contract, Contract monetisation, Merchant financing
3.	Foreign Exchange Risk (Lifson 1988)	Forward Contracts, Parallel loans, Barter, Currency Swaps, Commodity

		lending, Commodity lending
4.		
5.	Environmental Risks (Schaufelberger 2005)	Environmental Management, Rehabilitation Management, Emergency Response, Environmental Warranty, Environmental Insurances, Rehabilitation guarantees, Pollution Control Boards
6.	Infrastructure Risks (Benjamin2004)	Infrastructure Contracts, Government Contracts, Pooled Infrastructure, Transport studies, Free on Board offtakes
7.	Political Risks (Tiong 1995,1996)	Development agreement, Insurance (Export credit agencies)(against currency inconvertibility, creeping nationalisation, etc), Tax indemnification (by sponsors), Offshore payment agent (EOUs), Currency inconvertibility agreements, Local national participation (Local equity/ debt), Co-financing (National/Super-national Bodies)
8.	Participant Risks (Dailami,2003)	Cross Collateralization, Deficiency Agreements, Share Pledge, Board Control
9.	Completion Risks (Macmillan,2000)	Performance bond, Maintenance bond, Completion undertaking, Overrun standby facilities

3. DATA AND METHODOLOGY

The demand and supply gap of power in the country is increasing day by day and despite the numerous steps taken by the Government this gap is continuously increasing. A developing Country like India which heavily depends on the thermal power plants for its generation of power needs to definitely think of alternatives to increase the generation and reduce this gap. Even though many steps have been taken by the Government to move towards the alternative sources of energy, the major reliance is still on the coal industry

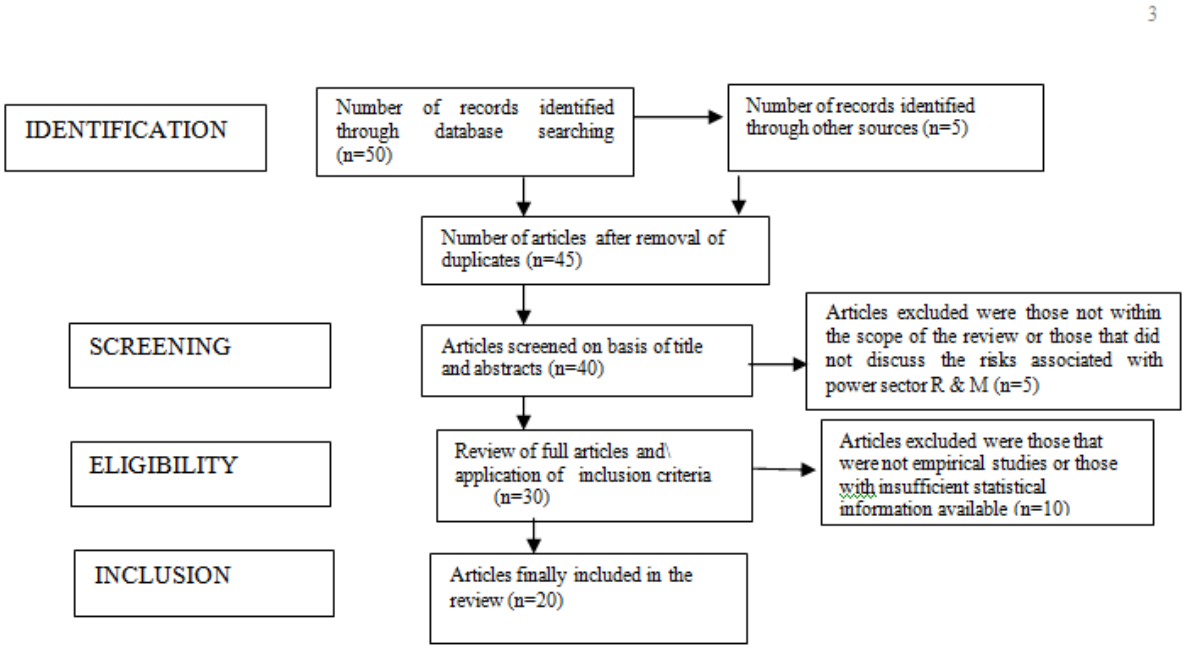
Various studies carried out in the past have highlighted one common concern of the poor and deteriorating efficiencies of the thermal power plants in India. Keeping in mind the various constraints with respect to the green field projects like that of land availability, legal and political concerns and majorly the lack of funds, the only alternative is to look at the other ways of improving efficiencies from the prevailing one. Hence this has generated a need to look into the various risks associated with the implementation of R & M projects in India and how these can be effectively mitigated to increase the efficiency of the thermal power plants in India.

A systematic literature review has been used which attempts to gather all factors that will help to answer the specific research question. It uses systematic methods that are selected with a view to minimizing bias, thus providing reliable findings from which conclusions can be drawn and decisions made.The key characteristics of a systematic review are: (a) a clearly stated set of objectives with an explicit, reproducible methodology; (b) a systematic search that attempts to identify all studies that would meet the eligibility criteria; (c) an assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias; and (d) systematic presentation, and synthesis, of the characteristics and findings of the included studies.

The PRISMA approach for reviewing literature has been followed (figure 1 below). This technique uses a detailed literature review approach to identify and analyze the factors and risks effecting the implementation of Renovation and Modernization projects in India using the databases like google scholar, and web of science, reputed journals and conference proceedings etc. The inclusion criteria's considered were peer-reviewed, empirical, original articles, research and review papers.

In this approach a total of 50 articles have been considered. 5 articles were removed from the 50 initial considered articles. Further the remaining 45 articles were studied and after further screening on the basis of the abstracts and titles 40 articles were taken up for further 5 articles were excluded which were not within the scope of the review or those that did not discuss the risks associated with power sector R & M. Finally 30 articles were taken up for full review out of which 10 were excluded as these articles were those that were with insufficient statistical information available The remaining 20 articles were then assessed and inclusion criteria applied and were included in this review

Figure 1: Work Flow Process for PRISMA Technique



4. FINDINGS AND DISCUSSIONS

Based on the extensive literature review as stated above the following risks have been identified as the key challenges to R& M of the thermal power plants in India.

Table 2: Summary of Risks identified

S. NO	RISK DESCRIPTION
1.	**Market related Risks**
1.1	Award of Contract through nomination and not competitive bidding
1.2	Non Existence of market for technical consultants
1.3	Non availability of schedule for sale of power generated.
1.6	Limited ability of utilities to infuse equity investment in R & M projects
1.7	Lack of awareness of possible market options
1.8	Credit exposure to power sector is likely to reach its limit set by most commercial banks
1.9	Lower vendor participation because of limited competition which leads to higher procurement cost.
1.10	Rebidding/Reward/Delays in award of R & M packages/Contract
1.11	Mismatch (delay) in supply of critical equipment and the shutdown period
2.	**Regulatory related Risks**
2.1	Lack of appropriate incentives sharing mechanism for Gencos
2.2	No bench marks set for R & M works
2.3	Lack of Government support and incentives
3.	**Funding Risk**
3.1	Non availability of funds with utilities for regular O & M procedures
3.2	Non availability of funds with utilities to take up R & M projects
3.3	Credit limit for power sector reaching limits with commercial banks
3.4	Focus only on technical criteria with limited focus on financial & economic concerns
3.5	Lack of confidence of financers on R & M due to limited success stories in the Country.
4.	**Planning Risks**
4.1	Inadequate assessment for R & M Works and scope of work

4.2	Poorly defined objectives
4.3	Lack of past operating & performance data with the utility which makes it difficult to identify the actual improvement areas.
4.4	Incomplete studies carried out and the studies carried out majorly focus on only the major components of the plant
4.5	Lack of energy Audits
5	**Contractual Risks**
5.1	Unfair and misbalanced risk and reward system between the utilities and contractors
5.2	Delays in bid evaluations and award of work
5.3	Inappropriate packaging strategy for works
5.4	Inappropriate contractual conditions for delay in works and changes in scope of work
5.5	Improper contractual conditions for the execution of R & M works.
5.6	Poor dispute resolution mechanism
6.	**Management Risks**
6.1	Limited project appraisal skills
6.2	Inadequate exposure of the utility professionals in the area of planning & execution of R & M
6.3	Adequate personnel not dedicated to the R & M Activity
6.4	People working in R & M Projects get transferred to other departments in middle of the project.
6.5	Lack of Authority amongst the officials involved.
6.6	Lack of decision making by the utility
6.7	Lack of long term generation plan and awareness of available market options
6.8	Reactive approach for identification of plants for taking up R & M
7.	**Institutional Risks**
7.1	Limited capacity and capability of state utilities in undertaking R & M projects
7.2	Poor implementation support from utilities while executing R & M projects
7.3	Delay in obtaining unit shutdown for undertaking technical studies

REFERENCES

1. Ahluwalia, Montek Singh (2011), "Prospects and Policy Challenges in the Twelfth Plan", www.planningcommission.nic.in/aboutus/speech/spemsa/spe_21052011.pdf.
2. Bagchi, Amaresh (2007), "Role of Planning and the Planing Commission in the New Indian Economy; Case for a Review", Economic & Political Weekly, Vol.42, No. 44 (Nov. 3-9, 2007), pp. 92-100.
3. Batra, R K and S K Chand (2011), "India's coal reserves are vastly overstated: is anyone listening" The Energy and Resources Institute, New Delhi.
4. Banker, R.D., A. Charnes and W. W. Cooper, 1984, Some models for estimating technical and scale efficiencies in data envelopment analysis. Management Science 30:1078–1092.
5. Barros, C.P. and N.Peypoch, 2008, Technical efficiency of thermoelectric power plants. Energy Economics 30: 3118-3127.
6. CEA, 2005. Technical Standard on Operation Norms for Coal/Lignite Fired Thermal Power Stations.Central Electricity Authority, Government of India.
7. CERC, 2000a. Availability Based Tariff, Petition No.2/1999. Central Electricity Regulatory Commission.
8. CERC, 2000b. Operational Norms for Thermal Generation, Petition No. 4/2000, Central Electricity Regulatory Commission. ⟨http://⟩cercind.gov.in/2612/Order%20Final.pdf
9. CERC, 2003. Operational norms for thermal generation, Petition No. 4/2000, Central Electricity Regulatory Commission. ⟨http:// www.cercind.gov.in/280203/4-00.pdf⟩.
10. Chikkatur, Ananth P (2008), "Coal Initiative Reports: A Resource and Technology Assessment of Coal Utilization in India", Pew Center on Global Climate Change.
11. Chikkatur, Ananth P. et al, (2009), "Sustainable development of the Indian coal sector", Energy 34 (2009), pp. 942-953.
12. CEA, 2013,Study Report on developing Markets for Implementation of R&M Schemes in Thermal Power Stations in India, India: Coal Fired Generation Rehabilitation Project
13. Chand, S.K., Sarkar, S.K., 2006. Demystifying Coal Security. Economic Times, 14 April.
14. Chikkatur, A., 2005. Making the best use of India's coal resources. Economic and Political Weekly 40, 5457–5461.

15. Central ElectricityAuthority,2006. All India Electricity Statistics. General Review, 2004–2005. Central Electricity Authority, Ministry of Power, Government of India. Available online /http://www.cea.nic.in/S.
16. Central ElectricityAuthority,2008a .17th Electric Power Survey,2007.Central Electricity Authority, Ministry of Power, Government of India.
17. Central Electricity Authority, 2008b. Thermal Performance Review,2007–2008. Central Electricity Authority, Ministry of Power, Government of India .Avail-able online /http://www.cea.nic.in/S.
18. Central ElectricityAuthority,2009a .All India Electricity Statistics. General Review, 2009. Central Electricity Authority, Ministry of Power, Government of India.
19. Dadwal, Shebonti Ray, (2012) "India's Overseas Assets: do they contribute to energy security", Strategic Analysis, vol. 36, issue. 1
20. Deo Sharma, S.C., 2004. Coal-fired power plant heat rate and efficiency improvement in India. In: Workshop on Near-Term Options to Reduce CO2 Emissions from the Electric Power Generation Sector in APEC Economies, Asia Pacific Economic Cooperation (APEC), Queensland, Australia.
21. Dubash, N.K. (Ed.), 2002. Power Politics: Equity and Environment in Electricity Reforms. World Resources Institute, Washington, DC.
22. Dubash, N.K., Rajan, S.C., 2001. Power politics: process of power sector reforms in India.Economic and Political Weekly 36, 3367–3390.
23. Dubash, N.K., Singh, D., 2005. Alternating currents: introduction to an international review of electricity restructuring. Economic and Political Weekly 40, 5242–5248.
24. Farrell, M.J., 1957, The measurement of productive efficiency. Journal of royal statistical society 120:7-24.
25. Halder, Dulai (2012), "Overseas investment in Upstream oil and gas: rational, policy, current situation and assessment of strategy", conference presentation, ONGC Videsh Limited, 23May 2012.
26. H. Robinson, P. Carrillo, C.J. Anumba and M. Patel, Governance & Knowledge Management for Public-Private Partnerships, Wiley-Blackwell, 2010
27. Jagadeesh, A. (2000), "Wind energy development in Tamil Nadu and Andhra Pradesh, India Institutional dynamics and barriers – A case study", Energy Policy 28 (2000), pp. 157-168.
28. Joseph, Kelli (2010), "The politics of power: electricity reform in India", Energy Policy 38 (2010), pp. 503-511.
29. J.A. Walewski, International project risk assessment, Dissertation for the degree of Doctor of Philosophy, The University of Texas at Austin, 2005.
30. Kalam, Abdul (2011), former President of India, in address during 21st SMIRT Conference, www.dae.gov.in/press/pr071111.pdf.
31. Kumar, Shukla, Umesh (2011) & Ashok Thampy, "Analysis of competition and market power in the wholesale electricity market in India," Energy Policy 39 (2011), pp. 2699-2710.
32. L.Y. Shen, A. Platten and X.P. Deng, Role of public private partnerships to manage risks in public sector projects in Hong Kong, International Journal of Project Management, 24(7), (2006), 587-594.
33. Madan, Tanvi (2006). The Brookings Foreign Policy Studies Energy Security Series: India, Washington, DC.
34. Mathew, Jose (2003), "Development Policies of India with Special Reference to New Economic Policy", Mahatma Gandhi University, Kerala.
35. Mathur, Neerja (2010) "Overview of Indian power sector" presentation at Joint IEA-India Workshop on Industrial Energy Efficiency, New Delhi, 27 January 2010.
36. Mathur, Neerja (2011) "Indian Power Sector – An Overview" presentation at Joint IEA-India Workshop on High Efficiency, Low Emission Coal Technology, New Delhi, 29 November 2011.
37. M.H. Sobhiyah, M.R. Bemanian and Y.K. Kashtiban, Increasing VFM in PPP power station projects- Case study: Rudeshur gas turbine power station, International Journal of Project Management, 27(5), (2007), 512-521
38. Pachauri, P K (2005), "Securing India's Energy Future", The Indian Express, 4 August.
39. R.J. Sawant, Infrastructure Investing: Managing risks & rewards for pensions, insurance companies & endowments, John Wiley & Sons, Inc, 2010.
40. Tavanir, Energy Planning Office Report, Energy Balance Sheets Between1996–2006, Available from Power Ministry library and information centre, 2007.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

DETERMINATION OF ZIP COEFFICIENTS FOR RESIDENTIAL LOADS

DOI: 10.17261/Pressacademia.2017.587
PAP-WCTIE-V.5-2017(25)-p.176-180

Mehmet Bircan [1], Ali Durusu [2], Bedri Kekezoglu [3], Onur Elma [4], Ugur Savas Selamogullari [5]
[1]Yildiz Technical University, Electrical Engineering Department, mbircanee@gmail.com
[2]Yildiz Technical University, Electrical Engineering Department, adurusu@yildiz.edu.tr
[3]Yildiz Technical University, Electrical Engineering Department, bkekez@yildiz.edu.tr
[4]Yildiz Technical University, Electrical Engineering Department, onurelma@yildiz.edu.tr
[5]Yildiz Technical University, Electrical Engineering Department, selam@yildiz.edu.tr

ABSTRACT

In this study, constant-impedance, constant-current and constant-power ZIP models have been analyzed for residental loads. Measurement based approach has been used to obtain ZIP models of selected loads. The measurement-based approach for load modeling is important because it reflects the real dynamics of the system and represents the load characteristics more accurately. The ZIP model is used to estimate the power drawn by the load depending on the voltage changes. Since the ZIP load model is a well-known model that provides power dependence in terms of the physical sense of voltage, studies are taking place in this area. In this study, the voltage and power values of selected residental loads at YTU Smart Home Laboratory are measured. Then, the ZIP coefficients have been calculated by using the least squares algorithm developed in MATLAB. The measured data is compared with the data from the obtained ZIP model for each appliance.

Keywords: Load models, ZIP model, residential appliances, least squares method.

1. INTRODUCTION

Load modeling is an important issue in power system analysis and it is an essential part of a static and dynamic analysis of power system. Load models are mathematical expressions defining the relationship between active and reactive power and power system voltage and frequency changes. Load models can be considered as the mathematichal representation of relationship between power and voltage. There are mainly two types of load models: static models and dynamic models. Static load models are generally not time-dependent (B. Zhao,2010). Most common static load models are exponential and polynomial load models. Also, a static load model can be defined as a composite ZIP model, which is a well-known model that represents the relationship between applied voltage and power consumption. Dynamic load models can be examined in two parts as input-output load model and physical load model. The most common dynamic load model is the induction motor model (K. Rudion, 2009).

Load characteristics have a significant effect on system performance and it is known that the results of load flow, voltage and transient stability simulations are dependent on the assumed load model (M. Sadeghi, 2009). Thus, the right models for all power system components including loads should be developed for power system analysis. Measurement based and component based approaches are two main approaches used to obtain the load models. Both produce appropriate polynomial or exponential static load models for load flow or stability studies where the load voltage and system frequency do not vary significantly with time (B. Zhao,2010). The polynomial model is commonly named as the ZIP model since the model consists of constant impedance (Z), constant current (I), constant power (P) components. The model parameters are the coefficients defining the proportion of each component (A. Bokhari,2014). Thus, each load has a unique coefficent set that describes the load.

The aim of this work is to obtain the ZIP models of 4 residential appliances at YTU Smart Home Labarotary. For thsi purpose, meassurements are completed to collect data on the power consumption versus applied voltage change. Then, ZIP model

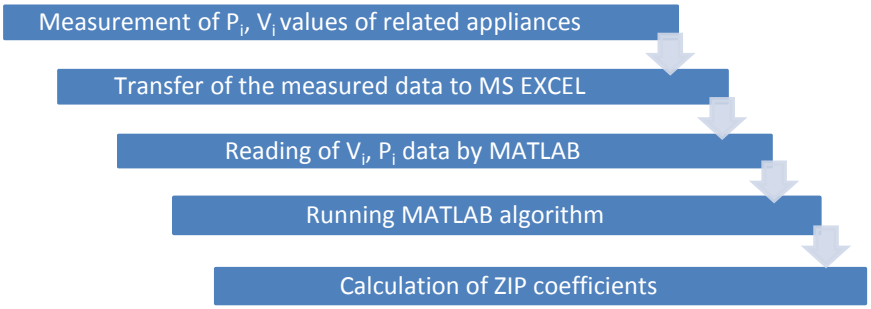
coefficients are calculated by using the least square algorithm developed in MATLAB. As a result of the study, calculated active power values of examined appliances are compared with actual measurements.

The rest of the paper is organized as follows: Section II presents the mathematical background of ZIP model calculation, Section III describes the evaluated case studies and related results. Section IV finalize the paper with concluding remarks.

2. DETERMINATION OF ZIP COEFFICIENTS

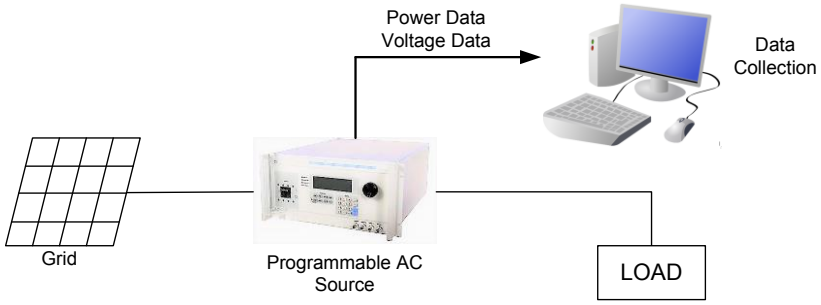
Power system loads can represent different characteristics. Accordingly, several different load models could be obtained using mainly three approaches: constant impedance, constant current, or constant power. However, power system loads cannot be always accurately modeled by applying each mentioned approaches individually. Therefore, ZIP models that include all three approaches are used to reach most accurate load models. ZIP models are extensilvely used in power system analyses. In this study, ZIP model coefficients of selected residential appliances in a Smart Home Labaratory at YTU Department of Electrical Engineering are obtained by using experimental results. The steps to determine ZIP coefficients are shown in Fig 1.

Figure 1: The Steps to determine ZIP Coefficients



A programmable AC source is used to obtained voltage-dependent load power measurements. The measurement system is shown in Fig 2. Collected data for each appliance is stored in an excel file and then is processed with MATLAB to calculte ZIP coefficients.

Figure 2: Measurement System



In experiments, the active power consumption of the tested appliance versus 10V increments of the applied voltage is collected. The starting value of the applied voltage is 100V and the ending value of the applied voltage is 240V. Using a MATLAB based least squares fit algorithm and measured data, ZIP coefficients of tested appliance are found. Mathematical background of developed algorithm is presented below.

2.1. ZIP Model

The ZIP model is a polynomial expression which describes the relationship between the applied voltage and constant impedance, constant current, and constant power components (K. Rudion, 2009). The ZIP model also reveals the quadratic, linear and constant change of the power curve of an element according to voltage (F. L. Quilumba, 2011). Active power and reactive power expressions of a ZIP model are given in Eq. 1 (K. Rudion, 2009):

$$P = P_0 \left[Z_p \left(\frac{V}{V_0} \right)^2 + I_p \left(\frac{V}{V_0} \right) + P_p \right] \quad (1)$$

$$Q = Q_0 \left[Z_q \left(\frac{V}{V_0} \right)^2 + I_q \left(\frac{V}{V_0} \right) + P_q \right]$$

Here, P and Q are active and reactive power corresponding to the operating voltage (V), P_0 and Q_0, are active and reactive power corresponding to nominal voltage (V_0); Z_p, I_p and P_p are ZIP coefficients of active power component; Z_q, I_q and P_q are ZIP coefficients of reactive power component.

2.2. Least Squares Method

Diffirent algorithms can be used to obtain ZIP coefficients based on measured voltage-power values. In this study, Least Squares Method (LSM) is employed due to its simplicity. The objective function is shown in Eq. 2.

$$\lambda = \sum_{i=1}^{n} \left(Z_p \left(\frac{V_i}{V_0} \right)^2 + I_p \left(\frac{V_i}{V_0} \right) + P_p - \frac{P_i}{P_0} \right) \quad (2)$$

Here, V_i and P_i are measured values of voltage and active power, respectively. To solve Least Squares problem, partial derivatives of objective function for each variable should be equal to zero. Thus, three equations are obtained for three ZIP coefficients. The solution is achieved by solving these three equations. The solution matrix is given in Eq. 3 (M. Sadeghi,2009).

$$\begin{bmatrix} \sum_{i=1}^{N} V_i^4 & \sum_{i=1}^{N} V_i^3 & \sum_{i=1}^{N} V_i^2 \\ \sum_{i=1}^{N} V_i^3 & \sum_{i=1}^{N} V_i^2 & \sum_{i=1}^{N} V_i \\ \sum_{i=1}^{N} V_i^2 & \sum_{i=1}^{N} V_i & n \end{bmatrix} \begin{bmatrix} Z_P \\ I_P \\ P_P \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} P_i V_i^2 \\ \sum_{i=1}^{N} P_i V_i \\ \sum_{i=1}^{N} P_i \end{bmatrix} \quad (3)$$

3. FINDINGS AND DISCUSSIONS

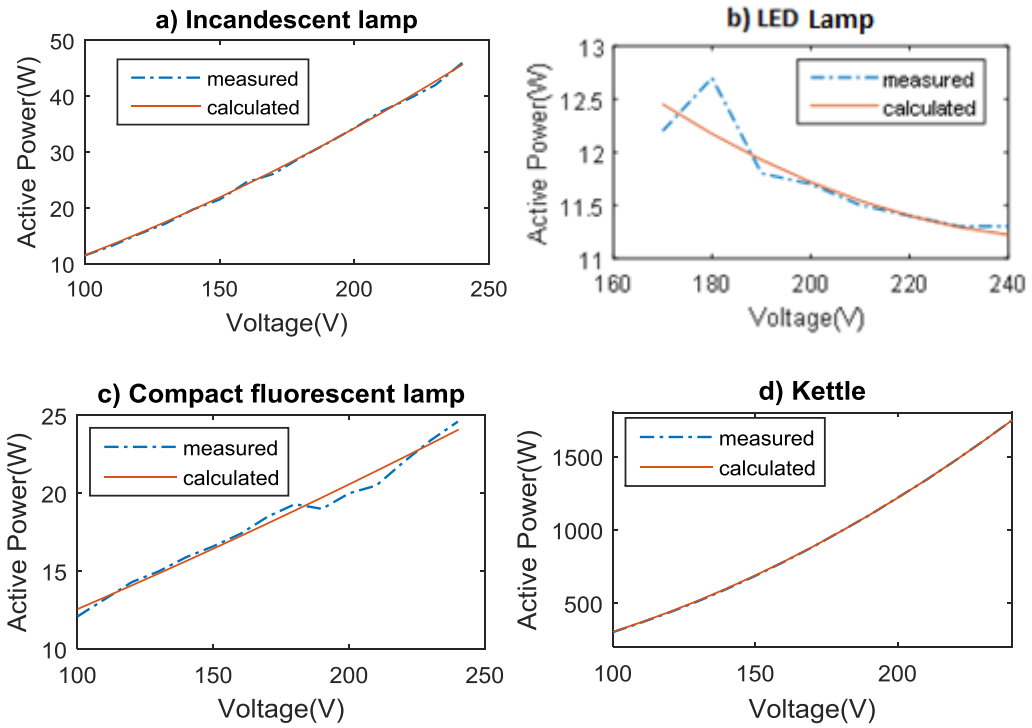
Three different lighting equipment, incandescent lamp, compact fluorescent lamp and LED lamp, and a kettle are selected as residential appliances and active power versus 10V voltage increments between 100V and 240V are measured and stored within the scope of this study. Then, measured data is processed for each appliance using the LSM algorithm developed in MATLAB. Calculated ZIP coefficients of each appliance are summarized in Table 1.

Table 1: ZIP Coefficients for Each Appliance

Appliance	Z	I	P	Z+I+P
Incandescent lamp	0.49	0.59	-0.08	1
LED lamp	0.73	-1.7	1.97	1
Compact fluorescent lamp	0.11	0.64	0.25	1
Kettle	0.95	0.07	-0.02	1

As seen, constant impedance model is convenient for the kettle while it is more appropriate to use ZIP model for other appliances. Comparision of the actual measurements and obtained ZIP model ouputs are given in Fig. 3.

Figure 3: Comparison of ZIP Model Outputs with Measured Values



Calculated and measured values of incandescent lamp are shown in Fig. 3a. Nominal voltage of lamp is 220 V and the active power at nominal voltage is 39.8 W. By examining Figure 3a, it is seen that calculated values are very close to the measured values and active power characteristic of the incandescent lamp can be correctly modeled with the ZIP model.

Power changes of LED lamp is given in Fig. 3b. Nominal voltage of LED lamp is 220V and the active power at nominal voltage is 11.4W. Fig. 3b shows that calculated values are very close to the measured values at above 200V voltage level. For the practical purposes, the obtained ZIP model for the LED lamp can be a valid model considering that the nominal voltage of appliances in Turkey is around 200-220V level. However, it is seen that the difference between the measured and calculated power values is rising at low voltage levels.

Calculated and measured values of compact fluorescent lamp are compared in Fig. 3c. Nominal voltage of the lamp is 220V and the active power at nominal voltage is 22.3W. It is seen that the measured and calculated values are very close.

Power changes of the kettle is given in Fig. 3d. Nominal voltage of the kettle is 220V and the active power at nominal voltage is 1470W. It has been observed that the measured and calculated curves for the kettle overlapping.

4. CONCLUSIONS

Load models have significant importance for accurate power system analysis. Power system loads can be modeled in different ways. In this study, ZIP coefficients of residential loads at YTU Smart Home Labarotory are obtained by using actual measurement data. Primarily, power measurements are performed at the operating voltage ranges of considered loads, Obtained data is then processed with help of a solution algorithm developed in MATLAB and ZIP coefficients of the selected appliances are calculated. Examination of ZIP coefficient results shows that the kettle is suitable for constant impedance model, however it is observed that ZIP model is more appropriate for the other appliances. When the ZIP model output and measured data are compared for each appliance, it is seen that calculated and measured values of incandescent lamp and compact fluorescent lamp are very close. The power characteristic for the LED lamp is accurate at voltages above 200V, however deviations are observed at low voltages. It is seen that the measured and calculated power curves overlap for the kettle.

REFERENCES

B. Zhao, Y. Tang, W. Zhang, and Q. Wang, “Modeling of common load components in power system based on dynamic simulation experiments,” in Proc. Int. Conf. Power Syst. Technol., 2010, pp. 1–7.

K. Rudion, H. Abildgard, Z.A. Styczynski ''Non-linear load modeling — Requirements and preparation for measurement,'' Power & Energy Society General Meeting, 2009, pp. 1–5.

M. Sadeghi and G. A. Sarvi, “Determination of ZIP parameters with least squares optimization method,” in Elect. Power Energy Conf., 2009, pp. 1–6.

A. Bokhari, A. Alkan, R. Dogan, M. D.-Aguiló, F. León, D. Czarkowski, L. Birenbaum, A. Noel, R. E. Uosef, '' Experimental Determination of the ZIP Coefficients for Modern Residential, Commercial, and Industrial Loads'', IEEE Transactıons On Power Delıvery, Vol. 29, No. 3, June 2014.

F. L. Quilumba, W. Lee, H. Huang, D. Y. Wang, and R. L. Szabados, “Load model development for next generation appliances,” in Proc. Ind. Appl. Soc. Annu. Meeting, 2011, pp. 1–7.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

INFRASTRUCTURE WITH R PACKAGE FOR ANOMALY DETECTION IN REAL TIME BIG LOG DATA

DOI: 10.17261/Pressacademia.2017.588
PAP-WCTIE-V.5-2017(26)-p.181-189

Zirije Hasani
Faculty of Conputer Science, University "Ukshin Hoti" Prizren, Kosovo. zirije.hasani@uni-prizren.com

ABSTRACT

Analyzing and detecting anomalies in huge amount of data are a big challenge. On one hand we are faced with the problem of storing a large amount of data, on the other to process it and detect anomalies in reasonable or even real time. Real time analytics can be defined as the capacity to use all available enterprise data and sources in the moment they arrive or happen in the system. In this paper, we present an infrastructure that we have implemented in order to analyze data from big log files in real time. Also we present algorithms that are used for anomaly detection in big data. The algorithms are implemented in R language. The main components of the infrastructure are Redis, Logstash, Elasticsearch, elastic-R client and Kibana. We explore implementation of several filters in order to post-process the log information and produce various statistics that suit our needs in analyzing log files containing SQL queries from a big national system in education. The post-processing of the SQL queries is mainly focused on preparing the log information in adequate format and information extraction. The other interesting part of the paper is to compare the anomaly detection algorithms and to conclude which of them is better to us for our needs. Also we add the elastic-R client to the infrastructure we develop for big data analytic in order to detect anomalies. The purpose of the analysis is to monitor performance and detect anomalies in order to prevent possible problems in real time.

Keywords: Big data, anomaly detection elgorithm, log data, logstash, elasticsearch, elastic-R client, kibana.

1. INTRODUCTION

With the increased number of internet users, the need for analyzing data and specifically log data is increasing too. The two general requirements of big data projects are common: analysis of the (near) real time information extracted from a continuous inflow of data and persisting analysis of a massive volume of data. Log management is complex and time consuming process, even harder when we have to deal with big log files that came in real time. Log file is a file that records all the events that happened during one software or/and operating system is running. Also, it may register all the exchange of personal messages between different users under some communication software. The content of log files could be diverse, e.g. it could be structured, semi-structured and weakly structured. Our special interest is log files that contain SQL queries.

Building an infrastructure for analyzing big log files in real time is a computational, storage and scalability challenge. To make proper choice of infrastructure we have done extensive investigation reported in [7], [8], [9] and [10]. In this paper, we present adjusted infrastructure proposed by Ian Delahorne [2]. Among open source and free software tools, we find it appropriate because it is possible to modify it when needed, by adding various other components (like Hadoop), or scale up or down by adding (duplicate, triplicate,…) some of existing components.

The main components of the infrastructure are: Redis, Logstash, Elasticsearch, Kibana and elastic-R client.

Redis is used for temporary buffering of the log data, Logstash utilizes different filters to manipulate and analyze the data, Elasticsearch is used for indexing and storing the data, elastic-R client is used for anomaly detection and Kibana is a user interface used to visualize the results. If some of the parameters rise above expected values, the elastic R- client will visualize it in order to know where the anomaly happen.

Study and experiments are motivated by need to use such an infrastructure for analyze of the log files from a system called e-Dnevnik (ednevnik.edu.mk[1]) and detecting anomalies in such system. e-Dnevnik is an electronic system for managing the data records of students in Macedonian schools. System enables daily communication between teachers, parents and students and various statistical analyzes used by Ministry of education and research of RM and other public institutions.

System receives a big number of requests during a day and analyze of these requests is required before they are saved to database in order to reduce the amount of logs that is necessary to be saved. The idea is to save into database just the information that is of interest for future processing and other to be ignored. Even more, analyze of log files in real time can signalize and detect errors, track CPU usage, monitor parameters and similar. The main part of the paper is to show the robust algorithms for anomaly detection in streaming data and how we implement them for our needs. We have examined different algorithm used for anomaly detection in real time big data as: MAD[13], runMAD[16], DoubleMAD[15], DBSCAN[13], moving average, Statistical Control Chart Techniques (aggregation, difference) [20], etc. Will be compared the algorithm based on the execution time because for real time big data execution time is very important.

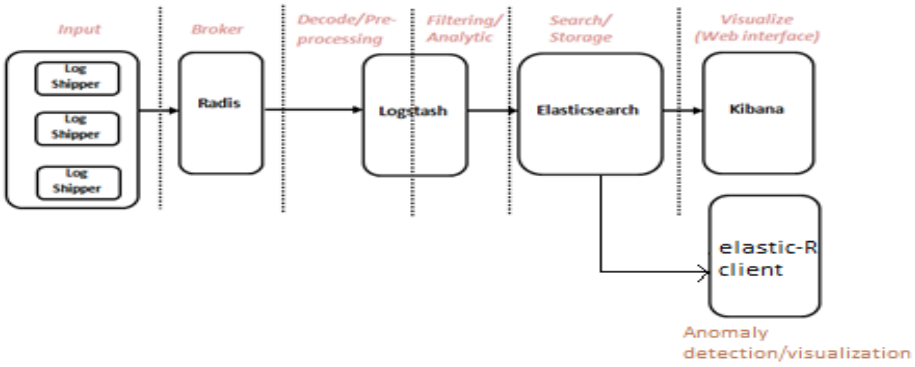
Paper is structured as follows: In the second chapter we explain which are the components and functions of this infrastructure; in the third chapter, we show the robust algorithms for anomaly detection in real time big data; in the fourth chapter, are compared this algorithm in order to chose the best one for our needs; in the fifth chapter, we demonstrate pre-processing of the SQL queries contained in log files and usage of several Logstash filters important for real time analytics; in the six chapter we implement the infrastructure for real time analysis of e-Dnevnik database log file; the last chapter is the conclusion for our work done thus far, including also ideas for future work.

2. INFRASTRUCTURE FOR ANALYZİNG LOG DATA IN REAL TIME

With aim to deal with big log files in real time produced by PostgreSQL server in order to analyze query performance, we start with the solution proposed by Ian Delahorne [2]. Since Elasticsearch [5], together with Logstash has evolved during the past several years, we include in our architecture several new and remove several unnecessary components. Our proposed architectural is shown in Figure 1. This architectural design is based on pipeline event processing, divided in following phases: input (collects and manages events and logs), buffering, decode/pre-process (extract structured data into variables, parse), filter (modify, extract information), anomaly detection and output (ship the data for storage, index, search and visualize).

An important characteristic of this architecture is the capability to scale up/out of every component, depending on the input stream size and rate, by running one or more of its components as separate threads/servers. For example, we can scale out the input phase with three shipper servers as shown in Figure 1, or we can scale up the Logstash filtering server on a bigger machine with more CPU cores/RAM.

Figure 1: Components of Infrastructure for Big Data Log Analytic and Anomaly Detection in Real Time



Flexibility is achieved also by possibility of adding various additional components as Hadoop, Cassandra, statistical or graphical tools like Statsd, Graphite and others. Generally in most cases when we run the Logstash server there will be two broad classes of Logstash host [3]:

- The first one is the host which runs the Logstash agent as an event "shipper" that sends application, service and host logs to a central Logstash server.
- The second one is central Logstash host which runs a combination of components of this architecture for pre-processing and filtering of events.

[1] http://ednevnik.edu.mk/

Broker (usually Redis [12]) acts as efficient temporary buffer for logs. This especially is important to enable interruptions in the processing of the log events in an occasion of upgrade process of the Logstash instances, or in the case of an unexpected raise of event size and number.

The main component of the infrastructure is Logstash [4]. It is written in JRuby and runs in a Java Virtual Machine (JVM) [4]. It is easy to deploy, as a single JAR file that can be started directly using a JAVA SE VM (no Apache Tomcat Containers are needed). Its architecture is simple comparing with other similar software architectures since it consists of a three phase pipeline (input, filter, output) and it provides an easy way of extension of functionalities in each phase using plugins.

Input phase collects the logs and sends the collected events to the filter phase. Logs can generally arrive from various sources: Files, TCP/UDP files, Syslog, Microsoft Windows EventLogs, STDIN, Key-value stores and a variety of others. In our case log file includes Postgres SQL CSV log files and Key-value stores (Redis [12]).

Logstash comprise a large collection of filters which enable us to extract structured data into variables, parse, modify and enrich the data, before they are pushed to the Elasticsearch.

Elasticsearch enables efficient indexing and storing of the event logs, and enables a full text search on them. It is an open-source distributed search engine library, built on top of Apache Lucene [9]. ElasticSearch [6] allows us to implement store, index and search functionality and as such help us in easier and more efficient computation of various data analytics. ElasticSearch is a NoSQL data store where data are stored as documents. Although it is mainly used by Java applications, the important thing is that applications need not to be written in Java in order to work with ElasticSearch, since it can send and receive data over HTTP in JSON to index, search, and manage our Elasticsearch cluster. Kibana[6] which is a HTML/JS frontend web interface to Elasticsearch for viewing the log data. The beauty of Kibana is that we can easily search in the data with different queries, produce charts, histograms and other visual products. Elastic-R client[14] gives access to local or remote Elasticsearch database. The anomaly detection algorithm are implemented in R package.

3. ROBUST ALGORITHMS FOR ANOMALY DETECTION

In order to specify which algorithm will be used for anomaly detection we have made a survey of existing algorithms. There are many algorithms for anomaly detection for different problems, in our case we focus on algorithms for contextual collective anomalies. The idea was to explore robust algorithms for anomaly detection and we chose some of them like MAD[13], runMAD[16], DoubleMAD[15], DBSCAN[13], sliding window.

The Median Absolute Deviation (MAD) algorithm is a robust measurement of variability, and can be viewed as the robust analogue for standard deviation [13]. This algorithm work as follow:

$$MAD = consistency.constant * median\ (abs(x - median * x))$$

The method works well for symmetric distribution of data. DoubleMAD is used if we have courved tail. MAD is better than standard deviation technique (SD) because in MAD will not take into calculation values which deviate a lot from normal data, this is not the same in SD for this reason MAD give better result than SD. For example if we have the streaming data (5,6,3,90) MAD will not take into calculation number 90, and this don't affect the result.

MAD is a good solution for not streaming data but in our case we have to analyze streaming data for this reason we use runMAD technique which use sliding window to find anomalies.

RunMAD [2] (Median Absolute Deviation of Moving Windows) for streaming data is the median of the absolute deviations from the data's median for defined window. As such does not make any distribution assumptions. Similar window functions are runmin, runmax, runmed, runquartile, etc. Depending on the stringency of the researcher's criteria, which should be defined and justified by the researcher, the author [17] proposes the values of k=3 (very conservative), k=2.5 (moderately conservative) or even k=2 (poorly conservative) for anomalies detection that are outside Median ± k*MAD.

DBSCAN algorithm is a density-based clustering algorithm. It works by greedily agglomerating points that are close to each other. Outliers are considered clusters with few points in them [13]. This algorithm has two main parts: a parameter ε that specifies a distance threshold under which two points are considered to be close; and the minimum number of points that have to be within a point's ε-radius before that point can start agglomerating.

Statistical control chart technique [20] is graph used to study how a process changes over time and control of repetitive processes. In general, the chart has a central line that represents the mean value for the in-control process and other two lines, upper control limit and the lower control limit. These control limits are chosen so that almost all the data points will

[2] http://svitsrv25.epfl.ch/R-doc/library/caTools/html/runmad.html

fall within these limits as long as the process remains in-control. Data could be chart of individual data, aggregated by time parameter (e.g. hour), moving range, moving average and others.

4. COMPARISON BETWEEN ROBUST ALGORITHMS FOR ANOMALY DETECTION

In order to define which algorithm is better for anomaly detection in streaming data we have to compare some of them. In the table below we made comparison between MAD, runMAD, DoubleMAD, DBSCAN, sliding window, HTM, Twitter ADVEC and Etsy-skyline based on some characteristics.

Table 1: Comparison of Anomaly Detection Algorithms

Characteristic	MAD	runMAD	DoubleMAD	DBSCAN	Sliding window	HTM	TWITTER ADVEC	Etsy skyline
Big data	yes	yes	yes	yes	yes	yes	yes	yes
Streaming data		yes			yes	yes	yes	yes
Symmetric distribution of data	yes				yes			
Mowing window		yes			yes	yes	yes	yes
Non symmetric distribution of data			yes		yes			
Contextual anomaly	yes	yes	yes		yes			
Collective anomaly				yes	yes			
Univariate data	yes	yes	yes	yes	yes			
Multivariate data						yes	yes	yes
Outlier detection	yes	yes	yes	yes	yes	yes	yes	yes

From the comparison we can conclude that the best algorithm for streaming data is runMAD because the data are computed in moving window in the time where they came.

For many anomalies both algorithms like MAD and DBSCAN perform well, but in some cases one differ from the other. If we have several streams coming from several servers and may one of them delay this will not be showed by MAD algorithm, in this cases it is better to use DBSCAN.

4.1. Algorithm Comparison Based on Execution Time

The challenge that face big data especially real time big data is anomaly detection, for this reason in our research we face with this problem and propose a infrastructure that will help to find anomalies in real time. The infrastructure is proposed in [19] but here we have added one more component elastic-R where are implemented the anomaly detection algorithms.

To realize the experiments we have used the log data from e-dnevnik for two working days 03.04.2015 (Friday) and 06.04.2015 (Monday). In the table below we have shown the result from execution time of different algorithms for different SQL queries in different dates. The attributes of data are timestamp and duration of query.

Table 3: Comparison of Algorithms for Real Time Big Data Anomaly Detection

	03.04.2015		06.04.2015	
	ee8eda021a8956da d5d9c208ee6a1aa7 (Q1) *72798 records*	66e46548f8a85602 98c537d34c468b3d (Q2) *2992 records*	ee8eda021a8956da d5d9c208ee6a1aa7 (Q1) *65070 records*	66e46548f8a85602 98c537d34c468b3d (Q2) *2110 records*
Algorithm	Execution time	Execution time	Execution time	Execution time
runMAD	4.987 secs	0.759 secs	4.439 secs	0.469 secs
DBSCAN	error	5.015 secs	error	2.006 secs
Twitter ADVec			30.306 secs	0.366 secs
Difference between data	1.691 secs	0.219 secs	1.648 secs	0.188 secs
Moving average	1.965 secs	0.272 secs	1.689 secs	0.263 secs

From the table 3 we can conclude that DBSCAN and Twitter ADVec algorithms are the slowest algorithm. The faster algorithm is statistical controll chart algorithm (difference between data in moving window).

5. PROCESSING OF SQL QUERIES

Processing of database transaction logs presents a big challenge due to their massive volume. The main target of the SQL queries analytics is to gather information and detects anomalies in query performance on an operational level. This means that we want an early detection of performance degradation of SQL queries in real time and alert adequately in order to remove the possible causes.

5.1. Log Data Pre-processing

In order to get more realistic results we must do a SQL queries pre-processing by performing a normalization procedure on them. The normalization of the SQL queries tries to remove all data and parameters from the queries in order to gather better grouping/clustering of SQL query types. These includes elimination of comments, start of transactions, string

content, null parameters, non essential numbers and hexadecimal numbers, the last line of code, removing of extra space, new line and tab characters and lower-casing. Similar normalization process can be referred in pgBadger [1] that is used for batch log file processing. Next is an example of Logstash configuration file for normalization of SQL queries using Logstash mutate filter. This filter allows performing of regular expression pattern matching and replacement for general transformation of event fields. Following is the piece of the Logstash filter configuration file for SQL query normalization:

```
mutate {
 # Set the entire query lowercase
 lowercase => [ "statement" ]
 gsub => [
  # Remove comments
  "statement", "\/\/\*(.*?)\*\/", "",
  # Remove extra space, new line and tab
  "statement", "[\t\s\r\n]+", " ",
  # Remove start of transaction
  "^\s*begin\s*;\s*/,"",
  # Remove string content
  "statement" , "\\'", "",
  "statement" , "'[^']*'", "''",
  "statement" , "''('')+", "''",
  # Remove NULL parameters
  "statement" , "=\s*null", "='",
  # Remove numbers
  "statement" , "([^a-z_\$-])-?([0-9]+)","\ 10",
  # Remove hexadecimal numbers
  "statement" , "([^a-z_\$-])0x[0-9a-f]{1,10}", "\10x"
 ]}
```

Other useful pre-processing plugging is the merge filter that lets us combine two events that occur within a period into a new single event. This can be helpful if information for a single SQL query is split into several log events. In our case, Postgres logs two events for a single query, first containing the SQL query, and the second containing the duration of the query execution. Merge plugin has the following options that are used:

- key => Unique identifier, used to match the two events you want to merge.
- order => 'first' or 'last', the order the events should arrive
- merge_tag => Tag(s) to add on the *new* event.
- period => Max length of time between events(seconds).

In the example below if the event is the first event to be merged we execute the following merge plugin. This can be controlled using conditional filter processing. The merging of events is based on the key values, i.e. in this case "session_id" and "session_line_num".

```
merge {
 key => [ "session_id", "session_line_num" ]
 order => 1
 period => 1
}
```

Finally for the second event, if that event contains the durraion of the SQL query, and matches the key fields "session_id" and "session_line_num", the event fields specified are merged. In this case we merge only the "duration" field.

```
merge {
key => [ "session_id", "session_line_num" ]
fields_to_merge => [ "duration" ]
 order => 2
}
```

At the end of pre-processing we remove the log message key-value for personal data protection, and further calculate a hash of the normalized SQL statement in order to optimize the analysis process so it will not involve the full complex SQL statements.

```
mutate {
 add_field => [ "sql_hash", "%{statement}" ]
 remove_field => [ "message" ]
}
anonymize {
 algorithm => "MD5"
 fields => [ "sql_hash" ]
 key => "<some seed>"}
```

5.1.1. Analytics Filter

In order to perform statistical analysis of the performance of SQL queries, we found that the use of the metrics filter [3] can be practical. The metric filter produces an aggregation metrics from the log events based on the selected key values. The metrics filter is invoked periodically (flush_interval), can filter the processed events based on a time frame (clear_interval) and can produce statistics of both event occurrence (count, rate of events) and event values (ex. sql statement duration). The timer parameter of the metrics filter gives us a variety of information as follows:

- "thing.count" - the total count of events
- "thing.rate_Xm" - the X-minute rate of events
- "thing.min" - the minimum value seen for this metric
- "thing.max" - the maximum value seen for this metric
- "thing.stddev" - the standard deviation for this metric
- "thing.mean" - the mean for this metric
- "thing.pXX" - the XXth percentile for this metric

Following is the Logstash configuration that uses the metrics filter in order to produce statistics on every 60 seconds, based on SQL events in the past 300 seconds. The statistics contain count, rate_1m and rate_5m for the event occurrence, and duration statistics per SQL query type (sql_hash). The statistics are produced as a separate log event.

```
metrics {
 add_tag => "metric"
 timer => [ "%{sql_hash}", "%{duration}" ]
 flush_interval => 60
 clear_interval => 300
 rates => [1,5]
}
```

Analyze of log files in real time can signalize and detect errors, track CPU usage, monitor parameters and similar. If some of parameters rise above expected values, or error occurs, built in (in future) triggers will indicate or even prevent possible problems in real time. Figure 2 shows how these metrics filters present results in Kibana.

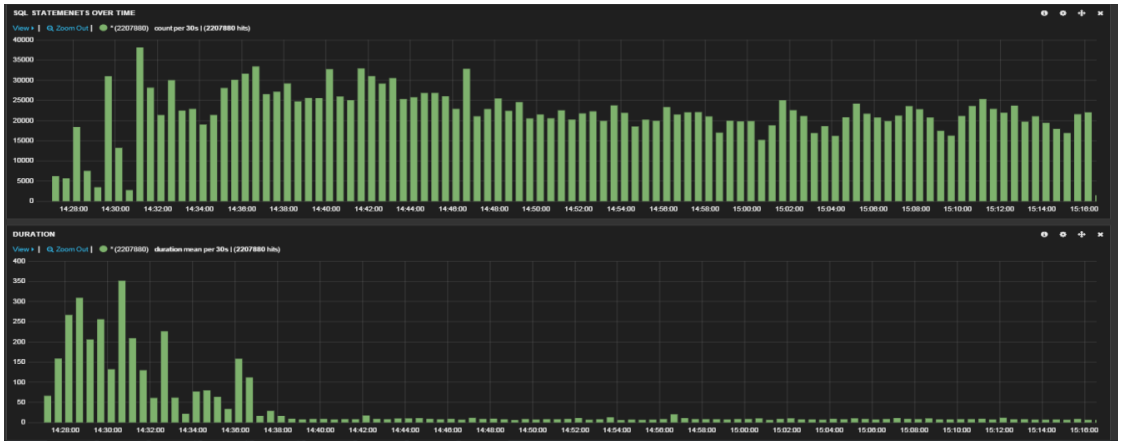
Figure 2: Result Fields of Metrics (meter and timer) Event in Kibana

0091e3390f763d542da76ea18d56717d.count	832
0091e3390f763d542da76ea18d56717d.rate_1m	4.50225627
0091e3390f763d542da76ea18d56717d.rate_5m	4.21929311
0091e3390f763d542da76ea18d56717d.min	0.008
0091e3390f763d542da76ea18d56717d.max	0.283
0091e3390f763d542da76ea18d56717d.mean	0.0620601
0091e3390f763d542da76ea18d56717d.stddev	0.20703094
0091e3390f763d542da76ea18d56717d.p1	0.009
0091e3390f763d542da76ea18d56717d.p5	0.009
0091e3390f763d542da76ea18d56717d.p10	0.01
0091e3390f763d542da76ea18d56717d.p90	0.113
0091e3390f763d542da76ea18d56717d.p95	0.124
0091e3390f763d542da76ea18d56717d.p99	0.15367
0091e3390f763d542da76ea18d56717d.p100	0.283

6. REAL TIME ANALYSIS OF E-DNEVNIK DATABASE LOG FILE

To illustrate possibilities of our infrastructure, we have analyzed log files generated from e-Dnevnik. e-Dnevnik is the electronic system for managing the student records of elementary and high schools in Macedonia. There is a huge number of requests in real time and we would like to take some statistics based on the traffic that is generated in a defined periods of time. The data analyzed are SQL queries saved in log file. In the Figure 3 below we present two histograms produced by Kibana. The first chart displays the distribution of the number of events in the system, calculated per 30 seconds intervals in the time period from 14:26 untill 15:16, having 2207880 hits all. The second chart shows the calculated mean duration of SQL queries execution time for the same period and intervals. This shows that the mean of the query duration is higher at the specific period of time. The higher mean duration time of SQL queries in this example is the consequence of the Postgres server restart and warming up of Postgres shared buffers.

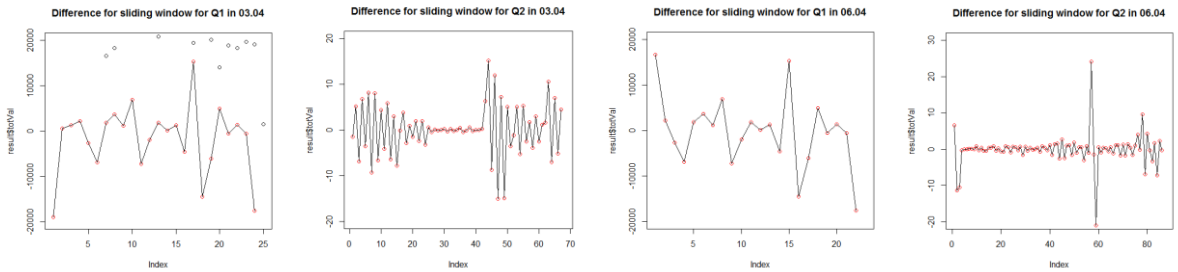
Figure 3: e-Dnevnik Number of Hits and Duration Mean per 30 Seconds Intervals, in Selected 14:26-15:16 Period of Time



6.1. Result Visualization with R

To show visually how our implemented/tested algorithm which is shown from the above experiments as faster algorithm for real time anomaly detection, we add in the infrastructure R package because the algorithm are implemented in R language. The figure 4 show the result visualization, we have visualize the result of difference algorithm and visually can see when the difference between points is larger it may be one anomaly. Object of study was 2 type of SQL query based on query duration and timestamp for two working days 03.04 and 06.04.

Figure 3: Statistical Control Chart (difference) Q1 and Q2 for 03.04 and 06.04



7. CONCLUSION

Analyzing Big Data in real time is a challenging process but the need for this analytics is emerging with enormous growth of incoming data and need of their fast analyze. The other challenge is detecting anomalies in real time big data. In this paper, we propose infrastructure we have adjusted in order to analyze big log files in real time, algorithms used for anomaly detection and demonstrate related analytics we made on system “e-Dnevnik” big log files that are produced daily by its PostgreSQL server.

The main components of the infrastructure are open source and free software tools, Redis, Logstash, Elasticsearch, elastic-R client and Kibana. The infrastructure design is based on the pipeline event processing, divided in phases: input (collects and manages events and logs), buffering, decode/pre-process (extract structured data into variables, parse), filter (modify,

extract information), anomaly detection and output (ship the data for storage, index, search and visualize). Proposed architecture is capable to scale up/out depending on the input stream size and rate, by running one or more of its components as separate threads/servers. Flexibility is achieved by possibility of adding various further components as Hadoop, Cassandra, statistical or graphical tools like Statsd, Graphite, or deploying extension of functionalities in each phase by using own plugins.

We illustrate the SQL queries database transaction logs analytics with implementation of the filters that produce various statistics enabling detections of anomalies in query performance on an operational level. This means that we are able to detect performance degradation of SQL queries in real time and alert adequately in order to remove the possible causes. In the same time in real time we do the pre-processing of the logs in order to reduce the amount of content of SQL queries that are necessary to be saved for further analyze. Also are elaborated robust algorithm used for anomaly detection and tested them by R language, the result is shown with plots. The testing phase with R for now is done offline but in the future we plane all the process to be online by integrating the elastic-R in our proposed architecture.

We made different comparisons in this paper for algorithms that are used for anomaly detection and in the future we plane to modify them in order to adapt for online use because now they are aimed just for offline use.

We plan to extend the pre-processing of the incoming logs by parameterization of the SQL queries to lower further the volume of the stored data and to enable easier future analyses. Depending on the input stream of data we will experiment with scale up/out of the system components/servers and including other (batch appropriate) components as R software.

REFERENCES

pgBadger. Retrieved April 04, 2015, from http://sourceforge.net/projects/pgbadger/.

Ian Delahorne. Postgresql Metrics With Logstash. Retrieved April 04, 2015, from http://ian.delahorne.com/blog/2014/06/10/postgresql-metrics-pipeline

Logstash. Retrieved April 05, 2015, from http://logstash.net/docs/1.4.2/filters/metrics.

James Turnbull. *The Logstash Book Log management made easy*. January 26, 2014.

Radu Gheorghe and Matthew Lee Hinman. *Elasticsearch in action*. Manning Publications 2014.

Mitchell Anicas. How To Use Logstash and Kibana To Centralize Logs On Ubuntu 14.04. Retrieved April 06, 2015, from https://www.digitalocean.com/community/tutorials/how-to-use-logstash-and-kibana-to-centralize-and-visualize-logs-on-ubuntu-14-04.

Zirije Hasani, Margita Kon-Popovska, Goran Velinov. Survey of Technologies for Real Time Big Data Streams Analytic. 11th International Conference on Informatics and Information Technologies. April 11-13, 2014 – Bitola, Macedonia.

Zirije Hasani, Margita Kon-Popovska, Goran Velinov. Lambda Architecture for Real Time Big Data Analytic. ICT Innovations 2014 Web Proceedings ISSN 1857-7288

Zirije Hasani. Performance comparison throw running job in Hadoop by defining the number of maps and reduces. 12th International Conference on Informatics and Information Technologies 2015. April 24-26, 2015 – Bitola, Macedonia.

Zirije Hasani. Virtuoso, System for Saving Semantic Data. 12th International Conference on Informatics and Information Technologies 2015. April 24-26, 2015 – Bitola, Macedonia

Apache Lucena. Retrieved April 30, 2015, from https://lucene.apache.org/.

Redis. Retrieved April 30, 2015, from http://redis.io/.

DBSCAN. Retrieved December 20, 2016, from https://cran.r-project.org/web/packages/dbscan/dbscan.pdf

elastic r client. Retrieved November 20 2016, from http://finzi.psych.upenn.edu/library/elastic/html/elastic.html

doubleMAD algorithm. Retrieved November 10 2016, from http://eurekastatistics.com/using-the-median-absolute-deviation-to-find-outliers/

runMAD algorithm. Retrieved November 10 2016, from http://svitsrv25.epfl.ch/R-doc/library/caTools/html/runmad.html

Christophe Leys, Christophe Ley, Olivier Klein, Philippe Bernard and Laurent Licata. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology. Elsevier Inc. 2013.

Miller, J. (1991). Reaction time analysis with outlier exclusion: Bias varies with sample size. The Quarterly Journal of Experimental Psychology, 43(4), 907–912, http://dx.doi.org/10.1080/14640749108400962.

Zirije Hasani, Boro Jakimovski, Margita Kon-Popovska and Goran Velinov. *Real time analytic of SQL queries based on log analytic.* ICT Innovations 2015

Mark Kasunic, James McCurley, Dennis Goldenson and David Zubrow. An Investigation of Techniques for Detecting Data Anomalies in Earned Value Management Data. Carnegie Mellon University. December 2011 https://pdfs.semanticscholar.org/b998/7cd7e7244b1235a21c72c5a6f6634a9ff430.pdf.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

BLACKOUT AND BLACKSTART ON POWER SYSTEMS

DOI: 10.17261/Pressacademia.2017.589
PAP-WCTIE-V.5-2017(27)-p.190-197

Cihan Ayhanci[1], Beyazit Yasar Yoldas[2], Bedri Kekezoglu[3]
[1] Yildiz Technical University, Department of Electrical Engineering, Istanbul, TURKEY, cihanayhanci20@gmail.com
[2]Turkish Electricity Transmission Company, Istanbul, TURKEY, yasarbeyaz@gmail.com
[3] Yildiz Technical University, Department of Electrical Engineering, Istanbul, TURKEY, bkekez@yildiz.edu.tr

ABSTRACT

Power systems are complex structures that energy production, transmission and consumption units. There are various reasons for energy interruptions on these systems. Power outage on the power system can lead to partial or complete loss of the network. Energy interruptions in large regions cause significant economic losses and precautions must be taken in this regard. In this study, Blackout situations in power systems were analyzed. Within the scope of the study, important Blackout situations encountered around the world have been examined. Blackout causes are listed and interpreted in power systems in the light of experienced events. Precautions have been put forward in order to avoid blackout situations. In addition, the effects of blackstart-equipped power plants on the recovery of the network have been examined.

Keywords: Power system analysis, blackout, blackstart, power system restoration, power grid

1. INTRODUCTION

One of the basic needs of human beings is electricity energy at nowadays. Everything from our life is utilized by electricity energy, from health to transportation, from fun to traffic (Pakize C., 2011). As a result, social and economic influences are inevitable in a interruption situation.

On energy systems that enable energy to reach consumers without problems in normal operating conditions, energy interruptions can occur due to disturbing currents (Zeng B., 2015). Especially when the power system, which is defined as power grid collapse, is partially or completely out of order, it is the most negative situation for power system operators. In the literature, blackout situations that can occur on power systems have been examined in detail. In one part of the studies done, the blakouts that have taken place on the world and have been realized so far have been examined and their causes and results have been questioned (Olga P, 2016), (Ji-Jen W, 2007), (ENTSOE, 2015), (UCTE, 2004), (Hordeski M., 2005), (Hines P., 2009). Restoration of the power grid after blackout occurred on the system is considered as an important issue (Kurup S., 2015), (En L., 2015), (Ketabi A., 2001), (Ancona J.J., 1995). The use of different energy sources for network restoration and the use of the micro,grid concept to eliminate blackouts has also been a major research topic (Gencheva R., 2015), (Castillo A., 2013), (Clean Technica, 2016).

In this study, blackout situations which can occur on power systems are defined, causes and results are examined. Some of the major blackouts on the world and in Turkey are presented. The use of production systems with blackstart for the restoration of the network has been investigated. The qualities to be found in the systems with blackstart feature are listed and their advantages are presented.

The rest of the work continues as follows. In section 2, informations about electricity networks and structures were made and power grid collapses were examined. Section 3 discusses the work done to restore power grid and the blackstart feature. Work with Section 4 has been terminated.

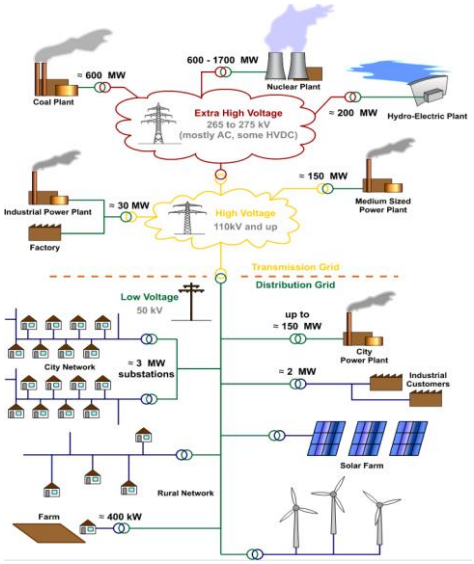
2. VOLTAGE COLLAPSE ON ENERGY SYSTEMS

Energy systems are the sum of equipment operated in coordination with each other. The goal of the energy system to operate smoothly is to ensure that energy is delivered to the consumer in an uninterrupted form. However, in some cases, all or partial of the system may be out of order. Voltage collapse also known as out of order of the system can occur under conditions of faults, overload, imbalance, inadequate reactive power support, and etc. (Ji-Jen W., 2007).

2.1. Electric Power Systems

Electric power systems are a integrated system that is built to deliver generated electricity to consumers. Basically, the power systems examined in three parts namely production systems, transmission systems and distribution systems, are shown in Figure 1 with their most basic form. Electric energy obtained by using different sources on a power system is transmitted to consumption points with high voltage lines as high as possible and delivered to consumers at the required voltage level.

Figure 1. Structure of the Electric Power Grid (MBizon, 2010)



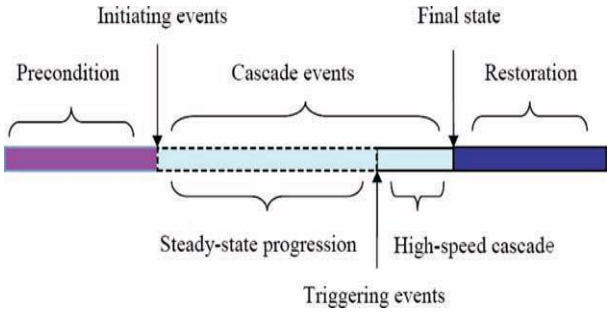
The continuous delivery of electricity to the user is directly related to the reliability of the system. Reliability can be improved by essential maintenance on the power systems, operating the system on optimal conditions, and expanding or renewing the system when the system is insufficient. On the other hand, partial out of service conditions may occur on the system, which in turn brings significant economic losses.

2.2. Blackout of Electric Power Grid

Electric power grid serves as a bridge between energy production plants and consumers. Energy supply-demand balance problems can ocur in time. These problems can cause the voltage and frequency of the power grid to increase or decrease. As a result of these occuring distortions, the power system is causing the power generation plants to be out of order through the protection relays. As a result, the synchronization in the power grid is starting to finish and the system is partially or completely collapsed. In other words, if the nominal frequency and voltage values of the system can not be nominal, the blackout of the system is inevitable (Zeng B., 2014).

Firstly, the preconditioning of the system must be investigated before the destructive effect, and the event that initiates the collapse should be investigated. The cascade events that come along with this event are the trigger of the collapse. If the system can not be intervened, the entire power grid can be lost. The first thing to do in the last phase of the collapse of the power grid is to initiate the restoration process (Yvon B.,2013). Phases of power grid collapse are shown in Figure 2.

Figure 2: The Phase of the Collapse of the Power Grid (Yvon B., 2013)



2.2.1. General Reasons for Power Grid Collapses

The collapses that occur in power systems arise due to five main factors in terms of system engineering (Yvon B., 2013). These are;

- Voltage loss
- Frequency loss
- Phase angle imbalance
- Production - consumption imbalance
- Synchronization loss

can be sorted as. The factors listed above do not occur under normal operating conditions. On the other hand, it can happen to collapse on the power grid with the effect of disruptive events that can occur inside or outside the system. The main disruptive effects that may cause blackouts on the system are listed below (Pakize C., 2011) :

- Overloading of the equipment used on the system
- Meteorological conditions
- Power system faults
- Fire, flood, earthquake etc. natural events
- Operation mistakes
- Wrong protection
- Sabotage, cyber attack, etc. situations
- Insufficient knowledge and experience

2.2.2. General Results of Power Grid Collapses

In the case of loss of energy in power systems or inability to reach consumers, both social and economic effects arise. In social terms, the problems that consumers may encounter are listed below:

- Health, education, transportation etc. services interruption
- Traffic problems beginng
- Lighting problems and related problems
- Communication network weakening
- Problems related to the inability to use electrical household equipment

In addition to all these, there are significant economic impacts due to the halt of production units, the failure of financial institutions, and the loss of industrial facilities.

2.2.3. Major Blackouts Occured on The World and Turkey

Blackout events can occur on power systems due to different reasons. Examples of headlines from the main blackout event on Earth are shown in Table 1 (Zeng B., 2015), (Yvon B., 2013), (ENTSOE, 2015), (UCTE, 2004), (UCTE, 2006).

Considering the interruptions that are taking place on the world, it is seen that 670 million subscribers in India in 2012 have been left without energy. In the case of Blackout in the Pakistani energy system, there are 160 million subscribers. In general, the duration of interruption varies between 5 and 10 hours. However, after the Blackout in Canada in 2003, the system recovered 192 hours of complete recovery (Olga P. V., 2016).

When the causes of the blackout events shown in Table 1 are examined, it is seen that there are interruptions due to overloading and heating problems especially in eastern countries. Besides these, operator errors, inadequate protection system, meteorological conditions and failures arise as other causes of blackout.

Table 2 presents the main blackout events and details seen in Turkey from 1999 to the present day. Table 2 shows that blackout events are caused by disasters, meteorological events and system operation problems. When the details of the Blackout incidents were examined, it was seen that in 1999 and 2016, the direct transmission system was interrupted as a result of the damage directly. The energy transmission lines were damaged due to the earthquake in 1999, and in 2016, demolition occurred in the poles due to excessive ice load (enerji.gov.tr, 2016). The beginning of the event in 2015 is emerging as a operation problem. During the event, production is concentrated in one region and consumption in another region. However, the energy system has been divided into two, the frequency drop in one part and the collapse in the other part due to the frequency increase (ENTSOE, 2015).

Table 1: Major Blackout Incidents on the World

	Date	Country		Date	Country
1	11.9.65	Northeast of US	**24**	12.1.98	San Francisco and California , US
2	5.1.77	Miami, US	**25**	7.1.99	New York City, US
3	7.1.77	New York City, US	**26**	3.11.99	Brazilian power system
4	12.19.78	France	**27**	1.2.01	India
5	1.1.81	Idaho, Utah and Wyoming, US	**28**	1.12.03	Croatia and Bosnia Herzegovina
6	3.1.82	Oregon, United States	**29**	3.31.03	Iran
7	12.27.83	Sweden	**30**	8.14.03	Northeast of US and Canada
8	7.23.87	Japan	**31**	8.28.03	South London
9	1.12.87	Western France	**32**	9.5.03	The West Midlands, UK
10	3.13.89	Qubec, Canada	**33**	9.23.03	Eastern Denmark and Sweden
11	8.24.94	Italy	**34**	9.28.03	Italian power system
12	12.14.94	Arizona and Washington, US	**35**	11.7.03	Most of Chile
13	1.17.95	Japan	**36**	7.12.04	The Athens and Southern Greece
14	6.8.95	Israel	**37**	3.14.05	Queensland, South Australia
15	3.12.96	Florida, US	**38**	5.25.05	Moscow, Russia
16	4.16.96	Southwestern of US	**39**	9.24.06	Pakistan
17	7.2.96	14 states in the US	**40**	11.4.06	European power system
18	8.7.96	Big Rivers Electric Corporation, US	**41**	4.26.07	Colombia
19	8.10.96	California Pacific Northwest, US	**42**	2.26.08	Florida, US
20	8.26.96	New York City, US	**43**	2.4.11	Brazilian power system
21	9.21.96	Allegheny Power System, US	**44**	9.8.11	Arizona and Baja California, US
22	10.30.96	New York City, US	**45**	7.31.12	India
23	1.1.98	Canada, New York, New England			

Each blackout event on the Turkish power system has been examined in detail and future projections have been established depending on the reasons. In this context, additional supply scenarios were developed, load and reserve power management were updated, necessary renewals were made in the energy system.

Table 2: Main Blackout Incidents in Turkey

Date	Initiating Point	Reason	Restoration Time	Result
August 17, 1999 (03:02)	Marmara Region	Natural disaster (Earthquake)	The full restoration with regional restoration continued until the evening of August 18th.	62 milyon kişi elektriksiz kalmıştır.
January 14, 2012 (13:43)	Marmara Region (Bursa)	Bursa Natural Gas Combined Cycle Power Plant's fall in natural gas pipeline due to cold weather	From 14:15 hours, electricity was supplied to Edirne, Tekirdağ and Kırklareli in Trakya region. As of 16:30 hours, 80% of the interruption was eliminated and 90% of the interruption was eliminated by 19:00 hours. At 19:15 the operation of the system returned to normal operation condition.	Istanbul, Sakarya, Kocaeli, Tekirdağ, Kırklareli and Edirne, more than 20 million people living in electricity without electricity.
March 31, 2015 (09:36)	Osmanca – Kurşunlu energy transmission lines	Out-of-service Osmanca,Kurşunlu energy transmission lines carrying 1127 MW / 1237 MVA because of overload	As of 11:30 am, the Black Sea and the Eastern Anatolia Region were synchronized. At 16:12, the entire system was restored.	The entire power grid is left without energy.
December 29, 2016 (14:20)	Istanbul	Seven different energy transmission lines feeding Istanbul have been affected by ice loads and storms	At 21:50, electricity could be supplied to the whole of Istanbul	Iron and steel factories and cement factories have not provided electricity to the organized industrial zones and the demand has been minimized with partial interruptions in the districts.

2.2.4. Solution Suggestions for the Prevention of Power Grid Blackouts

Electricity that is uninterrupted and reliable is so important for consumers. It is imperative to take some precautions not to cause any interruption in the power system. The following measures can be taken to prevent possible collapses in the system:

- Strategies for load management and risk factors should be identified
- Operators must be trained and informed
- The equipment used on the system must be regularly checked
- System should be developed depending on increasing supply and demand
- Necessary precautions should be taken for extreme weather conditions
- Security measures should be taken against sabotages and cyber attacks that may occur in the system.

3. RESTORATION OF ELECTRIC POWER GRID

In some cases, the operating conditions of the power system may be broke down. This distortion can cause the power system to collapse or break down. Once such disturbances have occurred, the power should be returned to normal operating conditions as soon as possible (Sreeram R. K., 2015). Hence, it is imperative that the network restoration program of each transmission system operating institution (NERC, 2015). Evaluation of restoration process consists of preparation of subsystems and establishment of target systems. The most important issue in the restoration process is that the evaluation is done carefully and the correct decisions are taken. A delay or wrong decision at the time of the determination causes very serious economic loss in the system.

The operating conditions in the restoration process differ from the work performed under normal operating conditions. There are some technical issues to consider during restoration conditions. These are the protection of the voltage and frequency values in the system, reactive power control, production and load balancing, energizing the units that need to be energized first, communication and coordination of production operators and operation operators that control the system are required. Another important issue in the restoration process is the creation of controlled islands around the units with blackstart features. With the aid of these islands, the system must be energized step by step and controlled to feed the loads.

It is a very important issue to reduce the recovery period to a minimum during the restoration of the electricity grid. Load distribution centers, transformer centers and operators in power plants should be in very good coordination during the restoration.

As a result of the partial collapse of the electricity network, the production units that have been disabled are energized by utilizing the neighboring systems. However, as a result of the complete system collapse, the system will be energized with blackstart-enabled power generation units.

3.1. Introduce of the Blackstart Feature of Power Plants

Blackstart is a feature that allows a production unit or station to switch from a blackout condition to an operating state that provides electrical energy without the help of an electrical grid (Wikipedia, 2014). In normal operating conditions, the electricity used in the power plant is supplied by the production units in the power plant. If all the main production units of the power plant are off, the electricity required by the plant is provided by the energy transmission lines connected to the transmission line of the power plant. If the power grid is blackout completely, the electricity grid is also de-energized. A power generation unit with a blackstart feature is required to return the electrical grid to normal conditions. Some blackstart generators are available to provide blackstart for power plant (Few MW capacities). These generators are used to start large power units. This process enables the entire plant to be energized and energize to the power grid (Agrawal V., 2010).

Nuclear and thermal power plant units must have a ratio of 7% to 8% of their energy production capacity to energize units with blackstart capability. Power capacity of hydroelectric power plants should be between 0.5% and 1% while the energy production capacity should be between 1.5% and 2% on natural gas power plants (The National Grid, 2001).

3.2. Energy Production Units Used to Provide Black Start

Generally, hydroelectric power plants are designated as blackstart sources for energizing grid connections because of the hydroelectric power plant needs very little energy to start. Therefore it is most common to use as a blackstart source (Wikipedia, 2014).

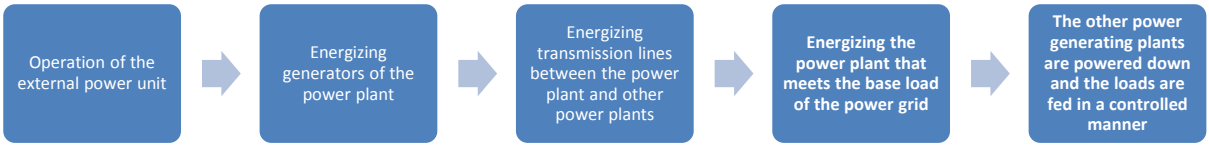
Providing blackstart capacity is not economical to create a large standby capacity in fossil fueled power plants because the equipment has a high internal energy requirement. However, generator sets with the blackstart feature are produced nowadays. Together with the hydroelectric units, diesel generator sets, gas turbine generator sets and gas turbines of large dimensions can be used as a blackstart source (The National Grid, 2001).

In February 2016, a Berlin-based company explained that a battery park with a capacity of 5 MW would be integrated with renewable energy sources and used as a blackstart source. It is announced that the project will last three year (Rob Compton, 2016).

3.3. System Energization with Blackstart Featured Power Plants

Blackouts in energy system are rare, but sometimes they can occur. As a result of the blackout of the energy system, the power plants are completely without energy. In this case, generators with the BlackStart feature should be fed primarily with external support units so that the system returns to the normal conditions again. Then the other units of the power plant are energized. Next, out of ordered energy transmission lines in the energized state are also energized. Then the base load plants are energized. Then, other power plants should be stepped in. Then, other power plants should take part step by step on power grid.

Figure 3: Energizing the Power Grid with the Blackstart Feature



4. CONCLUSION

Electric energy is utilized at every point of our daily life. As a result, interruptions that may occur on the power grid lead to significant loss. In this respect, blackouts which cause the network to be partially or completely out of service, need to be analyzed in detail.

In this study, the blackout situations that can occur on the electric networks have been examined. The important blackouts encountered on the world and Turkey have been taken into consideration and their causes and results have been analyzed. Measures to be taken to avoid blackout situations have been put forward. In addition to this, the effects of the blackstart stations on the power grid on the recovery of the power grid have been examined.

Considering this study, it has been observed that the use of different energy sources on the power grid both reduces blackout risk and contributes to power grid restoration by blackstart feature.

REFERENCES

Pakize C., 2011, “Examination of the Blackouts Occured in Interconnected Networks”, Graduate Theses, Marmara University, İstanbul.

Zeng B., Ouyang S., Zhang J., Shi H., Wu G., Zeng M., February 2015, “An analysis of previous blackouts in the world: Lessons for China's power industry”, Renewable and Sustainable Energy Reviews, Volume 42, Pages 1151-1163.

Olga P. V., Francisco S., September 2016, “Analysis of major blackouts from 2003 to 2015: Classification of incidents and review of main causes”, The Electricity Journal, Volume 29, Issue 7, Pages 42-49.

Ji-Jen W., Ching-Tzong S., Chi-Shuan L., Chung-Liang C., October 2007, “Study on the 729 blackout in the Taiwan power system”, International Journal of Electrical Power & Energy Systems, Volume 29, Issue 8, , Pages 589-599

TEİAŞ, ENTSOE Project Group Turkey, September 2015 “Report on Blackout in Turkey on 31st March 2015”

UCTE, April 2004, “Final Report of the Investigation Committee on the 28 September 2003 Blackout in Italy”

UCTE, 2006, “Final Report System Disturbance on 4 November 2006 in Europe”

Yvon B., Mircea E., Nikolai V., 2013, “Major grid blackouts: analysis, classification, and prevention”, IEEE Press, Part 3, Page 813-887 New Jersey.

Alberto B., Carlo A. N., Mario P., 2013, “Restoratıon Processes After Blackouts”, IEEE Press, Part 3, Page 888-923 New Jersey.

Jerry C. W., 2007, “AC Power Systems Handbook”, Third Edition, CRC Press, Page 330-341, California.

Hordeski M., 2005, Emergency and Backup Power Sources: Preparing for Blackouts and Brownouts, The Fairmont Press, Inc., United States of America.

Gencheva R., Stefanova A., Groudev P., 2015, ‘’ Investigation of station blackout scenario in VVER440/v230 withRELAP5 computer code’’ , Elsevier, Vol. 295, pp. 441–456.

Hines P., Apt J., Talukdar S., 2009, ‘’ Large blackouts in North America: Historical trends and policy implications’’, Elsevier, Vol. 37, pp. 5249–5259.

Welfonder E., Kurth M., 2007, ‘’ Limitation control procedures, required for power plants and power systems Possibility for reducing future blackouts’’, Elsevier, Vol. 31, pp. 93–104.

Pradhan P., Deki T., Wangmo P., Dorji D., Phuntsho D. and. Dorji C., 2015, ‘’Simulation and Optimization of Blackstart Restoration Plan in Bhutan using DIgSILENT’’, Proceedings of 2015 RAECS UIET Panjab University Chandigarh 21-22nd December 2015.

Castillo A., 2013, ‘’ Microgrid Provision of Blackstart in Disaster Recovery for Power System Restoration’’, IEEE Smart grid Comm 2013-Smart Grid Services and Management Models.

Kurup S., Ashok S., 2015, ‘’ Grid Restoration Planning After a Total Blackout’’, 2015 International Conference on Control, Communication & Computing India (ICCC) | 19-21 November 2015 | Trivandrum.

Wang H., Lin Z, Wen F., Ledwich G., Xue Y., Zhou Y., Huang Y., 2014, '' Black-start decision-making with interval representations of uncertainfactors'', Elsevier, Vol. 79, pp. 34–41.

"Black Start Restoration Procedures Of Southern Region 2013'', Southern Regional Load Despatch Centre Power System Operation Corporation Ltd, November 2012, Bangalore.

Agrawal V., Porwal R., Kumar R., Pandey V., 2010, '' Mock Blackstart Drills – An Excellent Learning Experience for Power System Operators'', CBIP- 5th International Conference on Power System Protection and Automation, 6-9, Dec 2010.

En L., Zhijun Q., Mingbo L., Yunhe H., Ning W., Jinyu W., August 2015, "Increasing Black Start Capacity by Fast Cut Back Function of Thermal Power Plants", International Journal of Smart Grid and Clean Energy.

Ketabi, A., Asmar, H., Ranjbar A.M, Feuillet, R., 2001, "An approach to initiation of optimal units during the restoration of the bulk power system"," 2001, Large Engineering Systems Conference on Power Engineering, Theme: Strengthening after 2001 (Kat. No.01ex490) , Halifax, NS, 2001, s. 190-194.

Hafner, Y.J., Duchen, H., Karlsson, M., Ronstrom, L., Abrahamsson, B., 2008, "HVDC with Voltage Source Converters-A Powerful Standby Black Start Facility", IEEE PES T&D Conference, 21-24 April 2008, Chicago.

The National Grid Company plc, 2001, "An Introduction to Black Start", Market Development, February 2001

Ancona, J.J., 1995, "A Framework For Power System Restoration Following a Major Power Failure", IEEE Transactions on Power Systems, Vol. 10, No. 3, August 1995.

Clean Technica, Restoring The Grid After A Blackout — Using Batteries, https://cleantechnica.com/2016/02/02/restoring-the-grid-after-a-blackout-using-batteries/, 24 April 2017.

Sun, W., Liu, C.C., Liu, S., 2011, "Black Start Capability Assessment in Power System Restoration", IEEE, 978-1-.4577-1002-5/11.

Lindenmeyera, D., Dommela, H.W., Adibi M.M., 2001, "Power system restoration- A bibliographical survey", Electrical Power and Energy Systems, 23: 219-227.

Feltes, J.W., Moran, C.G., 2008, "Black Start Studies for System Restoration", IEEE, NewYork 12305.

Zeng, S., Lin, Z., Wenb, F., Ledwich, G., 2012, "A new approach for power system black-start decision-making with vague set theory", Electrical Power and Energy Systems, 34: 114-120.

Joglekar, J.J., Nerkar, Y.P., 2008, "A different approach in system restoration with special consideration of Islanding schemes", Electrical Power and Energy Systems, 30: 519-524.

Barsali, S., Giglioli R., Poli, D., Sforna M., Salvati R., Zaottini R., 2008, "The restoration of an electric power system: International survey and discussion of possible innovative enhancements for the Italian system", Electric Power Systems Research, 78: 239-247.

Jiang Y., Chen, S., Liu C.C., Sun W., Luo, X., Liu, S., Bhatt, N., Uppalapati, S., Forcum, D., 2017, "Blackstart capability planning for power system restoration", Electrical Power and Energy Systems, 86: 127-137.

Barsali S., Polia, D., Praticò, A., Salvati, R., Sforna, M., Zaottini, R., 2008, "Restoration islands supplied by gas turbineS", Electric Power Systems Research, 78: 2004-2010.

Kleppingera, D., Broadwatera R., Scirbona C., 2010, "Generic reconfiguration for restoration", Electric Power Systems Research, 80: 287–295.

Yılmaz, O., 2011, "System Restoration and "Black start" Qualification of Technical Units, Technical Criteria, Evaluations and Test Procedures", Tübitak Space Technologies Research Institute –Teiaş-Ekh Tdep, 30 September 2011.

Lopes, J.A.P., Moreira, C.L., Resende, .O., 2005, "MicroGrids Black Start and Islanded Operation", 15th Power Systems Computation Conference, 22-26 August 2005, Liege.

Zhang, G., 2009, "EPRI Power Systems Dynamics Tutorial", Palo Alto, 1016042.

Sun, W., 2011, "New optimization techniques for power system generation scheduling", Graduate Theses and Dissertations, Iowa State University, Iowa City, p.1220.

Zhang, X.P., 2015, "Emerging Capability on Power System Modelling: System Security, Resilience and Recovery Modelling", The Instution of Engineering and Technology, 8:1-15.

MBizon, 2010, "Electricity Grid Schematic English", wikipedia.org, commons.wikimedia.org/wiki/File:Electricity_Grid_Schematic_English.svg.

Wikipedia, 2014, en.wikipedia.org/wiki/Black_start.

Nort American Electric Reliability Corporation(NERC), December 2015, "2015 Long-Term Reliability Assessment", Atlanta.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

AN OPTIMIZATION STUDY ON DISSOLUTION OF THE CALCINED COLEMANITE MINERAL IN METHYL ALCOHOL BY CO_2 IN AN AUTOCLAVE SYSTEM USING TAGUCHI METHOD

DOI: 10.17261/Pressacademia.2017.590
PAP-WCTIE-V.5-2017(28)-p.198-204

Meltem Kizilca[1], Mehmet Copur[2]
[1]Department of Chemical Engineering, Ataturk University, Erzurum, Turkey mkizilca@atauni.edu.tr
[2]Department of Chemical Engineering, Bursa Technical University, Bursa, Turkey

ABSTRACT

The process of optimizing the dissolution of colemanite ore ($2CaO.3B_2O_3.5H_2O$) in methyl alcohol by CO_2 in a high-pressure reactor was evaluated by using the Taguchi method. Optimum conditions for the cholemanite mineral were determined to be as follows: reaction temperature: 140°C, reaction time: 50 min, calcination time: 240 min, calcination temperature: 450 °C, solid-liquid ratio: 1/6 g.mL^{-1}, mixing rate: 500 cycle/min, pressure: 20 bar, grain size: -100 mesh and the amount of $CaCl_2$: 7.5 g.

Keywords: Colemanite, optimization, taguchi method.

1. INTRODUCTION

Boron is abundantly found in nature and the world's largest boron mineral reserve is located in Turkey. Boron is an economically valuable mineral that is not found in its natural state, but forms borate minerals with alkali and alkaline earth metal oxides of elements like sodium, calcium, and magnesium (Alkan and Doğan, 2004). Among these minerals, tincal, pandermite, ulexite, and colemanite have the highest percentages of boron reserve and production (Demirbaş et al. 2000). Colemanite chemically known as $2CaO.3B_2O_3.5H_2O$ is a boron mineral widely found in Turkey and is commercially valueable (Bayca, 2013). It is highly important to develop these boron reserves that form one of the most important underground treasures of Turkey, obtain new products from these reserves, and develop alternative economic and more environmentally-friendly technologies for present products. The most used source to obtain boron compounds in industry is boric acid (B_2O_3) (Şayan et al. 2006). Boric acid is formed as a result of reactions of boron minerals with various acid solutions.

In this study, the purpose is to develop an environmentally-friendly method to produce trimethyl borate by using colemanite ($2CaO.3B_2O_3.5H_2O$) and CO_2. The purpose of this study is to investigate optimum conditions for the dissolution of colemanite ore in methanol by CO_2 in a high-pressure controlled and high temperature reactor using the Taguchi method, and to develop an alternative process to produce trimethyl borate.

2. LITERATURE REVIEW

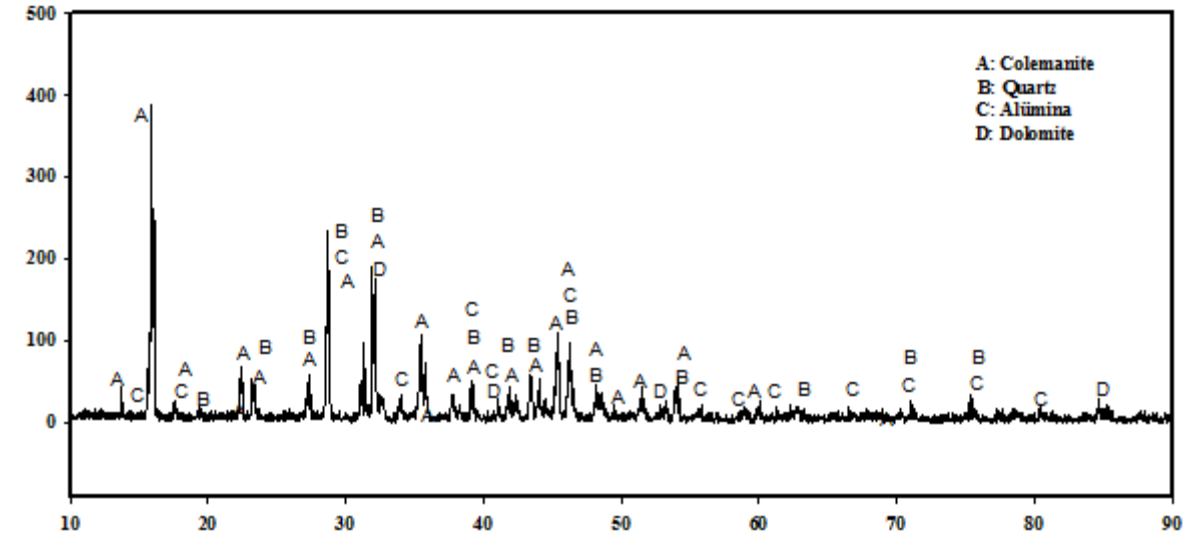
In the literature, the dissolution of boron minerals in various acid solutions was investigated, for example hydrochloric acid (Sis et al. 2015), phosphoric acid (Copur et al. 2000; Doğan and Yartaşı, 2014), sulphuric acid (Uysal et al. 2015), oxalic acid (Alkan and Doğan, 2004; Bayca et al, 2013), citric acid (Çiftçi, 2012), ammonium chloride (Alkan et al. 2013), ammonium sulfate (Kocakerim et al. 2007), propionic acid (Bulutçu et al. 2008) water saturated with chlorine gas, water saturated with both sulphurdioxide and carbondioxide (Kocakerim and Alkan, 1988; Şayan et al. 2006).

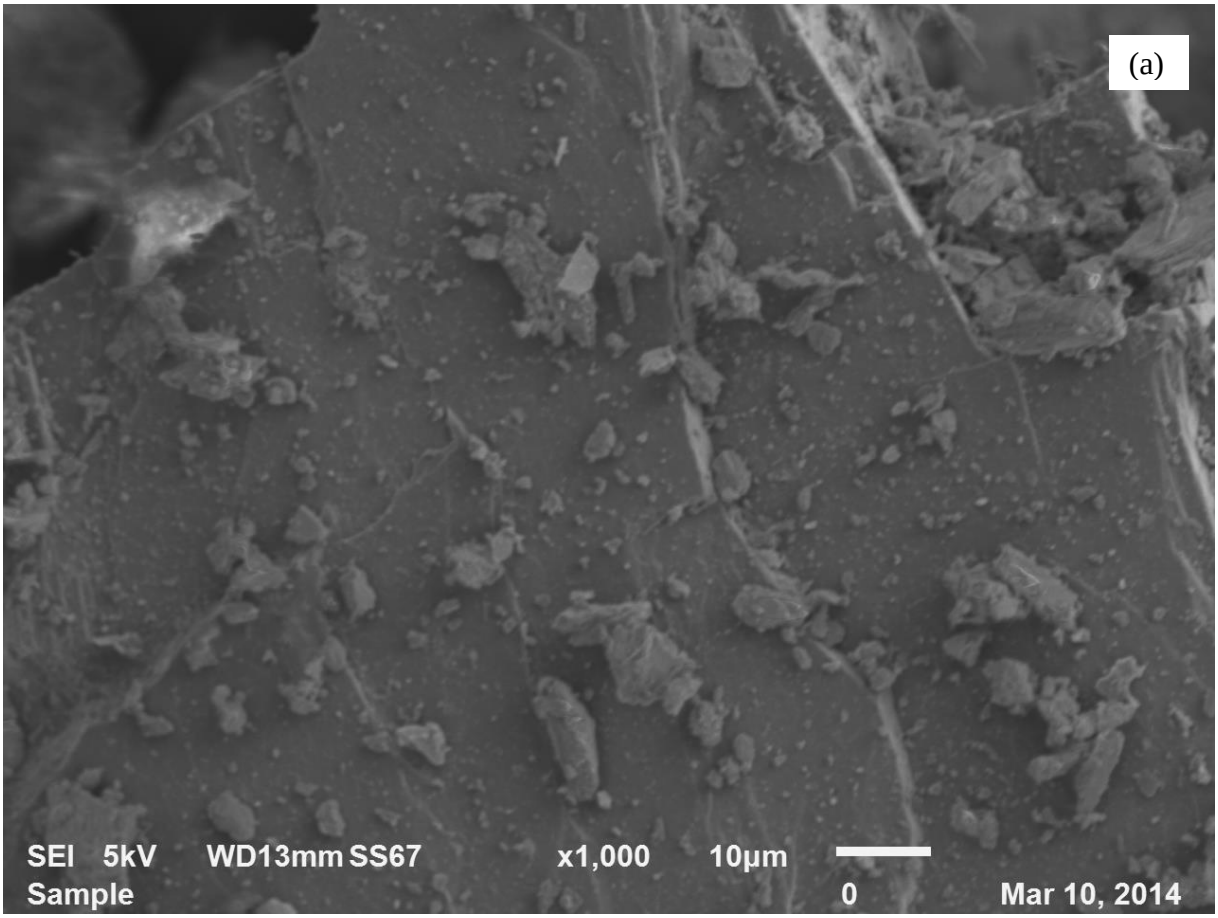
3. METHODOLOGY AND DATA

Colemanite ore used in this study was obtained from Bursa, in Turkey's Kestelek region. The ore was ground and separated after sieving by standart sieves into four fractions: 297+250, -250+177, -177+149, and -149+125 µm. Mineralogical analysis

of the ore was done by X-ray diffraction (XRD), while its elemental analysis was done by Scanning Electron Microscopy-Electron Dispersive Spectroscopy (SEM-EDS). The result of XRD analysis is shown in Figure 1. The colemanite mineral, based on EDS results, as shown in Figure 2, includes 24.13% B, 57.91% O, 12.40% Ca, 4.20% C, 0.24% Mg, 0.30% Al, 0.67% Si, and 0.50% K. Thus, a majority of colemanite consists of B, O, Ca and C, while Mg, Al, Si and K exist in small quantities.

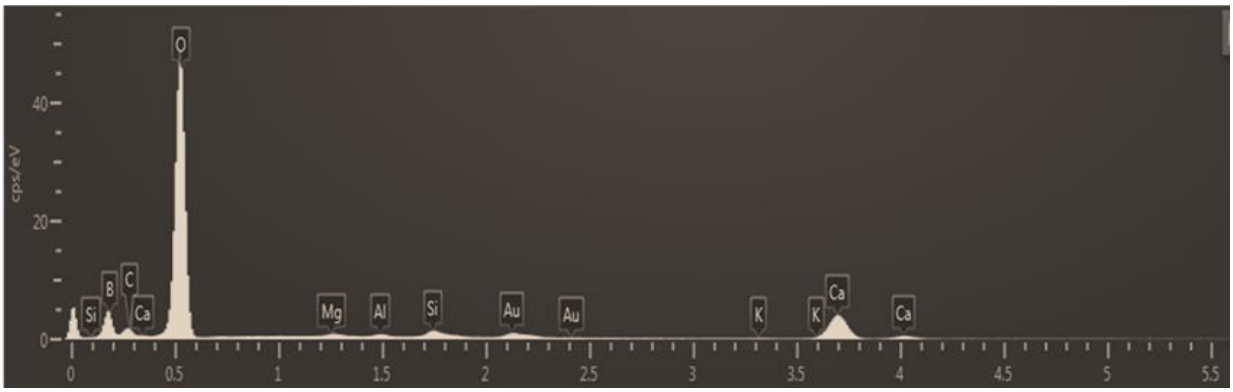
Figure 1: XRD Pattern of the Colemanite Mineral





(b)

Figure 2: SEM Micrograph (a) and EDS Pattern (b) of Colemanite Mineral



The Taguchi method is an experimental design method that tries to minimize variabilities in the product and process from uncontrollable and changeable factors by selecting the optimum combination of controllable factors. An orthogonal series a special form of the fractional factorial design is used in this method. How this method differs from other statistical experimental design methods is that it evaluates the parameters affecting an experiment in two separate groups: i. e. controllable and uncontrollable parameters, and enables the researchers to examine numerous parameters at more than two levels. Because uncontrollable parameters cause high costs, it is appropriate to investigate levels of controllable parameters to minimize effects of uncontrollable parameters. Generally, the performance characteristic of each product or process should have a nominal or target value. The purpose is to decrease the variability around that target value. Optimum operating conditions to be determined at the end of the experimental study should usually give the same or similar performance values under different operating conditions or times. The optimization criteria to be used for this purpose should control the variability around the performance value at a minimum level. In the Taguchi method, such an optimization criterion is the performance statistics, SN (signal to noise ratio). The equations below are used to calculate the performance statistics (Pignatiello, 1988; Phadke, 1989):

Larger the better:

$$SN_L = -10Log\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_1^2}\right] \quad (1)$$

Smaller the better:

$$SN_S = -10Log\left[\frac{1}{n}\sum_{i=1}^{n}Y_i^2\right] \quad (2)$$

Nominal is the best:

$$SN_S = -10Log\left[\sum_{i=1}^{n}Y_i^2/S^2\right] \quad (3)$$

where n is the number of repeated experiment; y_i is the performance value; SN_L, SN_S, and SN_T indicate performance statistics. If the purpose in a process is to reach maximum value, parameter levels, making the SN_L value maximum are optimum. If the purpose is to obtain the minimum, then parameter levels making the SN_S maximum are optimum. In the Taguchi method, if the experiment corresponding to optimum operating conditions is not available in the experimental plan, an additional model can be used by utilizing Equation 4 to estimate the performance value, i.e. dissolution percentage.

$Y_i = \mu + X_i + e_i$ (4)

where Y_i is the estimated performance value of the i^{th} experiment; μ is overall mean of the performance value; X_i is the overall activity size of parameter levels used in the i^{th} experiment; e_i is the experimental error. If experimental results are given in percentage, omega transformation of the obtained percentage values is performed before using the Y_i equation. Later, the omega value estimated for optimum conditions:

$$\Omega(\text{db}) = -10\text{Log}\left(\frac{1}{\text{P}} - 1\right) \quad (5)$$

where Ω (db) is the db value of the omega transformed percentage value and P is % value of the product obtained experimentally. Because the Y_i equation calculated based on experimental results is point estimation, a confidence interval for the error of estimation must be constructed to decide whether an additional model is sufficient. The confidence interval is found by using the equation below (Çopur and et al. 2000; Çopur and Kızılca, 2015):

$$S_e = \pm 2\sqrt{\left[\frac{1}{n_0}\right] + \sigma_e^2 + \left[\frac{1}{n_r}\right]\sigma_e^2} \quad (6)$$

$$\sigma_e^2 = \frac{sum\ of\ \ squares\ due\ to\ error}{degrees\ of\ \ freedom\ for\ \ erro} \quad (7)$$

$$\frac{1}{n_o} = \frac{1}{n} + \left[\frac{1}{n_{Ai}} - \frac{1}{n}\right] + \left[\frac{1}{n_{Bi}} - \frac{1}{n}\right] + \left[\frac{1}{n_{Ci}} - \frac{1}{n}\right] \quad (8)$$

Where S_e is the confidence interval for the effects of parameters; σ_e^2 is the variance of the error; n_r is the number of repetition of validation experiments; *nAi*, *nBi*, *nCi*…, *Ai*, *Bi*, *Ci*…indicate the levels of experiments (Phadke, 1989). If the estimated error is out of these limits, the additional model is not necessary. Validation experiments are a good tool to determine whether there is an internal interaction between parameters controlling the process. If the value estimated for optimum conditions is close to the observed value, this phenomenon indicates that there is no internal interaction between parameters. Otherwise, there is an internal interaction.

Experiments were conducted in a high-pressure resistant, temperature-controlled steel reactor having a mixing setup. For the calcination studies, colemanite ore was poured in a steel plate of 5x25 cm and 2 mm thick. Calcination treatment was performed in an ash oven of 20x32 cm under conditions determined in the experimental design. Parameters determined for this system are calcination temperature and time, reaction temperature, solid to liquid ratio, pressure, grain size and amount of calcium chloride. For optimization, orthogonal series were used as the Taguchi experimental design (Roy, 1990). The solid substance and methyl alcohol determined in accordance with the solid to liquid ratio were added into the reactor, which was adjusted for the conditions determined. After the reactor content reached the desired temperature value, the experiments were initiated by adjusting the reactor pressure to the desired value with CO_2 gas. At the end of reaction, B_2O_3 analyses were performed in the precipitates of solutions by filtering them with a vacuum pump.

Parameters used in this study and selected values are shown in Table 1 while the experimental plan based on the $L^{18}(3^7)$ orthogonal series is shown in Table 2.

Table 1: Parameters Studied in Experiments and Their Levels

Parameters		Levels		
		1	2	3
A	Mixing rate (rpm)	500	700	
B	Calcination temperature (^oC)	400	450	500
C	Calcination time (dak)	120	180	240
D	Solid-liquid ratio (gmL^{-1})	1/4	1/5	1/6
E	Grain-size (mesh)	-80	-100	-120
F	Reaction temperature (^oC)	130	140	150
G	$CaCl_2$ amount (g)	7.5	10	15

Table 2: L^{18} (3^7) Orthogonal Array Experimental Plan and the Amounts of B_2O_3 Dissolved in the Experiments

Exp.no	A	B	C	D	E	F	G	$\%B_2O_3$	$\%B_2O_3$	$\%B_2O_3$	$\%B_2O_3$
1	1	1	1	1	1	1	7.5	83.46	82.95	83.49	83.50
2	1	1	2	2	2	2	10	94.86	94.70	94.51	94.65
3	1	1	3	3	3	3	15	96.14	95.14	95.85	96.11
4	1	2	1	1	2	2	15	87.32	86.75	87.11	87.20
5	1	2	2	2	3	3	7.5	94.63	93.90	93.80	93.85
6	1	2	3	3	1	1	10	96.37	95.87	96.10	95.90
7	1	3	1	2	1	3	10	93.88	94.12	93.50	93.95
8	1	3	2	3	2	1	15	95.34	94.90	95.20	95.32
9	1	3	3	1	3	2	7.5	94.45	95.25	94.15	94.30
10	2	1	1	3	3	2	10	96.14	97.20	95.99	95.85
11	2	1	2	1	1	3	15	87.16	86.95	87.01	87.19
12	2	1	3	2	2	1	7.5	95.30	95.20	95.21	95.25
13	2	2	1	2	3	1	15	94.74	94.70	94.81	94.80
14	2	2	2	1	1	2	7.5	97.97	97.90	97.99	97.85
15	2	2	3	2	2	3	10	85.88	84.99	85.94	85.90
16	2	3	1	2	2	3	7.5	97.02	97.12	97.22	97.25
17	2	3	2	3	3	1	10	79.49	79.54	80.10	79.85
18	2	3	3	1	1	2	15	94.99	94.45	94.91	94.90

Table 3: L^{18}(3^7) Orthogonal Series

	A	B	C	D	E	F	G
Exp.No	**Mixing rate**	**Reaction temperature**	**Pressure**	**S/L ratio**	**Reaction time**	**Grain size**	**$CaCl_2$ amount**
1	500	130	10	20/200	40	-80	2.5
2	500	130	20	30/200	50	-100	7.5
3	500	130	30	40/200	60	-120	15
4	500	140	10	20/200	50	-100	7.5
5	500	140	20	30/200	60	-120	3.75
6	500	140	30	40/200	40	-80	10
7	500	150	10	30/200	40	-120	7.5
8	500	150	20	40/200	50	-80	15
9	500	150	30	20/200	60	-100	2.5
10	700	130	10	40/200	60	-100	10
11	700	130	20	20/200	40	-120	7.5
12	700	130	30	30/200	50	-80	3.75
13	700	140	10	30/200	60	-80	11.25
14	700	140	20	20/200	40	-100	2.5
15	700	140	30	30/200	50	-120	2.5
16	700	150	10	30/200	50	-120	7.5
17	700	150	20	40/200	60	-80	10
18	700	150	30	20/200	40	-100	7.5

3. FINDINGS AND DISCUSSIONS

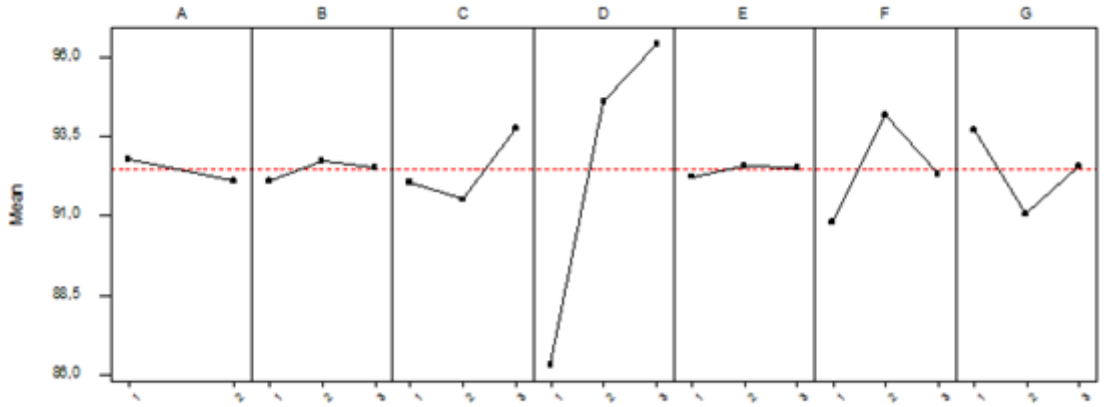
The reaction between colemanite, CO_2 and methyl alcohol can be expressed as follows:

$$2CaO.3B_2O_3.5H_2O+3CO_2+9CH_3OH \rightarrow 3H_3BO_3+3B(OCH_3)_3+2Ca(HCO_3)_2+4H_2O+CaCO_3 \quad (9)$$

Dissolution percentage values of B_2O_3 are shown in Table 3. Data obtained from the trials and analyses were evaluated by using the Minitab package program.

In the Figure 3, between the performance statistics and levels of parameters, the numerical value of maximum points shows the best value of the related parameter. Accordingly, the following were obtained: optimum operating conditions A_1 (mixing rate: 500 cycles/min), B_2 (calcination temperature: 450°C), C_3 (calcination time: 240 min), D_3 (solid/liquid ratio: 1/6), E_2 (grain size: -100 mesh), F_2 (reaction temperature: 140°C) and G_1 (the amount of $CaCl_2$: 7.5 g). Under these conditions, the passage ratio of B_2O_3 to the solution was found to be 98.4%. An F test was conducted to observe which parameters selected for optimization experiment designs are significant to the effects of impurities. Effective parameters were determined to be the solid-liquid ratio, reaction temperature, calcination time, and the amount of $CaCl_2$.

Figure 3: **The Effect of Each Parameter on the Optimization Criteria for B_2O_3.**



REFERENCES

Alkan, M., Dogan, M., 2004, "Dissolution kinetics of colemanite in oxalic acid solutions", Chemical Engineering and Processing, vol.43, 867–872.

Bayca, S.U. 2013,"Microwave radiation leaching of colemanite in sulfuric acid solutions", Separation and Purification Technology, vol.105, 24–32.

Bayca, S.U., Kocan, F, Abalı Y.2013, "Dissolution of colemanite process waste in oxalic acid solutions", Environmental Progress & Sustainable Energy, vol. 33, no.4, 1111-1116

Bulutcu, A.N., Ertekin C.O., Kuskay Celikoyan, M.B.,2008. Impurity control in the production of boric acid from colemanite in the presence of propionic acid. Chemical Engineering and Processing: Process Intensification, 47 (12), 2270–2274.

Copur, M., Temur, H., Yartası, A., Kocakerim, M.M. 2000, "The Kinetics of Dissolution of Colemanite in H3PO4 Solutions", Industrial & Engineering Chemistry Research, vol.39, 4114-4119.

Çiftçi, H. 2012, "Modelling and kinetic analysis of boric acid extraction from ulexite in citric acid solution", Canadian Metallurgical Quarterly: The Canadian Journal of Metallurgy and Materials Science, vol.51, no 1, 1-10.

Çopur, M., Kızılca, M.2015, "Determination of the Optimum Conditions for Copper Leaching from Chalcopyrite Concentrate Ore Using Taguchi Method", Chemical Engineering Communications, vol.202, no.7, 927-935.

Demirbaş, A., Yüksek, H., Çakmak, İ., Küçük, M.M., Cengiz, M., Alkan, M. 2000, "Recovery of boric acid from boronic wastes by leaching with water, carbon dioxide- or sulfur dioxide-saturated water and leaching kinetics. Resources", Conservation and Recycling, vol.28, no.1-2, 135-146.

Doğan, H., Yartaşı, A.2014" Optimization of Dissolution of Ulexite in Phosphate Acid Solutions", Journal- Chemical Society of Pakistan,vol. 36, no.4, 601-605.

Kocakerim, M.M., Alkan, M. 1988, "issolution of ulexite in SO2-saturated water". Hydrometallurgy, vol.19,385.

Kocakerim, M.M., Tunç, M., Küçük Ö., Aluz, M. 2007, "Dissolution of colemanite in (NH4)2SO4 solutions", Korean Journal of Chemical Engineering., vol.24, 55-5.

Phadke, M. S., 1989, "Quality Engineering using Robust Design,Prentice Hall: New Jersey, 61–292.

Pignatiello, J. J. 1988, IEE Transactions, vol. 20, 247.

Roy, R.K. 1990,"A Primer on the Taguchi method, Van Nostrand Reinhold Company", New York, USA,.

Sayan, E., Ekinci Z., Bese, A.V., Ata O.N., 2007, "Optimization and modeling of boric acid extraction from colemanite in water saturated with carbon dioxide and sulphur dioxide gases" International Journal of Mineral Processing, vol.82, no.4,187–194.

Sis, H., Bentli, İ. Ekmekyapar, A., Demirkıran, N.2015. Dissolution Kinetics And Variation Of Particle Size Of Kırka (Turkey) Ulexite in HCl Solutions. Journal of Ore Dressing. 17 (33), 27-37.

Tekin, G., Onganer, Y., Alkan, M. 2013, "Dissolution Kinetics of Ulexite in Ammonium Chloride Solution" Canadian Metallurgical Quarterly: The Canadian Journal of Metallurgy and Materials Sciences. Vol. 37, no.2, 91-97.

Uysal, B. Z., Demirel, H. S., İnce T. E., Uysal, D. 2015, "Boric Acid Production From Sodium Metaborate With Sulfuric Acid". CBÜ Fen Bil. Dergi., vol.11, no.3, 379-382.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

A SURVEY AUTOMATIC TEXT SUMMARIZATION

DOI: 10.17261/Pressacademia.2017.591
PAP-WCTIE-V.5-2017(29)-p.204- 213

Oguzhan Tas[1], Farzad Kiyani[2]
Istanbul Sabahattin Zaim University, Computer Engineering Department. oguzhan.tas@std.izu.edu.tr
Istanbul Sabahattin Zaim University, Computer Engineering Department. farzad.kiyani@izu.edu.tr

ABSTRACT

Text summarization is compress the source text into a diminished version conserving its information content and overall meaning. Because of the great amount of the information we are provided it and thanks to development of Internet Technologies, text summarization has become an important tool for interpreting text information. Text summarization methods can be classified into extractive and abstractive summarization. An extractive summarization method involves selecting sentences of high rank from the document based on word and sentence features and put them together to generate summary. The importance of the sentences is decided based on statistical and linguistic features of sentences. An abstractive summarization is used to understanding the main concepts in a given document and then expresses those concepts in clear natural language. In this paper, gives comparative study of various text summarization techniques.

Keywords: Text summarization, natural language processing, NLP, statistical approach.

1. INTRODUCTION

A summary can be defined as a text that is produced from one or more texts, that contains a siginificant partion of the information in original text(s), and that is no longer than half of the original text(s) [1]. According to another reference, text summarization is the process of distilling the most important information from a source (or sources) to produce abridged version for a particular user (or users) and task (or tasks)[2].

Automatic Text summarization is a challenging problem and it gaining much significance these days. The automatic summarization means an automatically summarized output is given when an input applied. Although the research on the Automatic Text summarization has been started 1950's in IBM Research Laboratories[3], in the recent years the field of Text Summarization has experienced an exponential growth because of the Internet. It is very difficult to manually summarize large documents of the text because there is a huge amount of information on the Internet. On the other hand, Internet is a rich library that provides more information than is demanded. Therefore, it is crucial for searching relevant document through an huge number of documents available. The target of text summarization is shorten the source text into a diminished version conserving its information content and overall meaning.

Text summarization methods can be classified into *extractive* and *abstactive* summarization. Abstractive Summarization is a method for novel phrasing describing the content of the text which requires heavy machinery from natural language processing, including grammars and lexicons for parsing and generation. *Extractive Summarization* is a method for determining salient text units (typically sentences) by looking at the text unit's lexical and statistical relevance or by matching phrasal patterns [4] Abstraction approaches provide sophisticated summaries and adapt well to high compression rates while extraction approaches are easy to adapt larger sources although the resulting summaries may be incoherent. Extractive Summaries are formulated by extracting key text segments such as sentences or passages from the text, based on statistical analysis of individual or mixed surface level features[5].

There are two different groups of text summarization. *Inducative* and *Informative*. Indicative summarization gives the main idea of the text to the user. The length of this summarization is around 5 percent of the given text. The informative

summarization system gives brief information of the main text. The length of the informative summary is around 20 percent of the given text [6].

Furhermore, summarization methods can be classified according to source which can be *single or multiple document summarization*. In single document summarization, only one document is provided for summary generation. It is a simple and earliest approach for summarization. Extractive and abstractive both summaries methods can be applied on single document summarization. Multi document summarization is also very important part of summarization. More than one information sources are provided for summary generation. Many web based clustering systems like news were inspired from multi document summary. However, task of multiple document summarization technique is more difficult and complex than single document techniques. The real aim is not only to remove redundancy and indentify correct text for summary but also to provide novelty and ensuring that final summary should be coherent and complete in itself. So it was a challenge for them to consider all the documents and relate the summary[7].

2. PREPOROCESSING AND EXTRACTION FEATURES

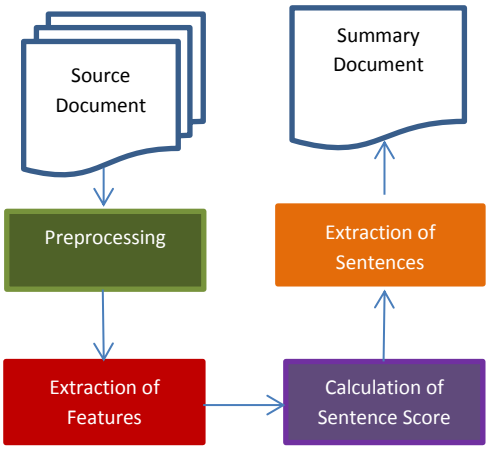
In this paper, we mention about single and multiple document Extractive Summarization techniques because the most of the current automated text summarization systems use extradiction methods. Extractive summarization process can be divided into three phases. First phase is Pre-Processing, second phase is Processing.

In Pre-processing phase, there are many techniques, such as part of speech(POS) tagging, stop word filtering and stemming.

2.1. Part of Speech(POS) Tagging

Part of Speech Tagging is the the process of grouping or specifying the words of text according to speech category such as noun, verbs, adverbs, adjectives etc. The POS tagging process can be implemented by many algorithms such as Hidden Markov Model and Viterbi.

Figure 1: Step by Step Text Summarization



2.2. Stop Word Filtering

Stop words are words which are filtered out before or after processing of text(corpus), it is fully non-objective depends upon the situation. For example, a, an, in, by can be considered as a stop words and filtered from plain text.

2.3. Stemming

Stemming is the process of reducing inflected (or sometimes derived) words to their word stem, base or root form. For example, removing from –ed or –ing from verbs, using singular instead of plural noun, etc. There are many NLP tools which are using stemmer algorithms for executing the stemming.

2.4. Feature Calculation

In order to calculate the score of sentence S, indicating the degree whether it belongs to summary or not, the following feautures are generally used in the score of sentences.

2.4.1.Title Similarity

A sentence has good score if it has maximum number of words similar in the title. The title similarity can be calculated by the number of words in the sentence that occur in the title and the total number of words in the title. The title similarity is defined as;

$$f1 = \frac{S \cap T}{t}$$

In formula;
S = Set of words in sentence
T = Set of words in title
$S \cap T$ = Similar words in sentence and title of document.

2.4.2.Sentence Position

In this feature, a sentence is evaluated by its position in the text. Whether it is the first 5 sentences in the paragraph, sentence position in text gives the importance of the sentences. We consider the first 5 sentences in the paragraph. This feature score is calculated by[6, 12].

f2 = 5/5 for 1st, 4/5 for 2nd, 3/5 for 3rd,
2/5 for 4th, 1/5 for 5th,
0/5 for other sentences.

On the other hand, simple and different formula is used when calculating sentence position value[13].

f2 = 1, if sentence is the beginning sentence in the text,
f2 = 0, if sentence is the middle of paragraphs in the text,
f2 = 1, if sentence is the the last position in the text.

2.4.3.Term Weight(Term frequency)

The total term weight is calculated by computing *tf* and *idf* for document. Here *idf* refers to inverse document frequency which simply tells about whether the term is common or rare across all documents[12,13]. The score of important score w_i of word *i* can be calculated by the traditional *tf.idf* methods.

$$w_i = tf_i.idf_i = tf_i . log\frac{N}{n_i}$$

tf_i = the term frequency of word *i* in the document
N = the total number of sentences
n_i = the number of sentences in which word *i* occurs
This feature can be calculated as follows.

$$f3 = \frac{\sum_{i=1}^{k} W_i(S)}{Max(\sum_{i=1}^{k} W_i(S_i^N))}$$

k= number of words in sentence

2.4.4.Sentence Length

This feature is suitable when eliminating the sentences which are too short such as datelines or author names. The short sentences are not expected to belong to the summary.

$$f4 = \frac{Number\ of\ Words\ in\ Sentence}{Number\ of\ words\ occuring\ in\ longest\ sentence}$$

2.4.5.Thematic Word

This feature is related with domain specific words which occur frequently in a document are probably related topic. The score for this feature is calculated as the ratio of the number of thematic words that occur in the sentence over the maximum summary of thematic words in the sentence.

$$f5 = \frac{Number\ of\ thematic\ word\ in\ S}{Maximum(Number\ of\ Thematic\ Word)}$$

2.4.6.Proper Nouns

Generally, the sentence that involves proper nouns is essential and it is most possibly contained in document summary. The score for this feature is counted as the ratio of the number proper nouns that occur in sentence over the sentence length the important sentence which contains maximum number of proper nouns.

$$f6 = \frac{Number\ of\ Proper\ nouns\ in\ S}{Sentence\ Length}$$

2.4.7.Sentence to Sentence Similarity

The cosine similarity formula gives the similarity between S and each other sentence. The term weight w_i and w_j of term t to n term in sentence S_i and S_j are represented as the vectors.

$$Similarity(S_i, S_j) = \frac{\sum_{t=1}^{N} W_{it} W_{jt}}{\sqrt{\sum_{t=1}^{n} W_{it}^2} \sqrt{\sum_{t=1}^{n} W_{jt}^2}}$$

$$f7 = \frac{\sum Similarity(S_i, S_j)}{Maximum(\sum Similarity(S_i, S_j))}$$

2.4.8.Numerical Data

Generally, the sentence that involves *numerical data* is essential and it is most possibly contained in document summary. The score fort his feature is calculated as the ratio of the number of numerical data that occur in sentence over sentence length. Its score is calculated as**[**12**]:**

$$f8 = \frac{Number\ of\ Numerical\ Data\ in\ S}{Sentence\ Length}$$

3. SUMMARIZATION METHODS

3.1.Query Based and Generic Summarization

In query based text summarization the scoring of the sentences of a given document is based on the frequency counts of words or phrases. Higher scores are given to the sentences containing the query phrases rather than the ones with single query words. The sentences with highest scores are then extracted for the output summary together with their structural context. Portions of text may be extracted from different sections or subsections. The resulting summary is the union of such extracts. In the sentence extraction algorithm, whenever a sentence is selected for the inclusion in the summary, some of the headings in that context are also selected**[**8**]**.

The query based sentence extraction algorithm[**9**] is as follows;

Algorithm:

1: Rank all the sentences according to their score.

2: Add the main title of the document to the summary.

3: Add the first level-1 heading to the summary.

4: While (summary size limit not exceeded)

5: Add the next highest scored sentence.

6: Add the structural context of the sentence: (if any and not already included in the summary)

7: Add the highest level heading above the extracted text (call this heading h).

8: Add the heading before h in the same level.

9: Add the heading after h in the same level.

10: Repeat steps 7, 8 and 9 for the next highest level headings.

11: End while

In generic summaries are not query based[10]. Query based summaries are biased because they do not provide the overall review of the source document. They deal with the user queries only hence not suitable for content overview. To define the category of the document and to describe the main key points of the document generic summaries are required. A best generic summary considers the main topics of the documents and tries to minimize the redundancy as less as possible.

Artificial intelligence (AI), deep learning, fuzzy logic and neural networks represent incredibly exciting and powerful machine learning-based techniques used to text summarization.

3.1.1. Bayesian Classifier

Given a training set of documents with hand-selected document extracts, develop a classification function that estimates the probability a given sentence is included in an extract. New extracts can then be generated by ranking sentences according to this probability and selecting a user-specified number of the top scoring ones.

For each sentence *s* we compute the probability it will be included in a summary *S* given the *k* features f_j; *j = 1…k*, which can be expressed using Bayes' rule as follows[22]:

$$P(s \in S \mid f1, f2, \ldots, fk) = \frac{P(f1,f2,\ldots,fk \mid s \epsilon S)\ P(s \epsilon S)}{P(f1,f2,\ldots,fk)}$$

Assuming statistical independence of the features:

$$P(s \in S \mid f1, f2, \ldots, fk) = \frac{\prod_{j=1}^{k} P(f_j \mid s \in S)\ P(s \in S)}{\prod_{j=1}^{k} P(fj)}$$

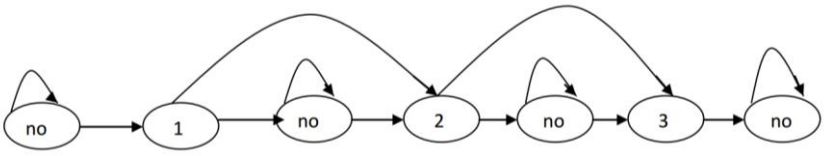
3.1.2. Hidden Markov Model

Hidden Markov Model(HMM)[25] is an another method to extract a sentence from a document. In contrast to a naive Bayesian approach, the Hidden Markov model has fewer assumptions of independence. In particular, the HMM does not assume that the probability that sentence i is in the summary is independent of whether sentence i-1 is in the summary.

The main idea is using a sequential model to account for local dependencies between sentences. In HMM Model, three features were used:

- position of the sentence in the document,
- number of terms in the sentence,
- likeliness of the sentence terms given the document terms.

Figure 2: Hidden Markov Model to Extract Three Summary Sentences from a Document.



The HMM was structured as follows: it contained 2s + 1 states, alternating between s summary states and s+1 nonsummary states. The authors allowed *"hesitation"* only in nonsummary states and *"skipping next state"* only in summary states. Fig. 2. shows an example HMM with 7 nodes, corresponding to s = 3. Using the *TREC* dataset as training corpus, the authors obtained the maximum-likelihood estimate for each transition probability, forming the transition matrix estimate, whose element (*i ,j*) is the empirical probability of transitioning from state *i* to *j*. Associated with each state *i* was an output function, (*O*) = Pr(*O* | state *i*) where *O* is an observed vector of features. They made a simplifying assumption that the features are multivariate normal. The output function for each state was thus estimated by using the training data to compute the maximum likelihood estimate of its mean and covariance matrix. They estimated 2s+1 means, but assumed that all of the output functions shared a common covariance matrix. Evaluation was done by comparing with human generated extracts.

3.1.3. Neural Networks Based Text Summarization

Artificial Neural Networks(ANN) are one of the most popular and powerful classes of machine learning algorithms. ANN is used to generate summaries of arbitrary length news articles. A neural network is trained on a corpus of articles. The neural network is then modified, through fusion to produce a summary of the most ranked sentences of the article. Through feature fusion, the network discovers the importance of various features used to determine the Summary-worthiness of each sentence[16].

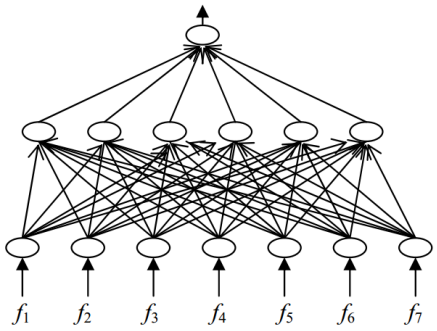
ANN has training and testing phases. In training phase, sentences in several test paragraphs where each sentence is identified as to whether it should be added in the summary or not. This is carried out by a human reader. The neural

network learns the patterns which are inherent in sentences that should be added in the summary and those that should not be added.

The Neural Network has three layered feed-forward structure. It consists of seven input layer neurons, six hidden layer neurons and one output layer neurons. Each sentence is represented as a vector [*f1, f2,…, f7*] which consists of 7 features. In this network, few features are different from before, but most of them are same. The features are selected according to position of document or position of the sentence.

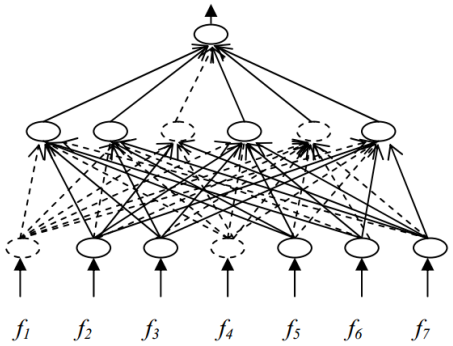
f1 = Paragraph follows title (Paragraph Position)
f2 = Paragraph location in document
f3 = Sentence location paragraph
f4 = First sentence in paragraph
f5 = Sentence Length
f6 = Number of thematic words in sentence
f7 = Number of title words in sentence

Figure 3: The Neural Network after Training



Text Summarization process consists of three phases: training, feature fusion and sentence selection. The first step involves training a neural network to recognize the type of sentences that should be added in the summary. In the second step, reduce the neural network and collapses the hidden layer unit activations into discrete values with frequencies. The third step, sentence selection, uses the modified neural network to filter the text and to select only the highly ranked sentences[17].

Figure 4: The neural Network after Pruning



The hidden layer activation values for each children layer neuron are clustered utilizing an adaptive clustering technique. Each cluster is identified by its centroid and frequency. The activation value of each hidden layer neuron is replaced by the centroid of the cluster, which the activation value belongs to. This corresponds to collapsing the effect of common features. The combination of these two steps corresponds to generalizing the effects of features, as a whole, and providing control parameters for sentence ranking.

Another technique[18] which is using Neural Network for text summarizing add “Numerical Data Feture” to feature input list so their network uses 8 input neurons. After finding high ranked summary sentences by neural network we feed these sentences to rhetorical structure to find the discourse structure from that and find rhetorical relation in sentences which may help in finding better summary sentences, which further might be used to form better summary.

Rhetorical Structure Theory[19] using rhetorical relations provide a methodical way for an analyst to analyse the text. An analysis is usually constructed by reading the text & building a tree using the relations. The example given below is a title and summary, the original text, broke down into units having numbers, is:

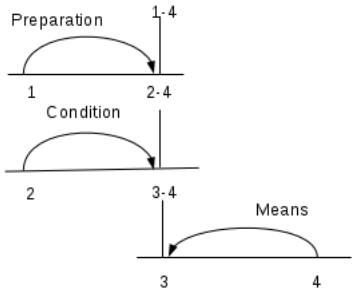
1. The Perception of Apparent Motion

2. When the motion of an intermittently seen object is ambiguous

3. the visual system resolves confusion

4.by applying some tricks that reflect a bulletin knowledge of properties of the physical world.

Figure 5: Rhetorical Relations



In the figure 3 number 1, 2, 3, 4 displaying the correspond units as explained above. 4th unit and 3rd unit forming a relation neans. 4th unit is the important part of this relation. So it is known as nucleus of the relation and 3rd unit is known as satellite of the relation. Similarly 2nd unit to 3rd and 4th unit is forming relation condition. spans may be composed of two or more units[18].

3.1.4. Fuzzy Logic Based Text Summarization

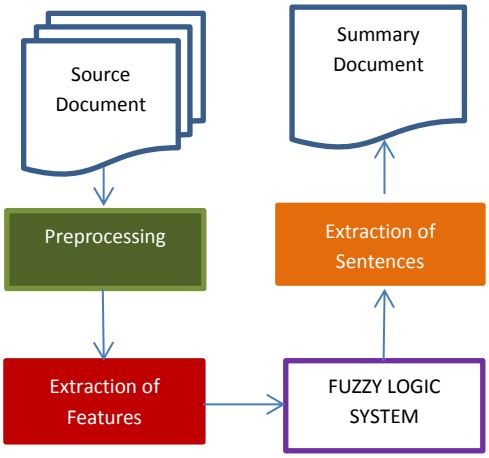
In this method, Fuzzy Logic rule and fuzzy logic set are used to extract the important sentences based on their features. Fuzzy logic techniques provide decision-support and expert systems with strong reasoning capabilities. Fuzzy logic set proposed by Zadeh[29] is a mathematical tool dealing with uncertainity, imprecision, vagueness and ambiguity. A few studies were done in text summarization. Witte and Bergle[30] present a fuzzy-theory based approach to coreference resolution and its application to text summarization. Automatic determination of coreference between noun phrases is fraught with uncertainty. They show how fuzzy sets can be used to design a new coreference algorithm which captures this uncertainty in an explicit way and allows us to define varying degrees of coreference. Patil and Kulkarni[31] also use Fuzzy Logic to score sentences after feauture selection and pre-processing step. They use eight features for text summarization: title word, sentence length, sentence position, numerical data, thematic words, sentence to sentence similarity, term weight and Proper Nouns.

The system involve in following steps:

- Perused the source archive into the system,
- In preprocessing step, the system extracts the individual sentences of the original documents. After, separate the input document into individual words. Next, remove stop words. The last step for preprocessing is word stemming;
- each sentence is related in vector of eight features that mentioned in Section II, whose values are obtained from the content of the sentence;
- the features are calculated to obtain the sentence score base on fuzzy logic method shows in Figure 6.;
- a set of highest score sentences are extracted as document summary based on the compression rate.

Fuzzy Logic System includes four components: Fuzzifier, Inference Engine, Defuzzifier and Fuzzy Knowledge Base. In the fuzzifier, crisp inputs are translated into linguistic values using a membership function to be used to the input linguistic variables. After fuzzification, the inference engine refers to the rule base containing fuzzy IFTHEN rules to derive the linguistic values.

Figure 6: Text Summarization Flowchart with Fuzzy Logic



In the last step, the output linguistic variables from the inference are converted to the final crisp values by the defuzzifier using membership function for representing the final sentence score[12]. In order to implement text summarization based on fuzzy logic;

Thus each sentence is associated with 8 feature vector. Using all the 8 feature scores, the score for each sentence are derived using fuzzy logic method. The fuzzy logic method uses the fuzzy rules and triangular membership function .The fuzzy rules are in the form of IF-THEN .The triangular membership function fuzzifies each score into one of 3 values that is LOW, MEDIUM & HIGH. Then we apply fuzzy rules to determine whether sentence is unimportant, average or important. This is also known as defuzzification. For example IF (F1is H) and (F2 is M) and (F3 is H) and (F4 is M) and (F5 is M) and (F6 is M) and (F7 is H) and (F8 is H) THEN (sentence is important).

In sentence selection phase, all the sentences of a document are ranked in a descending order based on their scores. Top n sentences of highest score are extracted as document summary based on compression rate. Finally the sentences in summary are arranged in the order they occur in the original document.

4.EVALUATION

Evaluation methods are necessary to compare results between text summarization methods. Two main criterion for evaluating the proficiency of a system is precision and recall which are used for specifying the similarity between the summary which is generated by the system versus the one generated by human. The terms are defined in the following equation[23].

$$Precision = \frac{Correct}{Correct+Wrong}$$

$$Recall = \frac{Correct}{Correct+Missed}$$

In these formula;

Correct = the number of sentences that are the same in both summary which are generated by the system and human.

Wrong = the number of sentences presented in summary and produced by system but is not included in the Summary generated by human.

Precision = the number of suitable sentences which are extracted by the system and Recall specify the number of suitable sentences that the summarization system missed.

$$Compression\ Ratio(CR) = \frac{Length(S)}{Length(T)}$$

$$Retention\ Ratio(RR) = \frac{Information\ in\ S}{Information\ in\ R}$$

The ROUGE, a set of metrics called Recall-Oriented Understudy for Gisting Evaluation, evaluation toolkit [28] that has become standards of automatic evaluation of summaries. It compares the summaries generated by the program with the human-generated (gold standard) summaries. For comparison, it uses n-gram statistics[27].

5. CONCLUSION

Owing to fast growth of technology and use of Internet, there is information overload. This problem can be solved if there are strong text summarizers which produces a summary of document to help user. In the last decades, Automatic Text Summarization is a popular research area and attracts a lot of attentions from different science branches. In this paper, we discuss type of summarization methods which might be used in a system to produce summary. Text summarization methods can be classified into *extractive* and *abstactive* summarization. An extractive summary is selection of important sentences from the original text. Abstractive Summariy is a method for novel phrasing describing the content of the text which requires heavy machinery from natural language processing, including grammars and lexicons for parsing and generation. Especially, this paper is focussing on extractive summarization methods.

Extractive bases summarization methods Bayesian Classifier, Hidden Markov Model, Neural Networks and Fuzzy Logic have to an extent, succeeded in making an effective summary of a document.

REFERENCES

[1] E. H. Hovy, Automated Text Summarization. The Oxford Handbook of Computational Linguistics, Chapter 32, pages 583-598. Oxford University Press, 2005.

[2] I. Mani, D. House, G. Klein, The TIPSTER SUMMAC Text Summarization Evaluation. In Proceedings of EACL, 1999.

[3] H. P. Luhn, "The automatic creation of literature abstracts", IBM Journal of Research and Development, vol. 2, pp. 159-165, 1958.

[4] U. Hahn, & I. Mani, "The challenges of automatic summarization" IEEE-Computer, 33(11), 29–36, 2000.

[5] V. Gupta, G.S. Lehal, "A survey of Text Summarization Extractive Techniques", Journal of Emerging Technologies in Web Intelligence, Vol. 2, No.3, August 2010.

[6] S.A. Babar, P. D. Patil, "Improving Performance of Text Summarization", International Conference on Information and Communication Technologies, ICICT, 2014.

[7] Sherry, P. Bhatia, "A Survey to Automatic Summarization Techniques", International Journal of Engineering Research and General Science Volume 3, Issue 5, September-October, 2015.

[8] S. Karmakar, T. Lad, H. Chothani, A Review Paper on Extractive Techniques of Text Summarization, International Research Journal of Computer Science (IRJCS) Issue 1, Volume 2, 2015.

[9] F. C. Pembe and T. Güngör, "Automated Query-biased and Structure-preserving Text Summarization on Web Documents," in Proceedings of the International Symposium on Innovations in Intelligent Systems and Applications, İstanbul, June 2007.

[10] G. Yihong, X. Liu. "Generic text summarization using relevance measure and latent semantic analysis." Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval, ACM, 2001.

[11] S. Alfayoumy, J. Thoppil, A Survey of Unstructured Text Summarization Techniques, International Journal of Advanced Computer Science and Applications, Vol. 5, No. 4, 2014.

[12] L. Suanmali, N. Salim, M. S. Binwahlan, "Fuzzy Logic Based Method for Improving Text Summarization", International Journal of Computer Science and Information Security, Vol. 2, No. 1, 2009.

[13] P.Priya, G. and K Duraiswamy, "An Approach for Text Summarization Using Deep Learning Algorithm", Journal of Computer Science, 1-9, 2014.

[14] Amy J.C. Trappey, Charles V. Trappey, "An R&D knowledge management method for patent document summarization" Industrial Management & Data Systems, vol.108. pp. 245-257. 2008.

[15] D. Ravichandran, P. Pantel, and E.H. Hovy, "Randomized Algorithms and NLP: Using Locality Sensitive Hash Functions for High Speed Noun Clustering", Proceedings of the conference of the Association for Computational Linguistics (ACL), 2005.

[16] Khosrow Kaikhah, "Automatic Text Summarization with Neural Networks", in Proceedings of Second International Conference on intelligent systems, IEEE, 40-44, Texas, USA, June 2004.

[17] Khosrow Kaikhah, "Text Summarization using Neural Networks", Departments of Faculty Publications Computer Science, Texas State University, 2004.

[18] Sarda A.T., Kulkarni A.R., "Text Summarization using Neural Networks and Rhetorical Structure Theory", International Journal of Advanced Research in Computer and Communication Engineering, 2015.

[19] W. C. Mann, S. Thompson, Rhetorical Structure Theory: Toward a functional theory of text organization. Text, 8 (3), 1988.

[20] F.Kyoomarsi, H. Khosravi, E. Eslami and P.H. Dehkordy, "Optimizing Text Summarization Based on Fuzzy Logic", In proceedings of Seventh IEEE/ACIS International Conference on Computer and Information Science, IEEE, University of Shahid Bahonar Kerman, UK, 347-352, 2008.

[21] S.A.Babar, S.A.Thorat, "Improving Text Summarization using Fuzzy Logic & Latent Semantic Analysis",International Journal of Innovative Research in Advanced Engineering (IJIRAE), 2014.

[22] J. Kupiec, J. Pedersen, and F. Chen, "A trainable document summarizer", In Proceedings of the 18th ACMSIGIR Conference, pages 68-73, 1995.

[23] E. Lahari, D.V.N.S. Kumar, M.Ubale, "A Comprehensive Survey on Feature Extraction in Text Summarization", E Padma Lahari et al , Int.J.Computer Technology & Applications,Vol 5 (1),248-256, 2014.

[24] P.D. Patil, P.M. Mane, "An Overall Survey of Extractive Based Automatic Text Summarization Methods", International Journal of Science and Research (IJSR) 2014.

[25] J. Conroy, D.P. O'leary, "Text Summarization via Hidden Markov Models and Pivoted QR Matrix Decomposition", Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval 2001.

[26] C. Cığır, M. Kutlu, İ. Çiçekli, "Generic Text Summarization for Turkish", The 24th International Symposium on Computer and Information Sciences, ISCIS 2009, 14-16 September 2009.

[27] C.Y. Lin, E.H. Hovy, "Automatic evaluation of summaries using n-gram co-occurrence statistics", Proceedings of HLT-NAACL-2003, Edmenton, Canada, 2003.

[28] C.Y. Lin, "ROUGE: A Package for Automatic Evaluation of Summaries" In Proceedings of Workshop on Text Summarization of ACL, .Spain. 2004.

[29] L. Zadeh, "Fuzzy sets. Information Control" vol. 8, pp.338–353.1965.

[30] R. Witte and S. Bergler, "Fuzzy coreference resolution for summarization" In Proceedings of 2003 International Symposium on Reference Resolution and Its Applications to Question Answering and Summarization (ARQAS) 2003.

[31] P.D. Patil, N.J. Kalkurni, "Text Summarization Using Fuzzy Logic", International Journal of Innovative Research in Advanced Engineering (IJIRAE) Volume 1 Issue 3, 2014.

[32] S. Gholamrezazadeh, M. Salehi, B. Gholamzadeh "A Comprehensive Survey on Text Summarization Systems", Computer Science and its Applications, 2009.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

THE CONNECTION CALCINATION AND SULPHATION WITH THE THERMAL BEHAVIOR OF CHALCOPYRITE ORE CONCENTRATE

DOI: 10.17261/Pressacademia.2016.592
PAP-WCTIE-V.5-2017(30)-p.214-223

Meltem Kizilca[1], Mehmet Copur[2]
[1]Department of Chemical Engineering, Ataturk University, Erzurum, Turkey
[2]Department of Chemical Engineering, Bursa Technical University, Bursa, Turkey. mkizilca@atauni.edu.tr

ABSTRACT

The thermal behavior of chalcopyrite ore concentrate was used thermogravimetric analysis (TG) coupled by differential thermal analysis (DTA). The thermal analysis were done up to 900°C from 0°C at different heating rates (2, 5, 10, 15 and 20 °C /min) under air atmosphere. The ore was calcined at 30, 60, 90 min and 350, 400, 450, 600, and 800°C and investigated sulphation proportions. The products of obtained calcination were characterized by XRD, SEM/EDS, FTIR analysis. Based on the conclusions obtained was confirmed the thermal behavior of the ore.

Keywords: Chalcopyrite, thermal analysis, calcination, sulphation, characterization

1. INTRODUCTION

Copper, is one of the most important among industrial metals, has a wide use area in industry such as cooling and refrigeration tubing, electrical equipment, jewelry, electrical power cables, heat exchangers. It is present usually as metallic in small quantities, and mostly as sulphated minerals, oxidized minerals and complex minerals in nature. Chalcopyrite ($CuFeS_2$), a copper sulphide minerals, is the most exploited as primary sources of copper. But the dissolution of copper sulphide minerals is quite difficult (Parasad et al., 1999; Mikhlin et al., 2004). Such minerals need to be converted into sulphate in order to be converted into soluble structure. The transformation to the sulfate structure by preheating of the minerals in the sulfur structure is called roasting. There are three common roasting methods: chlorination roasting, oxidation roasting and sulfation roasting. In the chlorination roasting; for the formation of metal chlorides to metal sulphides are a type of roasting made by the addition of some chlorinating agent, chlorine gas or some chlorides. This methods is a threat to the environment due to released gas (HCl, Cl_2) from the furnace. In oxidation roasting made of roasting at high temperature copper will cause a loss by creating copper ferrite. The sulfation roasting is an important route to recover copper from a copper sulphide minerals because during this process water-soluble sulphate and acid-soluble oxide compounds of copper are dissolved in the environment. At the same time, sulfur gas released during the roasting process ores with sulphide content are both stored and converted to sulfuric acid. It is known that due to its hazardous nature, SO_2 is not environmentally friendly and harmful to health and difficult to handle. (Copur et al., 2015; Zhao et al., 2015). The ore is oxidized after roasting process. Chemical conversion reactions during roasting are utilized

to change the connection of the element to be recovered in metallurgy. Creating a technical possibility or going to a more economical work requires the application of such reactions. It is quite difficult to predict the reaction and mechanism of chalcopyrite ore, which is characterized as a complex mineral. This is mostly because of a few other reaction taking place with one another interlocutory products. Define an exclusive temperature range for a demand product is very important for the process. Additionally, take into account the complexity of the process, define the reaction sequence may be determined by using TG/ DTG (thermal analysis techniques). These techniques are very beneficial for comprehension and planning of phase transformations and the selectivity (Parasad et al., 1999).

In the present work, the contribution of sulfation roasting for the recovery of copper from ore a copper sulfide content was investigated. As a contribution to the better knowledge of this process, reaction and mechanism of roasting process of chalcopyrite ore were determined the thermal analysis techniques (TG/ DTG). Supported with XRD-SEM-EDS-FTIR results of roasted products at selected temperatures.

2. LITERATURE REVIEW

Sargsyan and Hovhannisyan (2010a) were determined as the optimal roasting temperature (650± 10^{o}C) and the roasting time (30 min) with sulfating roasting chalcopyrite concentrates of distinct pyrite contents for the recovery of copper. Sargsyan and Hovhannisyan (2010b) determined using thermal analysis techniques (TG/ DTG) the phase transformations forms in chalcopyrite cuprum concentrate. They reported that Cu-Fe-S transform into water-soluble sulphate phases at temperatures up to 676^{o}C and into oxides when the temperature is greater than 676^{o}C. Jain et al. (1982) inspected the optimum temperature of roasting $CuFeS_2$ ore for the optimum recovery of copper from $CuFeS_2$ ore in H_2SO4 solutions in a mullite tube furnace. They determined that they reached a maximum value of the copper concentration in owing to the formation of soluble sulfates at a temperature of 550^{o}C and afterwards begins decrescenting due to the formation of insoluble ferrites. Parasad et al. (1988) studied using thermal analysis techniques (TG/ DTG) the phase transformations forms with of some additives such as Fe_2O_3, Na_2SO4, FeSO4 to chalcopyrite ore. They display that these additives contributed to copper concentrate of the sulphation roasting at 773K. They determined that there was more sulphation with Fe_2O_3, Na_2SO4, and FeSO4 than without any addition. Gülfen and Aydın, 2010 were studied calcination of chalcopyrite ore with Fe_2O_3 as sulfation agent and without any agent addition. They were found that Fe_2O_3 addition contributed more copper sulphate conversion and as the optimal roasting temperature (873K) and the roasting time (60 min). Zhao et al. (2015) were studied sulfation roasting with Na_2SO_3 as sulfation agent and without any agent addition to recover copper from the $CuFeS_2$ concentrate. They were determined as the optimal roasting temperature (500˚C - 550˚C) and the roasting time (1200 min).

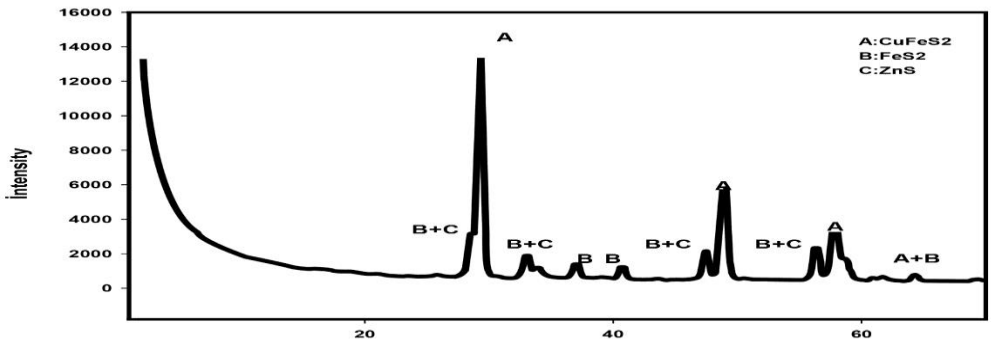
3. METHODOLOGY AND DATA

The chalcopyrite ore was from the Rize Çayeli of Turkey. The chemical analysis of the concentrate for 63 µm sieve is dedicated in Table 1.Mineralogical analysis of the concentrate was implemented using XRD (X-ray diffraction) (Bruker D8 Advance) powder diffraction. The results are shown in Fig. 1. $CuFeS_2$ and FeS_2 are a large amount of and other compounds such as ZnS, Cu_2S are smaller amounts of in the sample. In addition, SEM (energy dispersive X-ray spectroscopy) and EDS (energy dispersive X-ray spectroscopy), which combines surface microstruture and elemental compound of the chalcopyrite ore was defined by using X-MaxN attached to a JEOL 6010 LV Oxford Instruments in Fig. 2. The EDS analysis shows O, Fe, Cu, Al, Zn, Si, S, and Pb. (The percentage ratios of the elements: 1.21, 23.77, 26.54, 1.55, 0.17, 0.22, 22.34 and Pb- 24.20, respectively).

Table 1: The Chemical Analysis of Chalcopyrite Ore

Element	Cu	Fe	S	Zn	Pb
Percent (%)	24.88	29.46	34.23	2.19	1.19

Figure 1: XRD Recordings of Chalcopyrite Concentrate



2Theta /degree

The thermal analysis were performed using NETZSCH STA 409 PC Luxx instrument. The ore samples of 15-20 mg were used and Measurements carried out in air atmosphere at 2, 5, 10, 15 and 20 K/min heating rates at 298-1173 K. FTIR spectrum was recorded to identify the chemical structure of the ore using Perkin Elmer spectrometer 100 in Fig. 3. The peak 480 cm^{-1} and 600 cm^{-1} are FeS bands at, 612 cm^{-1} is ZnS bands, the peaks at 1390–1290 cm^{-1} are asymmetric stretching of (S–O) and the peaks at 1190–1120 cm^{-1} are symmetric stretching of (S–O) in Fig. 3.

Figure 2: SEM Image of Chalcopyrite Core and EDS Pattern (inset)

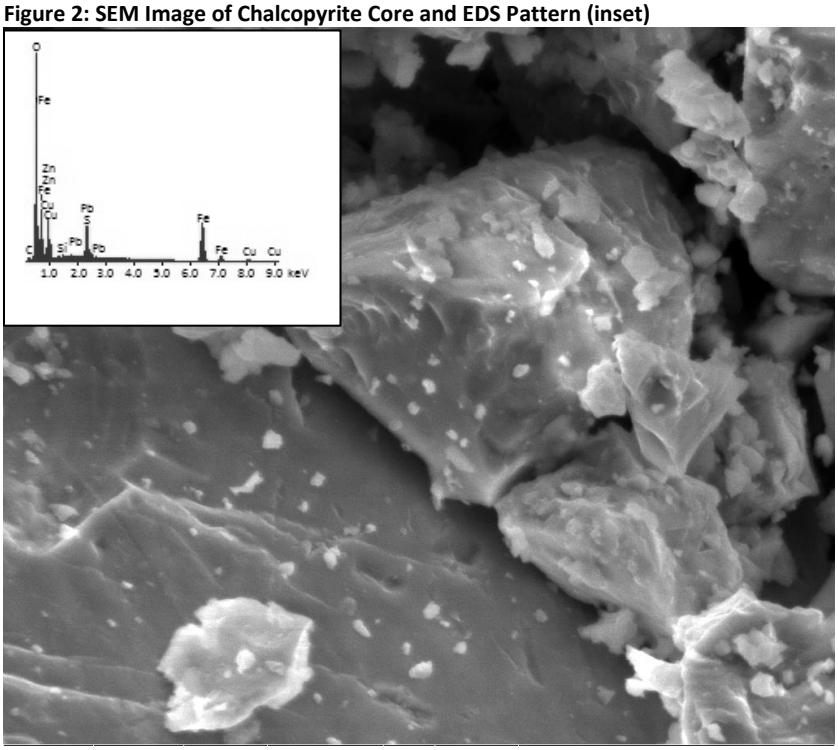
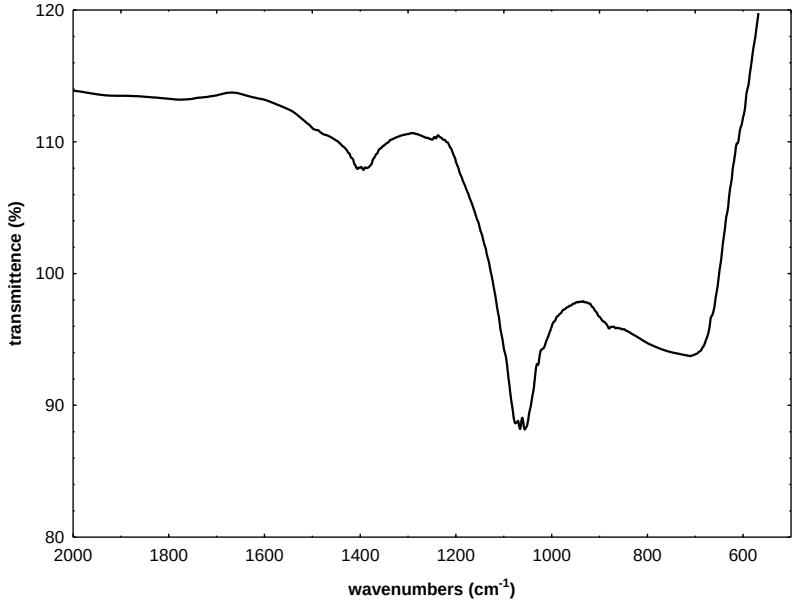


Figure 3: FT-IR Spectrum of Chalcopyrite Ore



The roasting process was performed on a porcelain capsules at 30, 60, 90 min and 350, 400, 450, 600, and 800^{o}C in an ash oven at a rate of 60 cm^{3} min^{-1} under the flow of air. Afterwards these roasted chalcopyrite patterns were weighed (in Table 2). XRD-SEM-EDS-FTIR results were applied to the samples obtained in the roasting experiments at selected temperatures.

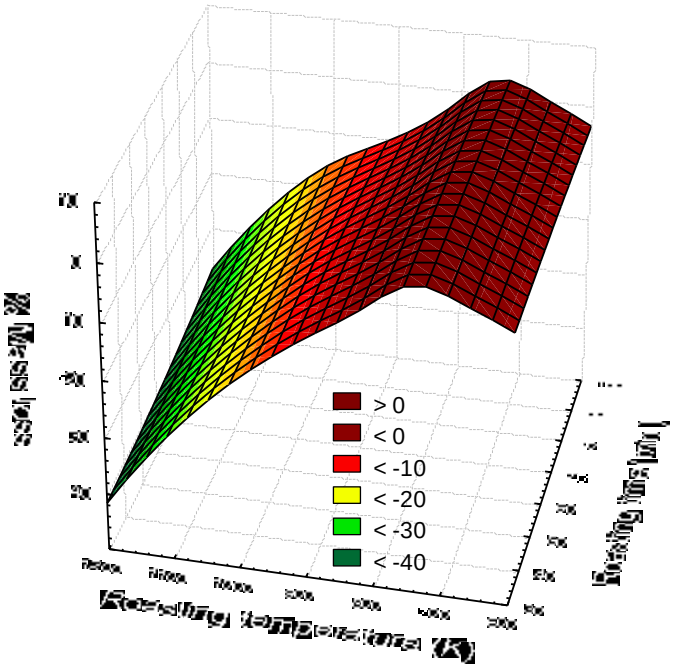
Table 2: Mass Changes in Roasting of Chalcopyrite Ore

Roasting time (min.)	Roasting temperature (g)				
	623K	673K	723K	873K	1073K
30	-1,2	-0,9	4,13	-6,42	-24,33
60	-1	-0,93	4,34	-6,12	-24,09
90	-1,6	-1,3	4,44	-7,42	-22,45

3. FINDINGS AND DISCUSSIONS

The roasting process was carried 350, 400, 450, 600, and 800^{o}C. The differences in mass amounts depending on the effect of roasting temperature are shown in the Fig. 4. In the roasting of chalcopyrite ore, weight changes that occur during calcination are shown in Table 2 and Fig. 4. Based on weight changes resulting from roasting the chalcopyrite ore, an increase was observed up to 723K. At 723K, a weight change of 4% was not significant.

Figure 4: Effect of Roasting Temperature



At 873K, a substantial change in mass did not occur, but there was a considerable mass loss (24%) observed at 1073K. Above 723K, a weight increase occurred from the formation of sulfate compounds; at temperatures much higher than 723K, the weight decreased because of the degradation of sulfate compounds and decomposition of sulphide components in the chalcopyrite ore. Because the sulfate compounds formed readily dissolve in a leach environment. The optimal roasting temperature was achieved at 500 °C. Also, thermal studies (DTA-TGA) were utilized to observe the transformation phases of chalcopyrite to copper sulfate. The overall graphs of DTA-TGA curves of chalcopyrite ore are shown in Figs. 5 and 6. With increasing heating rates from 2 K/min to 20 K/min, DTA curves recorded at different heating rates have similar patterns although peaks usually change at elevated temperatures.

Figure 5: TG Analysis of Chalcopyrite Ore

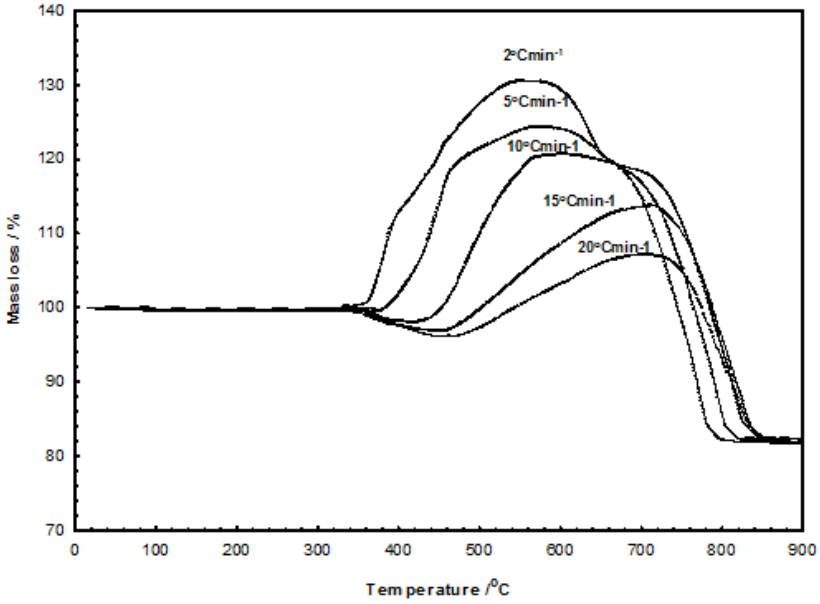
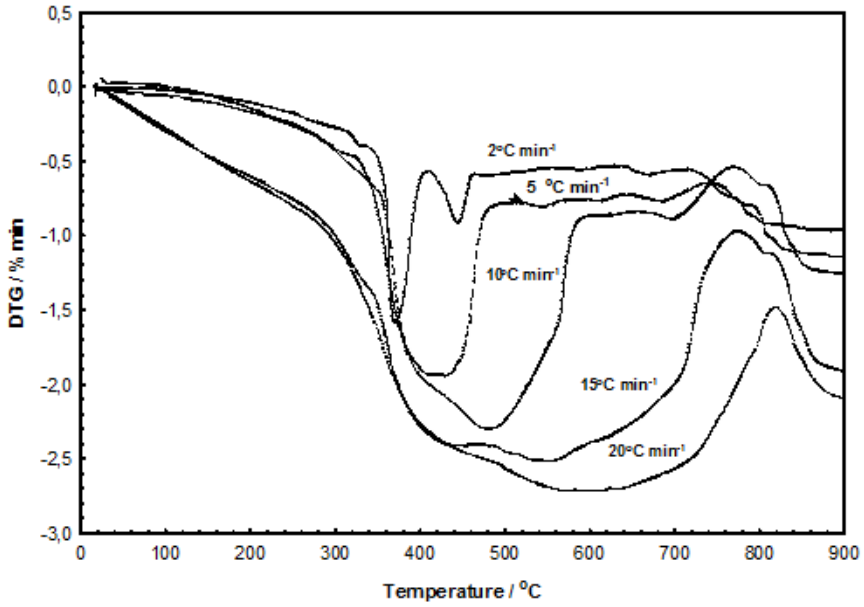


Figure 6: DTG Analysis of Chalcopyrite Ore



Furthermore, each DTA curve has two parts. The first part includes three exothermic peaks corresponding to oxidation of sulfides at low temperatures, whereas the second part includes endothermic peaks corresponding to the decomposition of previously formed sulfate and copper oxide-sulfate compounds. Figs. 4 and 5 reveal that the oxidation of chalcopyrite on its surface layer occurred at temperatures between 320 °C and 360 °C, depending on heating rates. A substantial increase is not observed in TGA curves at these temperatures, whereas small-scaled exothermic peaks are seen in DTA curves. At this temperature, copper matte a mixture of Cu_2S and FeS is obtained. The reaction, as stated by Sokic et al. (2008), can be represented by the following equation:

$2CuFeS_2+O_2(g) \rightarrow Cu_2S+2FeS+SO_{2(g)}$ (1)

Resulting from intense oxidation of the inner layer of chalcopyrite and simultaneous degradation of pyrite in the sample at temperatures between 370 °C and 400 °C, a high-intensity exothermic peak is clearly observed in the DTA curves. At this stage of the process, the formation of elemental sulfides decreased the amount of the original material.

$2CuFeS_2 + O_{2(g)} \rightarrow Cu_2S + 2FeS + SO_{2(g)}$) (2)

$2FeS_2+2O_{2(g)} \rightarrow FeS + 2SO_{2(g)}$ (3)

$2CuFeS_2 \rightarrow Cu_2S + 2FeS + S$ (4)

In subsequent heating operations, the weight of the sample increases and a moderate-intensity exothermic peak forms in DTA curves in the temperature range 420 °C to 520 °C. At this stage of the roasting operations, iron sulfate, copper sulfate, and copper oxide-sulfate compounds form. With further increases in temperature, based on the TGA curves, the weight of the sample decreases by small amounts resulting from the degradation of iron sulfates; in DTA curves, unclear endothermic peaks form in the temperature range 525 °C to 580 °C.

$2FeSO_4 \rightarrow Fe_2O_3 + SO_2(g) + SO_{3(g)}$ (8)

Sulfates start to degrade at temperatures above 700 °C.

$CuSO_4 \rightarrow CuO + SO_{3(g)}$ (9)

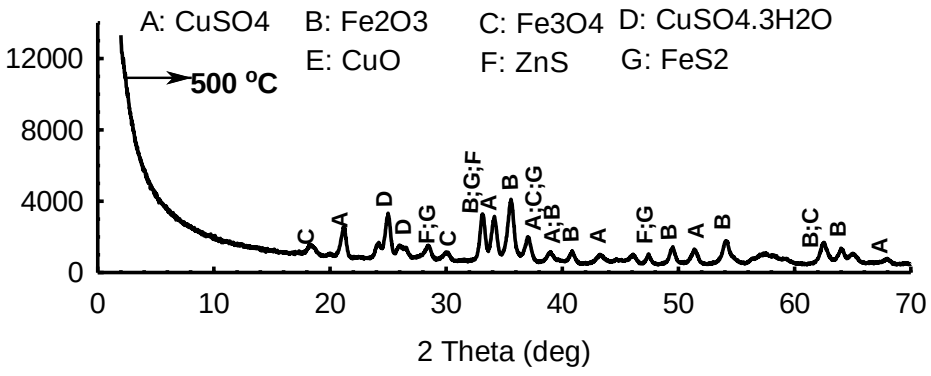
$CuO{\cdot}CuSO_4 \rightarrow 2CuO + SO_{3(g)}$ (10)

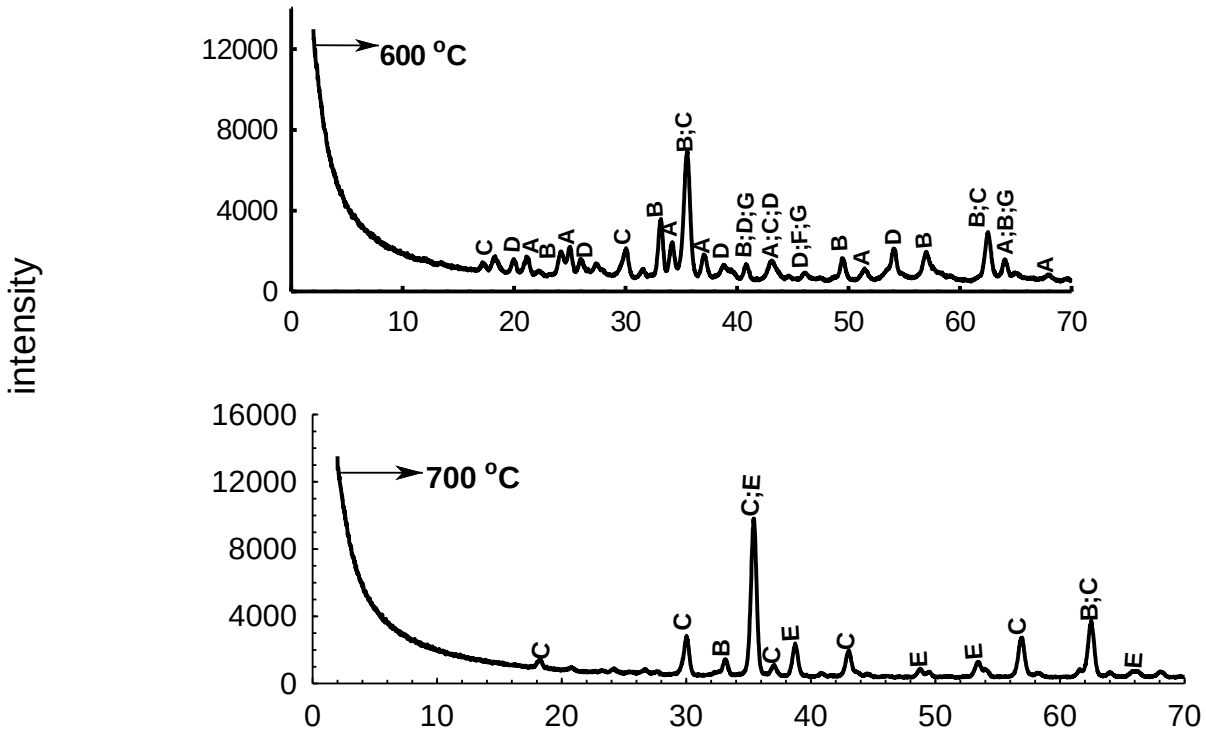
At these temperatures, weight loss is observed in the TGA curves, and two endothermic peaks are seen in the DTA curves. At a heating rate of 2 K/min, peaks representing the degradation of copper sulfate and copper oxysulfate are observed at temperatures of 725 °C and 710 °C. As heating rates are increased, copper sulfate and copper oxide sulfate begin to increase simultaneously at some stages; therefore, the peaks overlap. Prasad *et al*. (1999) indicated that during the roasting of chalcopyrite in air, the oxidation mechanism first results in formation of Cu_2S and FeS. Subsequently, Cu_2S transforms into Cu_2O, followed by the $CuO.CuSO_4$ transformation. Finally, $CuO.CuSO_4$ transforms into $CuSO_4$. According to these authors, the oxidation process of chalcopyrite follows the path shown below.

$CuFeS_2 \rightarrow Cu_2S \rightarrow Cu_2O \rightarrow CuO \rightarrow CuO.CuSO_4 \rightarrow CuSO_4$ (11)

XRD recordings of the patterns of roasted chalcopyrite at 500, 600 and 700^{o}C are showed in Fig. 7. In Fig. 7 is shown XRD recordings of patterns of roasted chalcopyrite at 500, 600 and 700^{o}C. Remarkably, the existence of iron and copper in different types of sulfates and sulfides can be seen: CuSO4, FeSO4, FeS at 500^{o}C and 600^{o}C.

Figure 7: XRD Recordings of Chalcopyrite Ore for 500, 600 and 700^{o}C





Furthermore, SEM (the surface microstructure) - EDS(the contents of the elements) techniques indicated that there are O, S, Fe, and Cu in Fig.8 (The percentage ratios of the elements: 36.33, 10.37, 25.70 and 23,12, respectively) and O, S, Fe and Cu in Fig. 9 (The percentage ratios of the elements: 20.76, 3.55, 49.63 and 11.90, respectively). It involving Fe_2O_3 and CuO as final products at 700^{o}C and SEM - EDS techniques indicated that there are O, S, Fe and Cu in Fig. 10 (The percentage ratios of the elements: 22.97, 0.32, 23.72 and 42.46, respectively).

a) **b)**

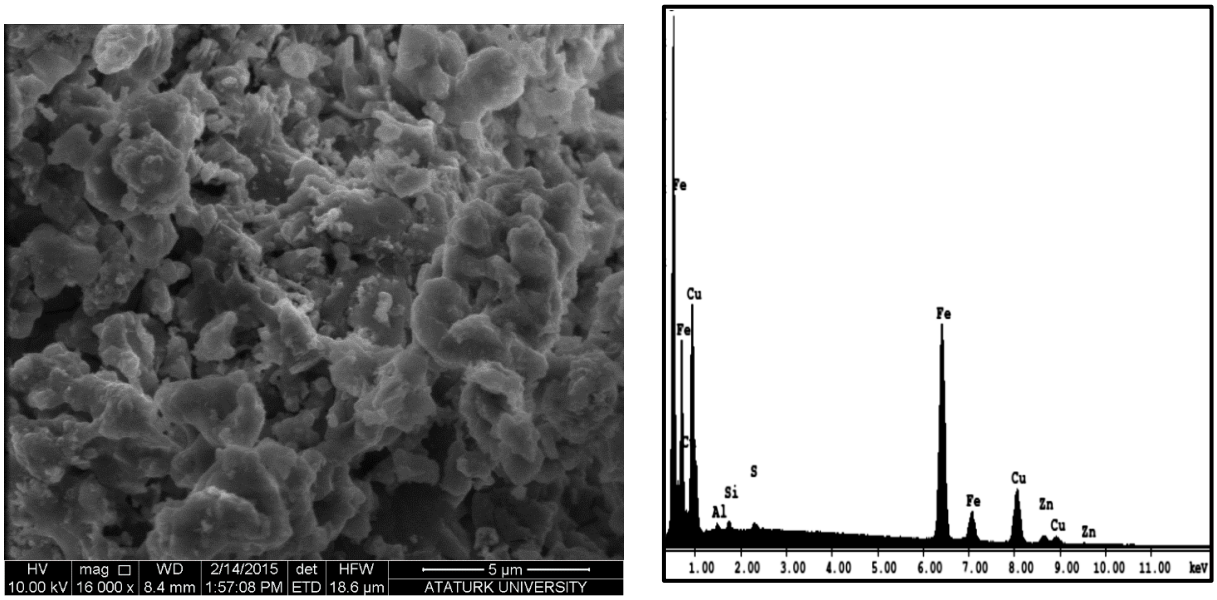


Figure 8: SEM Images of Chalcopyrite Ore for 500°C (a) and EDS Pattern (b)

a) **b)**

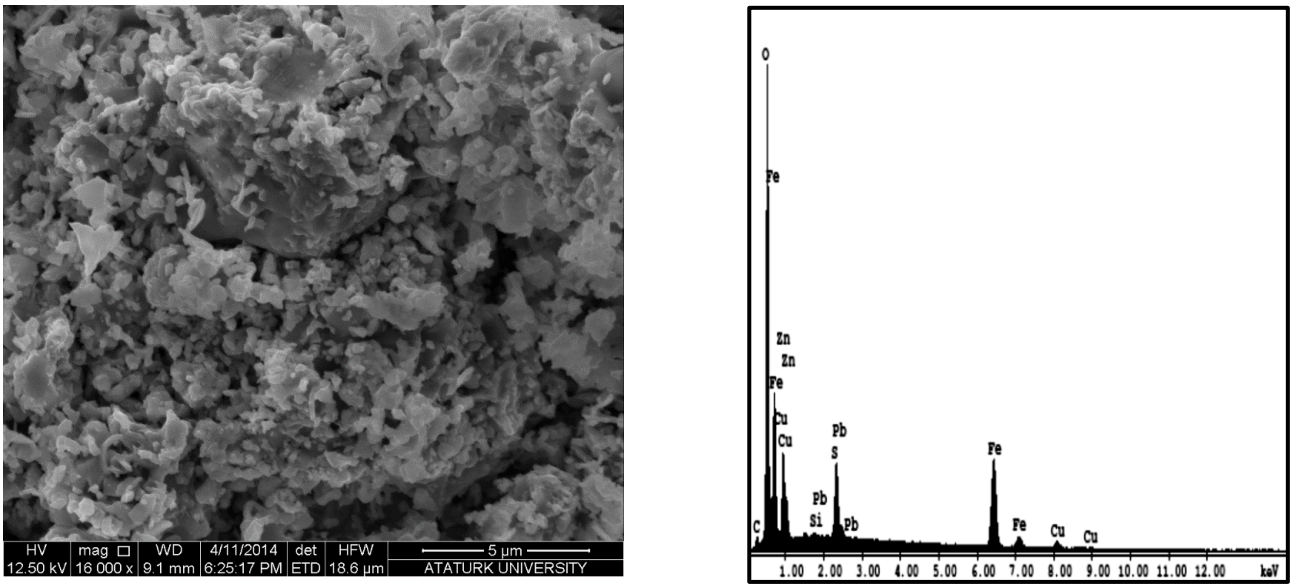


Figure 9: SEM Images of Chalcopyrite Ore for 600°C (a) and EDS Pattern (b)

a) b)

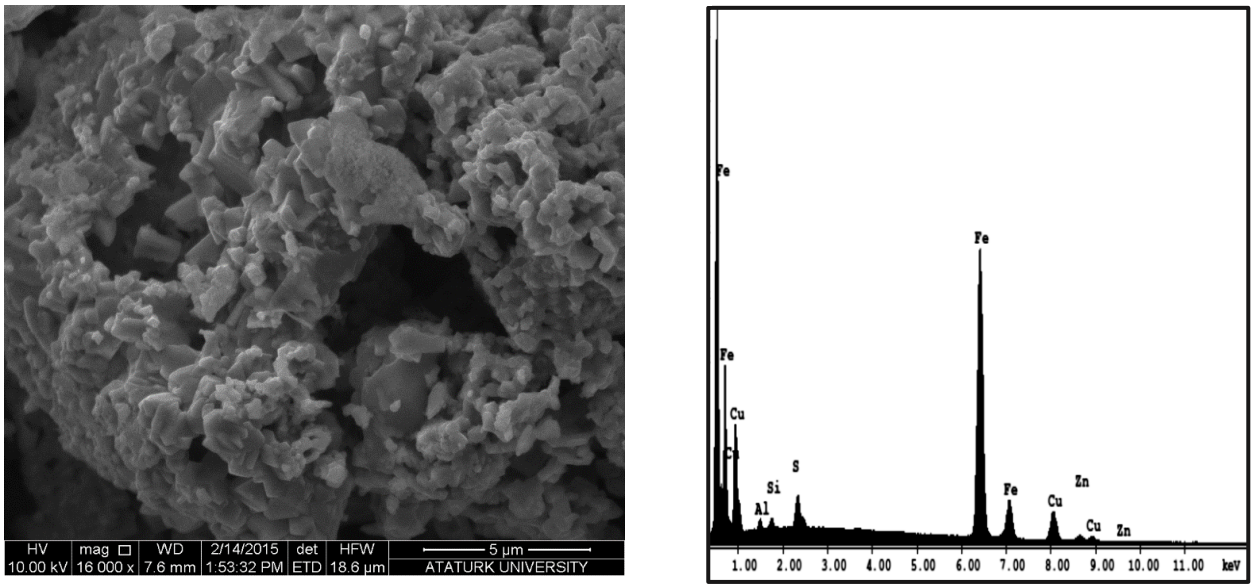
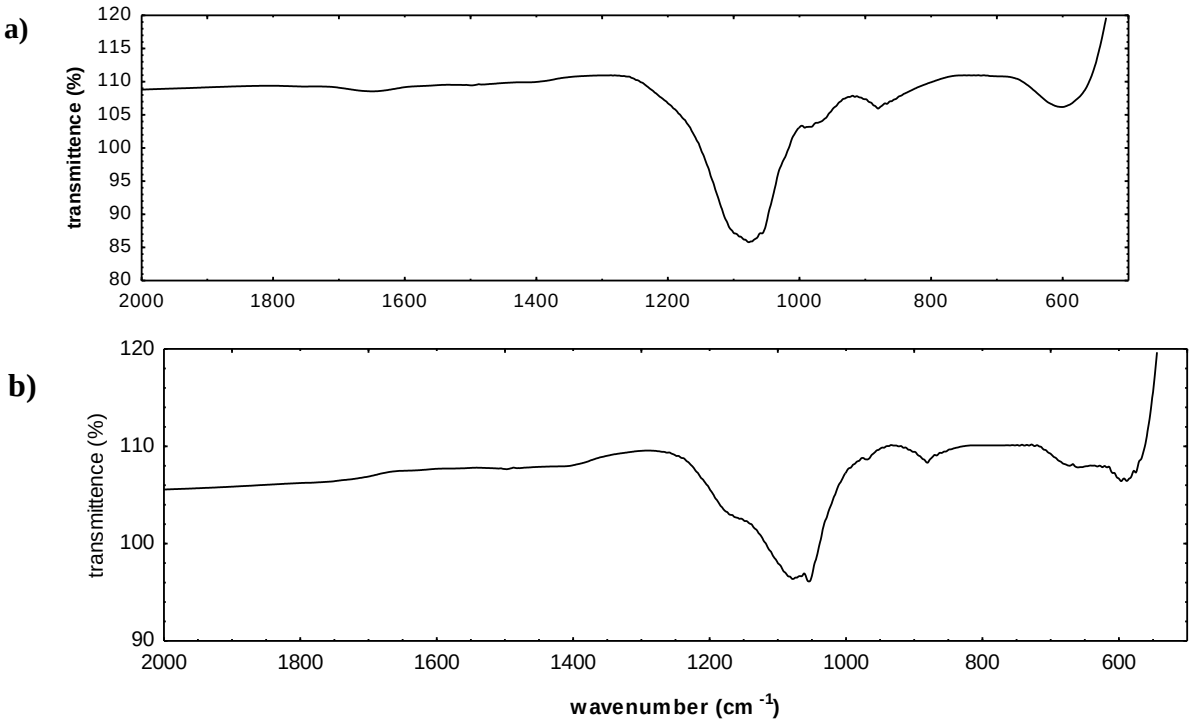
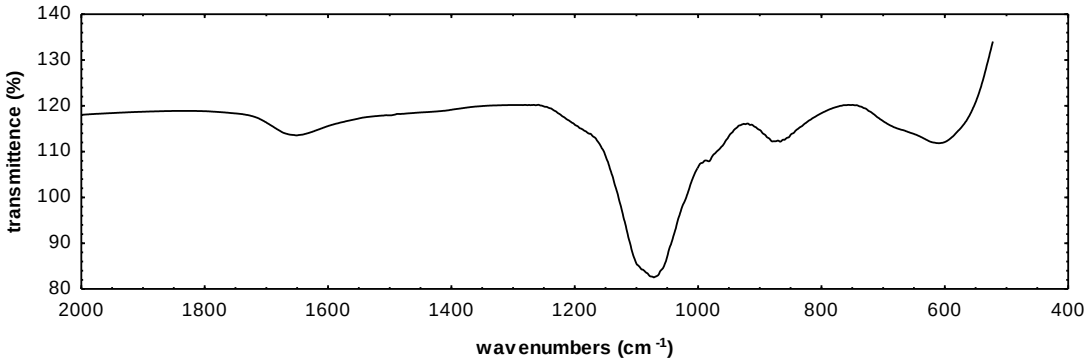


Figure 10: SEM Images of Chalcopyrite Ore for 700°C (a) and EDS Pattern (b)

Besides, FT-IR results confirmed these results. (Fig.11). The strong peak (S-O bands) at 1190–1120 cm^{-1} and 1390–1290 cm^{-1} confirmed the presence of sulfates and sulfides. Also, the peaks at 525cm^{-1} and 580 cm^{-1} demonstrated the presence of CuO at 462cm^{-1} and 560 cm^{-1} the presence of Fe2O3 (Rahman et. al., 2011; Suan et. al., 2011). The sulfation results chalcopyrite samples calcined at 400-900 °C are shown in Fig.11.





c)

Figure 11: FT-IR spectrum of chalcopyrite ore for 500°C (a), 600°C (b) and 700°C (c)

5. CONCLUSION

This work was determined as the optimal roasting temperature (500°C) and with sulfating roasting chalcopyrite concentrates for the recovery of copper. The phase transformations forms in chalcopyrite ore determined using thermal analysis techniques (TG/ DTG). The present study reported that Cu-Fe-S transform into water-soluble sulphate phases at temperatures up to 700 and into oxides when the temperature is greater than 700°C. The products of obtained roasted at selected temperatures for reaction and mechanism of roasting process were both characterized and supported by XRD, SEM/EDS, FTIR analysis

REFERENCES

Copur, M., Kızılca, M., Kocakerim, M.M. 2015, "Determination of the optimum conditions for copper leaching from chalcopyrite concentrate ore using taguchi method". Chemical Engineering Communications, vol. 202, no 7, 927-935.

Gulfen, M., Aydın, A.O. 2010, "Dissolution of copper from a primary chalcopyrite ore calcined with and without Fe_2O_3 in sulphuric acid solution", Indian Journal of Chemical Technology, vol. 17, 145-149.

Mikhlin, Y.L., Y.V. Tomashevich, Asanov, I. P., Okotrub, A.V., Varnek, V.A., Vyalikh, D.V. (2004). Spectroscopic and electrochemical characterization of the surface layers of chalcopyrite (CuFeS2) reacted in acidic solutions. Applied Surface Science, vol. 225,395–409.

Jain, R.C., Madhavan, S., Acharjee, D.K., Gupta Sen P. 1982," Optimum performance data for the leaching of roasted chalcopyrite concentrate in a jet contactor", The Canadian Journal of Chemical Engineering, vol. 60, 516-521.

Parasad, S., Pandey, B.D., 1998," Alternative processes for treatment of chalcopyrite -a review". Mineral Engineering, vol 11, no.8, 763-781.

Parasad,S., Pandey,B.D., 1999. Thermoanalytical studies on copper-iron sulphides. Journal of Thermal Analysis and Calorimetry, vol. 58, 625-637.

Sargsyan, L.E., Hovhannisyan, A. M. 2010a, "Actived sulfating roasting of the chalcopyrite concentrate for sulfuric-acid leaching", **Russian** Journal of **Non**-**Ferrous** Metals, vol. 51, 386-388.

Sargsyan, L.E., Hovhannisyan, A. M. 2010b, "Investigation of chalcopyrite cuprum concentrate coasting by thermogravimetric and differential-thermal analysis", Metallurgical and Minning Industry, vol.2, no. 3, 225-229.

Sokić,M., Ilić, I., Živković, D., Vućkovik N. 2008, "Investigation of mechanism and kinetics of chalcopyrite concentrate oxidation process". Metalurgija, vol. 47, 109-113.

Zhao Y., Hou Y., Cui Y., Liang H., Li L. 2015," Recovery of Copper from Copper Sulfide Concentrate by Sulfation Roasting". International Journal of Nonferrous Metallurgy, vol. 4, 9-13.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

NEW FORESIGHT GENERATION AND FRAMEWORK OF FORESIGHT

DOI: 10.17261/Pressacademia.2017.593
PAP-WCTIE-V.5-2017(31)-p.224-233

Nurdan Yuksel[1,2], Hasan Cifci[1,3] Serhat Cakir[4]
[1]Department of Science and Technology Policies Studies, Middle East Technical University
[3]Turkish Air Force
[4]Department of Physics, Middle East Technical University, Çankaya, Ankara, Turkey
[2]ASELSAN Inc.,Yenimahalle, Ankara, Turkey
*Corresponding Author: nurdan.yuksel@metu.edu.tr

ABSTRACT
The article starts with various definitions of technology foresight from the literature and new definition is given by the authors. Prominent definitions of foresight are analyzed and given according to their different elements in the literature. By assessing the evolution of foresight as a term and its five generations through time, the foresight of sixth generation is set forth. Based on the literature analysis, foresight generations are grouped into four main categories. Foresight frameworks are analyzed and a generic foresight functional model with nine consecutive phases named 'FORESIGHT' is suggested. Functions in the FORESIGHT model are matched with the phases of main foresight frameworks in the literature based on their actions and artifacts within specific phases. It is like a guide for practitioners of which activities are carried out in foresight studies.

Keywords: Foresight, foresight framework, foresight generations.

1. INTRODUCTION

Humans have a capacity to systematize the estimation of future events [1]. Voros [2] defines the Three Laws of Future based on the suggestion of Amara [3] as it is not predetermined, not predictable and the selection of today affects its outcomes. Since the future itself is a complex subject coming from its uncertain nature, futures oriented studies are difficult to determine and realize. Any statement related with future cannot be confirmed at the time when they are built [4].

It is a challenge or almost impossible to predict the future. Following quotes show the difficulty of predicting the future developments and prove the need a new approach to the future related matters:

- "Everything that can be invented has been invented", Charles H. Duell, Commissioner, US Office of Patents, 1899.

- "Airplanes are interesting toys, but of no military value", Marshall Foch, 1911.

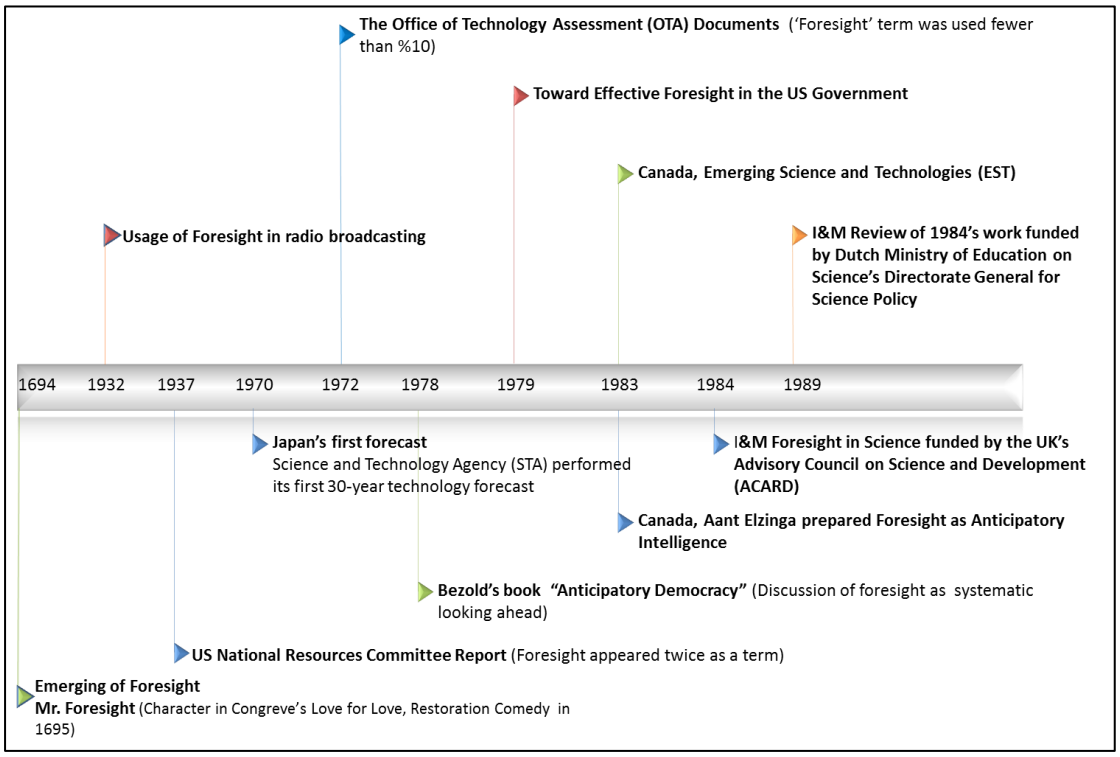
- "There is no reason for an individual to have a computer at home", Ken Olsen, Chairman of Digital Equipment Corporation, 1978.

Unlike future as a term, which invokes "looking ahead" [5], futures studies also include "looking back" [6] and looking present. Since *futures* is very wide subject, there is different terminology about it as *futures research*, "futures studies", sometimes *futures analysis*, *futurism*, *futuristics*, or even *futurology* [2]. According to Sardar [5], suffix of –logy in futurology which shows the meaning of certainty and objectivity is not suitable for futures having ambiguity. He mentions about the linguistic and pronunciation difficulty of futuristics and lethality of futurism's logic which implies that science will provide knowing the future. While Voros [2] prefers to use *Futures* since the term gives the plurality of different futures, Sardar [5] finds the best suitable term as *Future Studies* with the same reasoning of Voros. Futures are evaluated as extensive professional and academic area continuing

to develop with its methods, tools and research [7]. Futures studies are multi-disciplinary, trans-disciplinary and un-disciplinary[1] [5] activities for the future.

When looking at the literature, forecast and foresight as future studies come into prominence. While forecast is mainly passive attitude for the single future, foresight is systematic active attitudes of today for diverse futures. Foresight has wider aim such as involvement of broad participation, invoking the policy makers with its outputs [8] and providing greatest economic and social benefits [9]. Forecast is a set of techniques which are used to convert inputs to outputs whereas foresight is mainly a process of which techniques are used to create common understanding, commitment and networking [10]. According to Loveridge [11] foresight is structurally based on the "uncomfortable marriage of well-structured and tested information to its counterpart subjective opinion". The timeline in Figure 1 shows the history of foresight as a term.

Figure 1: Timeline of Foresight as a Term[2]



Hines and Gold [12] researched about the foresight profession and assessed its status against the common criteria of profession. Since they concluded that it corresponded the requirements of capacity, field and discipline more or less but it was not able to emerge as profession by 2013. Together with the knowledge economy in globalized world, it is undeniable that science and technology are the major determinants of the future. So, technology foresight becomes the most important type for future studies. Martin [13] defines Technology Foresight (TF) as "A process involved in systematically attempting to look into the longer-term future of science, technology, the economy and society with the aim of identifying the areas of strategic research and the emerging generic technologies likely to yield the greatest economic and social benefits". TF also creates networks between the stakeholders by invoking awareness about opportunities, threads, social and economic needs, benefits or damages. TF studies contribute to knowledge creation by building *Ba*[3] which is shared extent of the relations emerged from interactions between individuals, groups and environment [14] and filling the *Ma*[4] which represents the spaces between boundaries of Ba such as disconnection of stakeholders or physical distance between them.

TF is just one of the analyses of Technology Future Studies (TFA) or Future oriented Technology Analyses (FTA). Different forms of FTA, which covers emerging technology conditions and types,

[1] It consciously rejects the status and state of a discipline while being a fully-fledged systematic mode of critical inquiry.
[2] Miles [22] was used in building timeline of foresight.
[3] Japanese term come up with by philosopher Kitaro Nishida and developed by Shimizu
[4] Japanese word which can be roughly translated as "gap", "space", "pause" or "the space between two structural parts [43].

technology development paths and their impacts systematically, such as technology forecasting, technology intelligence, technology assessment, technology roadmapping exist concurrently with TF [15].

In Table 1, elements of various foresight definitions in the literature are shown. Elements are listed based on the number of references given. In accordance with the model suggested in this study, a new foresight definition is set forth as follows:

"*Foresight is a systematic and multidisciplinary process with proper methodology combinations for identifying technological, economic and social areas to prioritize investments and research in order to determine medium or long term future strategies by using all level of resources from organizational to international.*"

2.FORESIGHT GENERATIONS

Foresight studies can be divided into different generations based on the scope, goals and methods involved within foresight processes. In the literature, there are different approaches which can be categorized under 4 main groups towards generational models of foresight as shown in Figure 2.

Linstone [16] classifies foresight generations into three groups based on certain society.

I. First generation (ca. 1800 - Industrial Society): Foresight activities were mainly based on technology forecasting which dated back to industrial era with Taylor's scientific management. Quantitative, semi-quantitative and qualitative tools were developed in order to forecast and predict technological developments.

II. Second generation (ca. 1970 - Information Society): Information technology has noteworthy impact in effecting simultaneous centralization and decentralization or globalization and localization. Computer capabilities have been used for forecasting. It has become possible to use vast majority of databases and computer power in order to reach and process huge volumes of data.

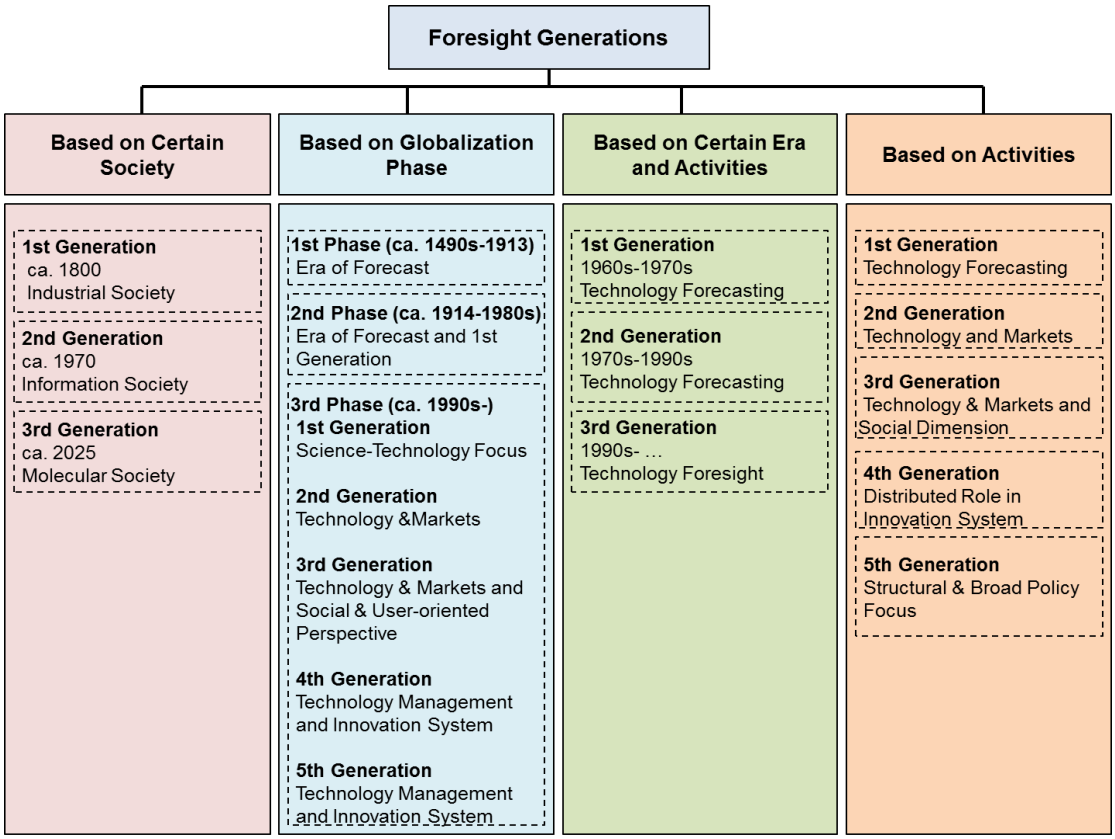
III. Third generation (ca. 2025 - Molecular Society): This era is founded on molecular level with nanotechnology, biotechnology, and materials science stepping forward. It is also referred to as the Nano-biotechnology era and the Micro and Nanotechnology era (MNT). Nanotechnology is expected to lead to a vast increase in computational power and intelligence which are exploited in foresight methods and activities. As one of the main triggers of the foresight studies, globalization plays an important role in emerging of foresight generations.

Table 1: Elements of Various Foresight Definitions

	Martin (1995)	Martin & Johnston (1999)	Slaughter (1995)	Georghiou (1996)	Georghiou, Harper, Keenan, Miles, & Popper (2008)	Barré (2001)	Miles & Keenan (2002)	Miles (2010)	Harper (2003)	Voros (2005)	Keenan et al. (2007)	Popper (2008)	Popper (2011)	Schmidt (2015)	Conway (2015)	Yüksel & Çifci & Cakir (2017)
Systematic studies/process	X			X		X	X						X		X	X
Looking at medium and long term future	X				X	X		X				X				X
Participatory, collective, networking process					X	X	X		X		X		X			
Building visions						X	X		X		X					
Gathering intelligence						X	X									
Learning process											X	X				
Mobilized actions						X	X									
Joining key agents of change and knowledge sources						X						X				
Assessing scientific and technological developments				X												
Looking into science, technology, economy and society	X															
Identifying technologies which have impact to create economic and social benefits, industrial competitiveness, wealth creation and quality of life	X															

Identifying critical technologies to the success of the country or industry								X								
Cognitive capacity															X	
Setting Priorities					X											X
Changing research system and administrative culture					X											
Identifying and examining weak signals														X		
Activities anticipating, recommending and transforming (ART) technological, economic, environmental, political, social and ethical (TEEPSE) futures.													X			
Actively engaging key stakeholders													X			
Prospective and policy-oriented process													X			
Context for predictive analysis and planning														X		
Promoting trans-disciplinarity research												X				
Underpinning areas of strategic research	X															
Universal Human Capacity			X													
Wiring up the national/regional system of innovation		X														
Strategic Thinking										X						
Common Ownership of Strategies									X							
Incubation of Strategic Approaches									X							
Using organizational, regional, national and international resources																X
Proper methodology combinations																X
Identifying technological, economic and social areas to prioritize investments and research																X

Figure 2: Main Approaches to Foresight Generations



Jemala [29] classifies five foresight generations according to their corresponding three globalization phases by pointing out first foresight studies of related generations according to date:

I. First globalization phase: The phase dating back to Christopher Columbus in 1490s and lasting to 1913 with the effect of "Laissez Faire" thought was generalized as Era of Forecast since there were not so many systematic approaches and strategic plans in general. Most of the future studies were being performed just because of inherent human reaction to its fear of undetermined future.

II.Second globalization phase: The phase from 1914 to 1980s mainly affected from world wars and information and communication technologies was named as Era of Forecast until 1956 and First Generation of Foresight from 1956 to 1990. First foresight generation focused on science and technology of post war conditions. Although 'forecasting' term continued to be used to 1980s, in 1956 China performed forecast having many features of foresight. Systematic manner with participatory methods and strategic planning to be beneficial in societies especially with the efforts of futurology and technology professionals [29], e.g. 30-year technology forecast of Japanese Science and Technology Agency (STA) in 1971, started to be dominant as First Generation of Foresight.

III.Third globalization phase: The phase from 1999 to 2009 based on the inequalities between the countries arising from the global trade and finance included all five generations of foresight. Second foresight generation was mainly triggered from industry and markets. It is performed with the participation of different stakeholders and science-technology relation with social-economic development was the main focus. Third foresight generation, approximately in 2000, expands the stakeholders to social actors to create foresight culture and to do studies on different areas like environment, education, ethics etc. apart from just science and technology. Fourth and fifth generations focus on systematic approach with system view more and technology and innovation management become dominant.

When it comes to mixed approach –i.e. generations based on certain era and activities- Reger [30] determines three generations based on technology foresight process and assigns certain time intervals per generation [31]:

I. First generation (from 1960s to beginning of 1970s): Technology foresight was carried out within organizations and focused on forecasting and prediction. Foresight was conducted as a sub-task of project planning.

II. Second generation (from 1970s to beginning of 1990s): Foresight was focusing on technology data and remained a forecasting action. Specialized and separated foresight units were responsible for foresight within organizations.

III. Third generation (from 1990s): Technology foresight became an important element of strategic management and decision-making. Not only technologic but also economic, social, environmental and legal trends were taken into account. Activities were extended to encompass networks of organizations.

In terms of foresight generations based on activities, Georghiou categorizes generations into five different groups based on the stakeholders involved and activities carried out [32]. Foresight generations are [33]:

I. First generation: Merely technology forecasting carried out by experts on the future studies. Main goals are making accurate predictions and expressing results to non-expert audiences.

II. Second generation: Foresight activities which began in 1990s in order to combine technology and markets. Industry and academia are bringing together to establish science and business networks.

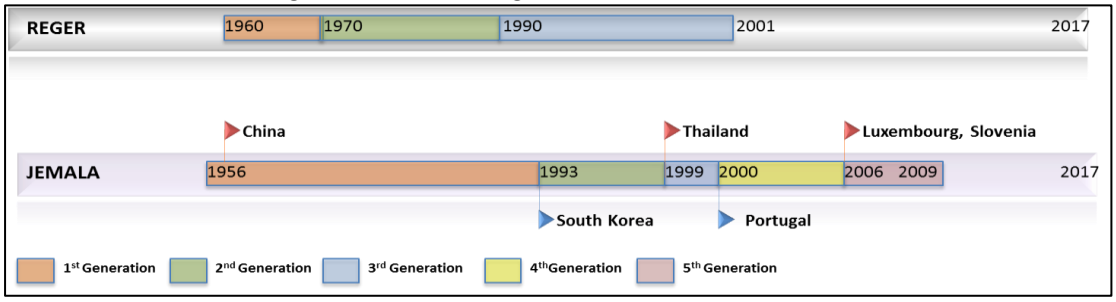
III. Third generation: While keeping technology and markets, social dimension is added to the studies. Industry, academia, government and civil society are the main actors. More stakeholders are involved in evaluation.

IV. Fourth generation: Foresight activities are moved away from integrated programs to different actors within science and innovation system. Multiple organizations carry out foresight studies according to their needs with some level of coordination.

V. Fifth generation: Structural activities focusing on complex policy issues of global science, technology and innovation systems.

The foresight generations in the literature classified based on the certain dates by authors Reger [30] and Jemala [29] (with the addition of the first performing countries) is depicted in the Figure 3. While Reger [30] assorted foresight as three generations up to 2001, Jemala [29] grouped them into five generations until 2009. The generations' dates were not coherent in their studies because of different arguments used in classification and terminological confusion between forecast and foresight. An individual foresight program can have elements of more than one generation. These generations are not sequential approaches and should be thought as "concurrent, overlapping and reflexive" [34].

Figure 3: Dates of Foresight Generations in the Literature



When foresight generations are evaluated for their methodologies approaches; first generation, which was dominated by forecasting mostly around 1960s, used mostly quantitative methods although some qualitative methods like scenario planning emerge in this generation [35]. For the second generation, in which foresight activities started to increase, participative methods together with qualitative methods became widespread. Third generation, which included social dimension, used the analysis like SWOT (Strengths, Weaknesses, Opportunities, and Threats) more. Fourth and fifth generations, which focus on global science, technology and innovation systems with different actors and complex structures, use mixed and internet based methods more prevalently.

Saritas [36] suggests new foresight approach which comprehends complex, inter-dependent, inter-connected nature and behaviors of the foresight system. As sixth foresight generation the assessment of the authors is added to Table 2, which is adapted from Georghiou et al. [19] and [37], and generations are shown with respect to its concentration/dimensions, participating actors, economic rationales and principle.

Table 2: Foresight Generations

Foresight Generations	Concentration Dimensions	Participating Actors	Economic Rationales	Principle
First	Technology	Technology Experts, Professional Futurists	Economic Planning	To follow the disciplinary taxonomies of science-engineering
Second	Technology-Markets	Academics, Industrial Researchers and Managers	Market Failure	To provide a bridge between industrial/service sector and economy
Third	Technology-Markets-Social Dimension	More Social Stakeholders (NGOs, Consumer Groups	System Failure (socio-economic system)	To solve socio-economic problems
Fourth	Science-Innovation System	More Participators of National Policy Exercise	Bridging institutions in socio-economic system	To build own structures in terms of object of analysis
Fifth	Global science-technology management-innovation systems	More experts, stakeholders and professionals with foresighting skills	Bridging institutions in socio-economic system	To build own structures in terms of object of analysis
Sixth	Netocracy[5], biotechnology, more values and ethics in chaordic social dimension	Netocrats[6], Netizens[7] (crowd-sourced from a much wider range of constituencies than the usual experts), Futurists, Futurizens[8]	Blurring the roles of consumers and producers in economy	To co-create by combining altogether the desirable visions of stakeholders with evidence from big data

[5] "It refers to a perceived global upper-class that bases its power on a technological advantage and networking skills" [44].
[6] "Those who can harness networks of information and master new forms of communication will control finance and legislation, forming the new business and government elites" [45]
[7] "People online who actively contribute towards the development of the Net (a Net Citizen)" [46].
[8] "Prescient, future-facing, forward-looking people, citizens of future [47] and futurium (The European Commission Project collecting people's contributions in futures policy constitution) community.

3.FORESIGHT FRAMEWORKS AND PRINCIPLES

The main dimensions of foresight, which are context, scope, artifacts, participating actors and methods, have evolved throughout the foresight journey within its generations until today. So, many scholars have been contemplating about foresight's theoretical and conceptual framework to catch this dimensional change. It has also brought the need for common vocabulary for foresight thus ontology has become important. "Ontology is a formal, explicit specification of a shared conceptualization" [38] and it provides clarification of knowledge [39]. Within this perspective, in the foresight literature, frameworks play important role to create common terminology. Accordingly, some of well-known foresight frameworks are analyzed in the following paragraphs.

Martin [13] analyses foresight as a process having three phases**.** In the first phase named *pre-foresight,* the decision taking to perform the process and preparation are the main steps. Second phase is *foresight* containing comprehensive design, strategic analysis of its possible utilities, consensus about the best *promising options,* determining the scientific priorities by disseminating the outputs. In the last phase which is *post-foresight*, implementation of foresight in policy-making and resources allocation take place [39].

Voros [40] defined generic foresight framework by treating it as a process. His framework is based on modification of Horton's three phased process, combining it with Mintzberg's strategy view and Slaughter's strategic foresight approach. While knowledge gathering and strategic intelligence are assessed as *inputs*; analysis which is related to the question of 'what seems to be happening?', interpretation based on the question of 'what is really happening?' and prospection related to the 'what might happen?' are handled in the *foresight* phase. *Outputs* are comprised of strategic choices resulted from broad understanding. By separating outputs and action, he added *strategy* as a last step that contains actions for decision-making in strategy development and strategy planning.

Hines & Bishop [41] viewed foresight as a process requiring six involvements of leaders and organizations defined as *framing, scanning, forecasting, visioning, planning* and *action.* Framing contains problem identification and determination of its and its solutions' expenses. Scanning is the step of being aware of the trends and related information. While forecasting is creating the possible futures, visioning is selection of desired future and determining the position of organization for that future. Planning is generating the road to the desired future and acting is bringing plans to actions on ongoing base.

Saritas [36] proposed a Systemic Foresight Methodology (SFM) comprised of *Intelligence*, *Imagination*, *Integration*, *Interpretation*, *Intervention* and *Impact*, all within *Interaction*, to generate and conduct foresight operations. All of the operations are performed in dynamic, repetitive and evolutionary way and considered in the extent of how systems in the range of foresight are handled. Foresight is implemented with the thought of the internal (e.g. management, processes, motivation, politics, culture, power, skills) and external context (e.g. STEEPV[9]) in the systemic approach [36]. The systemic approach assesses the foresight with its interrelated and interdependent elements and their complex connections [36]. Within this approach, *intelligence* phase contains gathering data, analysis of trends and changes to determine the scope and content of the foresight study. *Imagination* involves creating and interconnecting new ideas, scenarios. *Integration* is the phase of identifying priorities and creating *agreed model of future*. While the models and visions are transformed to strategies to be carried out in *interpretation* phase, plans and policies are required to be constituted in order to put them into practice in *intervention* phase. The foresight is reviewed and lessons learned are taken to renew it or for better applications in *impact* analysis. All of the phases in SFM are conducted with the participation of stakeholders in some forms interaction.

Schultz [42] took foresight as an integrated process and defined it as a five key activities *namely 'identify and monitor change'*, *'assess and critique impacts'*, *'imagine alternative outcomes'*, *'envision preferred vision'* and *'plan and implement change'* in order. He puts the change to the center of foresight process and starts with catching the pattern of change with trend analysis. As a second activity, the impacts of change are defined by sorting them according to their relevance and importance. Determining alternative futures and picturing them in the mind is the key element called imagining alternative outcomes. Selecting the desired future and putting it forth as stated vision is the following activity. Last activity is comprised of specifying stakeholders, designating resources for the clear purposes, developing strategies, organizing activities and creating change.

With the light of the foresight frameworks in the literature, in this study a generic foresight functional model named 'FORESIGHT' which is built to draw outlines of the functional system is developed. When one of the functions doesn't work properly, it can't be mentioned about the proper operation of the whole system. In the proposed functional model FORESIGHT, there are nine functional blocks as '*framing*', '*obtaining*', '*reviewing*', '*establishing*', '*synthesizing*', '*illustrating*', '*guiding*', '*handling*' and '*tracking*'. In the model, consecutive functions are interrelated and interdependent.

[9] STEEPV: **S**ocial, **T**echnological, **E**conomic, **E**cological, **P**olitical, Values analysis tool is a framework to gauge how the external environment will impact a given company's strategic plan to remain competitive.

Framing: Fulfilling the tasks of determining foresight purpose, scope, content and time horizon.

Obtaining: Collecting data and information, gathering participants also by using co-nomination in iterative way which are consistent with its frame stated in the previous function.

Reviewing: Sharing ideas and opinions on the accessed data and information related with past and present, summarizing, analyzing them to be processed.

Establishing: Thinking about the future with the knowledge created, picturing possibilities in the minds and imagining the alternatives to create futures.

Synthesizing: Combining all alternative future thoughts with the present state conditions and resources in an interpretive way. Discussion, negotiation, facilitation and conflict resolution takes place in this function.

Illustrating: Pointing out the possible futures, visioning and generating reports, broadcasting with multimedia, sharing in social media.

Guiding: Defining actions and changes those will be performed, determining sequencing of them to reach different futures, strategy development and planning.

Handling: Taking actions, making changes and solving application problems.

Tracking: Evaluating outcomes and results of handling, performing impact analysis to take lessons for learning process.

In Table 3, functions in the FORESIGHT have been matched with the phases of mentioned foresight frameworks based on their actions and artifacts within specific phases.

Table 3: Foresight Frameworks in the Literature

<table>
<tr><th></th><th>Yüksel&Çifci (2017)</th><th>Martin (1995)</th><th>Miles (2002)</th><th>Voros (2003)</th><th>Bishop&Hines (2006)</th><th>Schultz (2006)</th><th>Sarıtaş (2011)</th></tr>
<tr><th></th><th>Foresight Functions</th><th>Foresight Process</th><th>The Foresight Cycle</th><th>A Generic Foresight</th><th>Framework Foresight</th><th>Key Activities of Integrated Foresight</th><th>Systemic Foresight</th></tr>
<tr><td>F</td><td>Framing</td><td rowspan="2">Pre-Foresight (Decision, Preparation)</td><td>Pre-Foresight</td><td rowspan="2">Inputs</td><td>Framing</td><td></td><td></td></tr>
<tr><td>O</td><td>Obtaining</td><td>Recruitment</td><td rowspan="2">Scanning</td><td>Identify and monitor change</td><td>Intelligence</td></tr>
<tr><td>R</td><td>Reviewing</td><td rowspan="5">Foresight (Process Design, Strategic Analysis, Agreeing, Disseminating)</td><td rowspan="5">Generation</td><td>Analysis Interpretation</td><td>Asses and Critique Impacts</td><td>Imagination</td></tr>
<tr><td>E</td><td>Establishing</td><td rowspan="2">Prospection</td><td rowspan="2">Forecasting</td><td rowspan="2">Envision Preferred Futures</td><td rowspan="2">Integration Interpretation</td></tr>
<tr><td>S</td><td>Sythesizing</td></tr>
<tr><td>I</td><td>Illustrating</td><td>Outputs</td><td rowspan="2">Visioning Planning</td><td></td><td></td></tr>
<tr><td>G</td><td>Guiding</td><td rowspan="2">Strategy</td><td rowspan="2">Plan and Implement Change</td><td rowspan="2">Intervention</td></tr>
<tr><td>H</td><td>Handling</td><td>Post-Foresight (Implemantation, Allocation)</td><td>Action</td><td>Action</td></tr>
<tr><td>T</td><td>Tracking</td><td></td><td>Renewal</td><td></td><td></td><td></td><td>Impact</td></tr>
</table>

4.CONCLUSION

Future has always been wondered and some systematic studies have been done to get ready to different futures through the time. So, there are different futures studies in the literature including the foresight which is the most common one especially from 1990s. To cover the all elements of prominent foresight definitions in the literature, the new definition of foresight which is based on the resources, methodology and futures strategies is presented. By searching the literature for foresight generations, sixth foresight generation is evaluated as the one dominated with netocracy and biotechnology.

It is assessed that the intensive use of big data, Internet of Things, add on biological units and artificial cyborgs will bring ethical and social dimensions together. The sixth generation foresight studies involve the netocrats, netizens, futurists and futurizens as the actors who initiate, carry out the operations and affected from them.

Additionally, a generic foresight functional model with nine consecutive phases named FORESIGHT (**F**raming, **O**btaining, **R**eviewing, **E**stablishing, **S**ynthesizing, **I**llustrating, **G**uiding, **H**andling, **T**racking) is suggested with the search of frameworks in the literature. This functional model covers the phases of a generic foresight process regarding to its activities done.

Futures will continue to remain the latent area always been tampered by humans even the science allows to read

the thoughts and determine the people tendencies in the future. The combination of these different thoughts and tendencies is incredible process, which is also affected by diverse factors, and creates ambiguity. In contrast to this uncertainty for futures, foresight is improvable area and can be made more explicit and useful by using diverse resources, effective methodology based on new advanced techniques and by constituting strategies meticulously for alternative futures. Foresight can become more effective and successful by applying its evaluation and impact analysis; executing it iteratively and treating it as learning process. Nevertheless, the nature of the partiality of futures -the impossibility of comprehension- will remain the most important limitation of foresight process.

REFERENCES

[1] J. Schatzmann, R. Schäfer, and F. Eichelbaum, “Foresight 2.0 - Definition, overview & evaluation,” *Eur. J. Futur. Res.*, vol. 1, no. 1, p. 15, 2013.

[2] J. Voros, “A Primer on Futures Studies,” *Prospect Foresight Bull.*, vol. 6, no. 1, 2001.

[3] R. Amara, “The Futures Field: Searching for Definitions and Boundaries,” *Futurist*, vol. 15, no. 1, pp. 25–29, 1981.

[4] P. R. Kreibich, B. Oertel, and M. Wölk, “Futures Studies and Future-oriented Technology Analysis Principles, Methodology and Research Questions,” in *Berlin Symposium on Internet and Society*, 2011, pp. 1–53.

[5] Z. Sardar, “The Namesake: Futures; futures studies; futurology; futuristic; foresight-What’s in a name?,” *Futures*, vol. 42, no. 3, pp. 177–184, 2010.

[6] R. A. Slaughter, *The Foresight Principle: Cultural Recovery in the 21st Century*. London: Adamantine Press, 1995.

[7] M. Conway, *Foresight: an Introduction*. Thinking Futures, 2015.

[8] G. Steed and S. Tiffin, “A National Consultation on Emerging Technology,” in *Science Council of Canada*, 1986.

[9] B. R. Martin, “The origins of the concept of ‘foresight’ in science and technology: An insider’s perspective,” *Technol. Forecast. Soc. Change*, vol. 77, no. 9, pp. 1438–1447, 2010.

[10] K. Cuhls, “From forecasting to foresight processes—new participative foresight activities in Germany.,” *Forecast*, vol. 22, pp. 93–111, 2003.

[11] D. Loveridge, “Technology foresight and models of the future,” in *CEC ASTPP Network Conference Ideas in Progress*, 1996, no. Number 4.

[12] A. Hines and J. Gold, “Professionalizing foresight: Why do it, where it stands, and what needs to be done,” *J. Futur. Stud.*, vol. 17, no. 4, pp. 35–54, 2013.

[13] B. R. Martin, “Foresight in science and technology,” *Technol. Anal. Strateg. Manag.*, vol. 7, no. 2, pp. 139–168, 1995.

[14] I. Nonaka and N. Konno, “The concept of ‘Ba’: BUILDING A FOUNDATION FOR KNOWLEDGE CREATION,” *Calif. Manage. Rev.*, vol. 40, no. 3, pp. 40–54, 1998.

[15] V. Kamtsiou, A. Naeve, L. K. Stergioulas, and T. Koskinen, “Roadmapping as a Knowledge Creation Process: The PROLEARN Roadmap,” *J. Knowl. Manag.*, vol. 1, no. 3, pp. 163–173, 2006.

[16] H. A. Linstone, “Three eras of technology foresight,” *Technovation*, vol. 31, no. 2–3, pp. 69–76, 2011.

[17] B. R. Martin and R. Johnston, “Technology Foresight for Wiring Up the National Innovation System-Experiences in Britain, Australia, and New Zealand,” *Technol. Forecast. Soc. Change*, vol. 60, no. 1, pp. 37–54, 1999.

[18] L. Georghiou, “The UK Technology Foresight Programme,” *Futures*, vol. 28, no. 4, pp. 359–377, 1996.

[19] L. Georghiou, J. C. Harper, M. Keenan, I. Miles, and R. Popper, *The Handbook of Technology Foresight: Concepts and Practice*. Cheltenham, UK: Edward Elgar Publishing, Inc., 2008.

[20] R. Barré, “Synthesis of Technology Foresight, Strategic Policy Intelligence: Current Trends, the state of Play and Perspectives,” 2001.

[21] I. Miles and M. Keenan, “Practical Guide To Regional Foresight in the UK,” 2002.

[22] I. Miles, “The development of technology foresight: A review,” *Technol. Forecast. Soc. Change*, vol. 77, no. 9, pp. 1448–1456, 2010.

[23] J. C. Harper, “Vision Document, eFORESEE Malta ICT and Knowledge Futures Pilot, 2003.” 2003.

[24] J. Voros, “A generalised ‘layered methodology’ framework,” *Foresight J. Futur. Stud. Strateg. Think. Policy*, vol. 7, no. 2, pp. 28–40, 2005.

[25] M. Keenan, R. Popper, M. Alexandrova, D. Marinova, D. Tchonkova, and A. Havas, “Research infrastructures foresight (rif),” 2007.

[26] R. Popper, “Foresight Methodology : an overview and more...” 2008.

[27] R. Popper, "Dr. Popper's Foresight and Horizon Scanning Blog-Futures." 2011.

[28] J. M. Schmidt, "Policy, planning, intelligence and foresight in government organizations," *Foresight*, vol. 17, no. 5, pp. 489–511, 2015.

[29] M. Jemala, "Evolution of foresight in the global historical context," *Foresight*, vol. 12, no. 4, pp. 65–81, 2010.

[30] G. Reger, "Technology Foresight in Companies: From an Indicator to a Network and Process Perspective," *Technol. Anal. Strateg. Manag.*, vol. 13, no. 4, pp. 533–553, 2001.

[31] R. Boutellier, U. Deplazes, and K. Loffler, "Model of Technology Foresight: An innovative Approach," in *IEEE International Engineering Management Conference*, 2007, pp. 7–14.

[32] L. Georghiou and M. Keenan, "Evaluation of national foresight activities: Assessing rationale, process and impact," *Technol. Forecast. Soc. Change*, vol. 73, no. 7, pp. 761–777, 2006.

[33] R. Popper, L. Georghiou, M. Keenan, and I. Miles, *Evaluating Foresight Fully-Fledged Evaluation of the Colombian Technology Foresight (CTFP)*. Universidad del Valle (Colombia), 2010.

[34] S. Gieseke, "FLA as a means of participation in modern democratic decision-making," in *The Role of Forward-Looking Activities for the Governance of Grand Challenges, Insights from the European Foresight Platform*, 2012.

[35] M. Conway, "An Overview of Foresight Methodologies," *Thinking Futures*. pp. 1–10, 2006.

[36] O. Saritas, "Sytemic Foresight Methodology," in *Forth International Seville Conference on Future-Oriented Technology Analysis (FTA) FTA and Grand Societal Challenges – Shaping and Driving Structural and Systemic Transformations*, 2011, p. 34.

[37] J. C. Harper, "Impact of Technology Foresight," 2013.

[38] R. Studer, R. Benjamins, and D. Fensel, "Knowledge engineering: Principles and methods," *Data Knowl. Eng.*, vol. 25, no. 1–2, pp. 161–198, 1998.

[39] A. Abdollahi, M. Fasanghari, M. Mohammedpour, M. Soltani, and M. Nili, "An Ontology Based Model for Technology Foresight," in *2008 Third International Conference on Convergence and Hybrid Information Technology*, 2008, vol. 1, pp. 510–514.

[40] J. Voros, "A generic foresight process framework," *Foresight*, vol. 5, no. 3, pp. 10–21, 2003.

[41] A. Hines and P. Bishop, *Thinking about the Future:Guidelines for Strategic Foresight*. Washington, DC: Social Technologies LLC, 2007.

[42] W. L. Schultz, "The cultural contradictions of managing change : using horizon scanning in an evidence-based policy context," *Foresight*, vol. 8, no. 4, pp. 3–12, 2006.

[43] Ma, "In Wikipedia. Retrieved April 10, 2017, from https://en.wikipedia.org/wiki/Ma_(negative_space)." .

[44] Netocracy, "In Wikipedia. Retrieved April 22, 2017, from https://en.wikipedia.org/wiki/Netocracy." .

[45] A. Bard and J. Soderqvist, *NETOCRACY: The New Power Elite and Life After Capitalism*. Stockholm: Pearson Education, 2002.

[46] M. Hauben, "Preface: What is a Netizen." pp. ix–x, 1995.

[47] M. Lee, "Futurists and futurizens," no. July. pp. 1–3, 2013.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

EXPANSION OF A RELATIONAL DATABASE TO SUPPORT SEMANTIC WEB QUERIES

DOI: 10.17261/Pressacademia.2017.594
PAP-WCTIE-V.5-2017(32)-p.234-239

Ylber Januzaj[1], Samedin Krrabaj[2], Artan Luma[3], Arben Hajra[4], Besnik Selimi[5]
[1]University of Mitrovica. ylber.januzaj@umib.net
[2]University of Prizren. . samedin.krrabaj@uni-prizren.com
[3]South East European University. b.selimi@seeu.edu.mk

ABSTRACT

Nowadays various types of data, especially those with public character, are stored and represented in relational databases. While this way of storing data is a practice for most of the institutions, it turns them into isolated silos with low level of accessibility and interoperability in the web. Indeed not all datasets are readable from the World Wide Web, therefore to increase their access it is necessary to provide mapping from relational databases to any serialization format of Resource Description Framework (RDF), as one of the best practices for distributing and interlinking data on the web. While there are several ways to accomplish the mapping process, the most appropriate and fastest way is by using the Protégé Plugins application. In this paper we present this mapping method, and describe the complete conversion of a database from SQL Server relational Database to RDF using different scenarios. By using Sesame Repository, as the local repository for our dataset, several SPARQL queries will be generated on top of our data. As a main contribution this paper emphasizes the advantages and performance of querying and using serialized RDF statements versus those stored in the relational databases.

Keywords: Linked data, semantic technologies, RDF, Protégé.

1. INTRODUCTION

Lately, technology research especially those of the World Wide Web have reached a high point of research and development, however there is still a lot to be discovered. As mentioned in [1],[5] we claim there are a lot of changes not only in the functional way but in the organizational way as well. These changes have occurred following the developments of Semantic Web, but what is more important is that we still consider that the amount of data on the Semantic Web remains collateral, when compared to the traditional Web. Based on all these changes that occurred in the World Wide Web, the need to work with dynamic data in our Web site, arises.[3],[4],[14] Unfortunately, not all records are appropriate to be used in Web sites, so the importance to import data from a relational database to ontologies is frequently required. There are many priorities of RDF over a relational database, and we will mention the most important ones. [11] We all know that life-science is developing rapidly, and there are hundreds of subjects that have developed a data-warehouse with information that needs to be either private or public. If we have a large database company, for instance, which needs to create a large database, they hire a team of programmers which consist of beginner level programmers up to the most advanced ones. All of these programmers need approximately one year to tune the database, and after a while changes will happen, for example schema changes, format changes, and the requests for new data sources will grow up. In RDF no changes will happen, because triples are triples, and there is no need to change any schema, or any tables. The only thing that is required to do, is updating the queries. Second, SPARQL is more stable than SQL since SQL has many dialect changes from one implementation to another. And finally RDF has a better performance in grouped data. Let's imagine a website which responds to a single record, where each record consists of many fields, and after we have a normalized SQL database we will have 40-50 tables with many-to-many relations, and in order to display each page we need to join them. In RDF there is a simple way to call the records. If displaying a record is needed, for instance one in the 30th table, all we need to know in order to display it, is the URI of that record. Exhuming the relational data and making them available as RDF triples is not an unknown field, and there are many negotiability in this area. There are many tools which do the import, from an existing relational database to RDF, but in this paper we will present the automatic mapping way with Protégé Tools from

an existing SQL relational database, which uses the 'ready' plug-ins which need to be configured to achieve a successful connectivity and import. First, we will show the configuration of Protégé tools, which are necessary to connect to the SQL Server in the distance. The connection will be done using IP addresses. Then, after we have successfully connected, we will do a complete import of tables, including the data that is in them. Next, the automatically created Classes will be detected, along with Slots and Instances. These Slots and Instances will be able to be modified the way we want to, but we should always be careful because later all these Classes, Slots, and Instances will be used to query the data. At the end of the paper we will show some queries which will be implemented on Sesame server, where as an endpoint, will be our local machine, and after this point we consider that they are able to query from the Web sites and PHP programmers.

2. INITIAL STATE (MAPPING)

In [13],[15] is emphasized that obstacles of Semantic Web applicants are non-semantic information which are required for transformation in Semantic Web language (e.g., OWL, RDF). The goal of our project is to ensure that we will cover all OWL constructs, and support all OWL entities such as Classes, Properties and Individuals.

2.1. Ontologies Describing the Database Schema

Ontology is a language which serves to subscribe semantic web data, or also known as the vocabulary of semantic web. In our project, after a conversation we have ontologies which system reserves for us. We will nominate them in vocabulary, which is globally known for easy access, like Dublin Core (DC)[1], Friend of a Fried (FOAF)[2], The Bibliographic Ontology Specification (BIBO)[3], etc., for having an easy interoperability with other resources. In this context, the possibility to enrich them with other data from different repositories will be considered through ontology alignment and also by creating ongoing links by different services like sameas.org[4]. The database schema is described through ontologies in different ways, depending on the form that it is required and on the form that the application work.[1],[4],[7] It is important because in some cases it will be required to import all the database with the possibility to make "live" changes on it, in some other cases it will be required to import the database not in "live" form, because if we try to make changes in our database, then we should update the importation because we cannot see the changes that occur in database. In the next part, the advantages and benefits of an RDF repository will be described through the conversion of a bookstore database. There are thousands of records saved on a relational database, and we will see the delay of the importance. We will also see the OWL ontologies and the frame ontologies that will describe our database.

2.2. Using Protégé Plugin

Protégé is developed by Stanford Center for Biomedical Information Research at the Stanford University School of Medicine, and it is an open source project. [2],[6] There are many developed versions of Protégé but the most appropriate version to perform for this paper is 3.4.4. As we mentioned in the previous sections there are many tools that enable us to convert from relational database to RDF, as D2RQ[5], Triplify[6], etc. Our choice is Protégé because its comprehensive features, functionalities and a well-designed GUI (Graphical User Interface).

2.2.1. First Steps

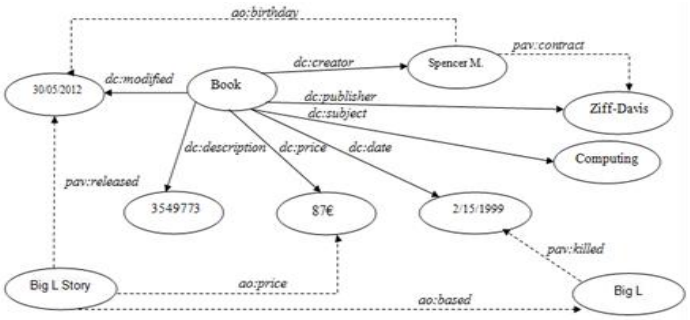
The goal is to convert the data from relational database to RDF in a way to be published and used as open data. For this purpose we want to share something that is on both sides' interest, citizens and the provider. There is chosen a Bookstore database, from a small Bookstore in our city. The database contains several entities, in different relationships, like Authors, Publishers, Subjects, and Titles. We will focus on a few of them and make these information public to the citizens, and at the same time the Bookstore employees will have the opportunity to extend the data at any time. Having such data bring several advantages to people (readers, buyers), more up to data information, interoperability with other repositories, enriching the data with various other data, etc.

2.2.2. Dealing with Ontology

In this section we will see the Ontology which will be created after we make a complete con-version from the relational database to the RDF.

[1] http://dublincore.org/documents/dces/
[2] http://xmlns.com/foaf/spec/
[3] http://bibliontology.com/
[4] http://sameas.org/
[5] http://d2rq.org/
[6] http://triplify.org/

Figure 1: Sketch of Books Ontology



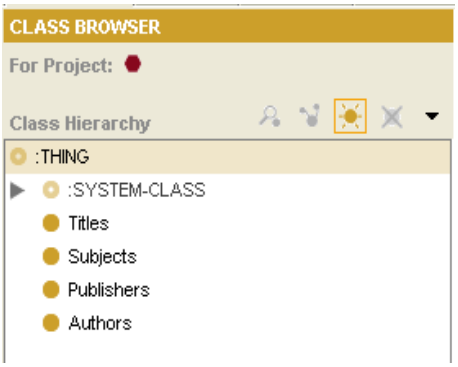
In [12] is emphasized that an important component in this context is the definition of Classes, Properties and Individuals. In the following section we will show the importance of these components:

- *Classes* tab – building class hierarchy
- *Properties* tab – creating object properties, data properties and assign domains
- *Individuals* tab – creating individuals

2.2.3. Classes

Resources which are separated in groups and are in RDF are called Classes. In our project we will see these classes which in relational database were known as entities. As it is mentioned in [6], after a successful conversion from the relational database to the RDF statements, we need to define the classes. These classes are created automatically but we can manipulate with those classes for having specific customization, by choosing the adequate class we want to work on. In Figure 2 is showed a print-screen sample of the classes created automatically by Protégé. This is our imported database ontology which has owl:Thing as parent of all classes.

Figure 2: Class Browser

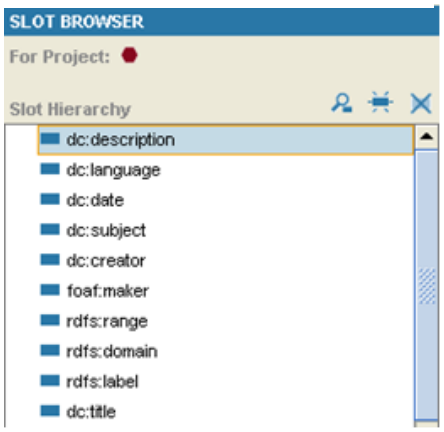


2.2.4. Properties

A property is nothing else but a relation between the subject and an object, which needs to be related by a property to enables us to see the necessary information. By Protégé, both types of Properties, the object and data properties can be manipulated in the same way. Also, below you can see a print-screen example of the properties created also automatically. We should mention that these properties in the relational database were attributes, for example the attribute Title from the table Titles, is being displayed as Titles. In the Figure 3 have a list of these properties. By default, during the conversation process, the properties are automatically created from the table attributes, and so they get the same name.

However, in section 2.1 is mentioned that the best practice is the implementation of well know ontologies and vocabularies. For this purpose there is applied Dublin Core in *dc:description, dc:date, dc:creator*, etc. instead of the attributed names that came from the table.

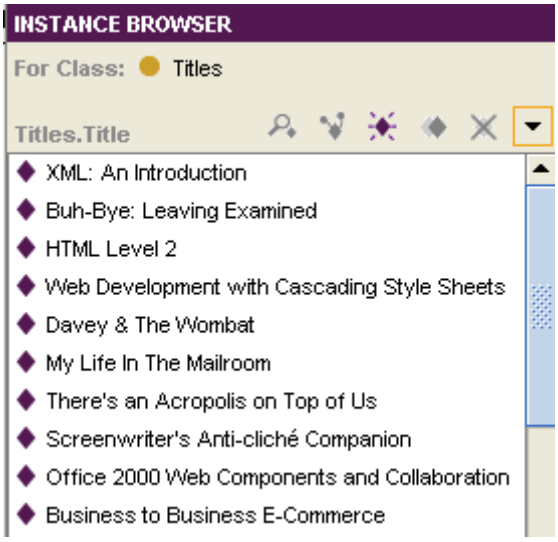
Figure 3. Slot Browser with Properties



2.2.5. Individuals

Unlike the class and property, individuals is the exact place where the information is stored. The relational database after the conversion to the RDF, will be totally converted with the records on it. These records on the database, after the conversion, are known as Individuals and they are connected to the Properties. [6] As default they will be displayed as Instance_Nr, where dhe Nr will refer to the number of the records, and in the configuration of the individual we can choose to display in the form that we want. For example for the Class Title, we can choose to display the individuals as Title, PubDate, ISBN, Info or whatever way they are described. We can also modify them, delete, rename, moreover we can insert a new instance as a new individual, and we can query it like it was converted from the relational database. Figure 4 will show an example of the Individuals after converted by the Protégé – OWL.

Figure 4. Instance Browser including Converted instances



2.2.5. Reasoner Usage

After the ontology creation, must be sure that if this ontology is consequent or not by using a reasoner. Protégé – OWL is distributed with a reasoner called HermiT which needs to be configured and installed as a plug-in. So, after the check

procedure, if there is no any error then we can conclude that the ontology is consequent. We can explain this error with an example: consider we have the Titles.PubDate property value, and if we try to put string values on it, then after we make a reasoner check we will see that we have an error. The error will say ontology is inconsequent, and this is because in the Titles.PubDate we were required to put data type values and not string values.

3. STORING RDF TRIPLES

After a successful conversion, the serialized data need to be stored in an RDF Repository. At the moment there is a relatively large list of choices for this purpose, where each of them has its advantages and disadvantages. Among these, the most know tools are AllegroGraph[7], OpenLink Virtuoso[8], Sesame[9] , 4store[10] , etc. For some practical reasons, in this project our choice has been Sesame, which includes parsing, storing, inferencing and querying such data. Thus, after running the Sesame, locally or in a server, we are ready to upload our data, as .N3 format.[8],[9],[10] In this way a successful upload guarantee that our data are stored and ready for consuming by other people. For this purpose we are offering a SPARQL endpoint as one of the best practices for consuming the data.

3.1. SPARQL Endpoint

In the following section we will show the possibility of retrieving data from our repository by querying the SPARQL endpoint deployed at the Sesame server. The current link where the datastore is saved locale looks like this: http://localhost:8080/openrdf-workbench/repositories/bookstore. In the Listing 1 is showed an example for illustrating the retrieved data with a simple SPARQL query. So, the Title, Published Year and the Price for all the books will be listed.

Listing 1.

```
SELECT ?Title ?PubDate ?Price
WHERE
{
?p dc:title ?Title.
?p dc:date ?PubDate.
?p rdfs:domain ?Price.
}
 LIMIT 10
```

The retrieved data of the query from the Listing1 are showed in the *Figure 5*.

Figure 5: Table after SPARQL Query Execution

Title	PubDate	Price
"There's an Acropolis on Top of Us"	"1985-01-08 00:00:00.0"	"27.0"^^xsd:float
"Kitchen Chemistry"	"1998-01-22 00:00:00.0"	"28.0"^^xsd:float
"Creating Your First Spiral Galaxy"	"1997-03-04 00:00:00.0"	"18.0"^^xsd:float
"The Magic Thingamajig"	"1994-01-03 00:00:00.0"	"37.0"^^xsd:float
"Algae: Smarter Than You Think"	"1990-01-16 00:00:00.0"	"19.99"^^xsd:float
"The All Igneous Rock Guide"	"1989-02-01 00:00:00.0"	"37.0"^^xsd:float
"The Road Taken"	"1992-01-22 00:00:00.0"	"42.0"^^xsd:float
"An American Story"	"1993-01-18 00:00:00.0"	"19.99"^^xsd:float
"The Ontologist's Guide To Happiness"	"1988-02-07 00:00:00.0"	"30.0"^^xsd:float
"No More Why: All The Answers"	"2000-02-01 00:00:00.0"	"41.0"^^xsd:float

Thus, everyone can perform various SPARQL queries for getting any kind of information. Everything is accessible without any special hardware or software requirements. The data can be consumed, integrated and visualized everywhere and for any purpose.

[7] http://franz.com/agraph/allegrograph/
[8] http://virtuoso.openlinksw.com/
[9] http://openrdf.org/
[10] http://4store.org/

4. CONCLUSION

In our paper we presented the advantages of serialized RDF statements versus Relational Databases using Data Master Protégé Plug-in, which allows us to convert the database from relational model to the RDF statements by using the Java mechanisms. All this import is done by using the JDBC driver which creates a kind of a bridge between the Protégé application and the SQL Server. We presented the ontologies, manipulation with them and a consummation through the SPARQL endpoint deployed in the SESAME Server. Offering the data as Linked Open data, especially by applying a SPARQL endpoint, bring several benefits to the readers, buyers and to the bookstores also. At any time these data can be used and visualized by anyone for any purpose. In addition, the data will be enriched with other data stored in different repositories.

REFERENCES

[1] Csongor Nyulas, Martin O'Connor and Samson Tu, "DataMaster – a Plug-in for Importing Schemas and Data from Relational Databases into Protégé," in Stanford Medical Informatics, Stanford University School of Medicine, Stanford, CA 94305, 2007.

[2] Farid Cerbah, "Learning Highly Structured Semantic Repositories from Relational Databases The RDBToOnto Tool" in Dassault Aviation, DPR/ESA, 78 quai Marcel Dassault 92552 Saint-Cloud , France, 2008.

[3] Li Ma, Jing Mei, Yue Pan, Krishna Kullkarni, Achille Fokue, Anand Ranganathan, "Semantic Web Technol-ogies and Data Management" in IBM China Research Laboratory, IBM Software Group, Ibm Watson Research Center, Bei Jing 100094, China San Jose, CA, 2012.

[4] Juan F. Sequeda, Syed Hamid Tirmizi and Daniel P. Miranker, "SQL Databases are a Moving Target Position Paper for W3C Workshop on RDF Access to Relational Databases" in The University of Texas at Aus-tin, Department of Computer Sciences, Austin, USA, 2007.

[5] Simeon Polfliet, Ryutaro Ichise, "Automated Mapping Generation for Converting Databases into Linked Data" in Ensimag engineering school Grenoble Institute of Technology (INPG) Grenoble, France, Principles of Informatics Research Division National Institute of Informatics, Tokyo, Japan, 2010 .

[6] Adam Jurcik,"Development of Visualization Plug-in for Protege" in Masarykova Univerzita Fakulta In-formatiky, Brno, 2010.

[7] Juan F. Sequeda, Marcelo Arenas and Daniel P. Miranker, "On Directly Mapping Relational Databases to RDF and OWL" in The University of Texas at Austin, PUC Chile, Lyon, France, 2012.

[8] Rafael Valencia-García, Francisco García-Sánchez, Dagoberto Castellanos-Nieves, and Jesualdo Tomás Fernández-Breis, " OWLPath: An OWL Ontology-Guided Query Editor" in Ieee Transactions On Systems, Man, And Cybernetics—Part A: Systems And Humans, 2011.

[9] A. Bernstein and E. Kaufmann, "Gino—A guided input natural language ontology editor," in Proc. Int. Se-mantic Web Conf., vol. 4273, Lecture Notes in Computer Science, I. F. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, and L. Aroyo, Eds.pp. 144–157, UK, 2006.

[10] H. Namgoong and H.-G. Kim, "Ontology-based controlled natural language editor using CFG with lexical dependency" in Proc. ISWC/ASWC, vol. 4825, Lecture Notes in Computer Science, K. Aberer, K.-S. Choi, N. F. Noy, D. Allemang, K.-I. Lee, L. J. B. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, and P. Cudré-Mauroux, Eds., pp. 353–366, Korea, 2007.

[11] H.-G. Kim, B.-H. Ha, J.-I. Lee, and M.-K. Kim, "A multi-layered application for the gross description using semantic web technology" in Int. J. Med. Inf., vol. 74, no. 5, pp. 399–407, Hindawi Publishing Corporation, 2005.

[12] Lei ZHANG, Jing LI, "Automatic Generation of Ontology Based on Database" in School of Computer Sci-ence and Technology, China University of Mining and Technology, Xuzhou 22111, Jiangsu Province, China, 2012.

[13] Rai Ghawi, Nadine Cullot, "Database-to-Ontology Mapping Generation for Semantic Interoperability" in Laboratory LE2I, University of Burgundy, Dijon France, 2007.

[14] Nikolaos Konstantinou, Dimitrios-Emmanuel Spanos, Michael Chalas, Emmanuel Solidakis and Nikolas Mitrou," VisAVis: An Approach to an Intermediate Layer between Ontologies and Relational Database Con-tents" in National Technical University of Athens, Division of Communications, Electronics and Information Systems, Heroon Polytechniou, Athens, Greece, 2006.

[15] Martin J. O'Connor, Christian Halaschek-Wiener, Mark A. Musen, "Mapping Master: a Flexible Approach for Mapping Spreadsheets to OWL" in Stanford Center for Biomedical Informatics Research Stanford, USA, 2010.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

SOLVOTHERMAL SYNTHESIS AND BIOLOGICAL ACTIVITY OF NI-DOPED ZINC OXIDE NANOPARTICLES

DOI: 10.17261/Pressacademia.2017.595
PAP-WCTIE-V.5-2017(33)-p.240-246

Sammia Shahid[1]*, Muhammad Mudassar Sher[2]
[1]Department of Chemistry, School of Science, University of Management & Technology, Lahore-, Pakistan. sammia.shahid@umt.edu.pk
[2]Department of Chemistry, School of Science, University of Management & Technology, Lahore-, Pakistan. mudassarsher786@gmail.com

ABSTRACT
Metal oxide nanoparticles are potential candidate for making future antimicrobials. Increased interest is due to change in fundamental properties at nanoscale. Ni-doped zinc oxide nano-particles were prepared for pharmacological studies. Co-precipitation and solvo-thermal methods were employed which yielded Ni-doped zinc oxide nano-particles and un-doped zinc oxide nano-particles were synthesized via solvo-thermal method. All prepared nano-particles were characterized using X-ray diffraction studies whereas doping was confirmed by Energy Dispersive X-ray analysis. Shape and morphology of these nano-particles was assessed using Scanning Electron Microscopy. The synthesized nano-particles have shown antibacterial activity against both *Gram*-negative and *Gram*-positive bacteria designating these nano-particles as future broad spectrum antibacterial. The optical properties were also studied by measuring the energy band gap and were found 1.50 *eV* for un-doped zinc oxide nano-particles, and it decreases to 1.47 *eV* for Ni-doped zinc oxide. Ni-doped zinc oxide nano-particles were proved to be active future pharmaceutical and biomedical agents.

Keywords: Antibacterials, morphology, nanoparticles, energy dispersive X-ray, doping

1. INTRODUCTION

The particles, rods and systems having one or more dimensions in nanometer scale range are the possible smallest things that we can produce up till now (Liewhiran et al., 2006). The particles having size ranges from 1-100 nanometers are called as nano-particles. These particles are unique in their properties and they also find applications in various fields (Hewakuruppu et al., 2013). For instance nano-particles are being used electronics, energy production, biosciences and medicines (B nuzea et al., 2007).

Similarly nano-particles of metallic origin found applications in electronics, medicine and in many other industries (Sahoo et al., 2007). Metals have been used as raw materials to produce distinct nanostructures such as nano-rods, nano-nails, nano-wires, nano-tubes and nano-belts besides nano-particles alone (Hsu et al., 2006). Oxides of some metals are also in use for their ease in fabrication to nano materials. For example Zinc oxide has been used to produce nano-particles, nano-sheets, nano-tubes, nano-wires, nano-belts and nano-rings, etc. (Hua et al., 2007).

Bio-inspired productions of Zinc oxide nano-particles were achieved using environmentally along with eco-friendly accepted systems (Cai et al., 2013). Different chemical methods have been used to prepare Zinc oxide nano-particles by using different materials of interest. The majority of the Zinc oxide crystals were synthesized by using traditional high temperature (solid state) method but in this method it is difficult to organize the particle properties along with energy consumption. Zinc oxide nano-particles can also be prepared on a big scale at low cost by using simple solution based methods, such as chemical precipitation, hydrothermal reaction and sol–gel synthesis (Zhong et al., 1996).

Zinc oxide nano-particles are most favored due to its applications such as transparent conducting films, gas sensor, piezoelectric transducers as well as photo catalytic activity (Zhang et al., 2009). Zinc oxide nano-particles are used in different commercial products for instance in cosmetics, sunscreens and for anti-bacterial activity (Xu et al., 2010). Other nanostructures of Zinc oxide are also found applications in optoelectronics, transducers, sensors, and biomedical sciences

(Wang et al., 2004)**.** Doping of Zinc oxide nano-particles with other metals such as Ni, Ga, In, Sn, Al, Y, and Sc, are also in practice to improve properties Zinc oxide nano-particles (Moriga et al., 2004). Nickle recognizes as one of the most proficient doping element to advance as well as tune the electrical and optical properties of Zinc oxide nano-particles. The main emission peak of Zinc oxide is found as a sturdy ultraviolet (UV) emission at 380 nm. (Al-Harbi et al., 2011)**.**The interest to study optical properties of Zinc oxide nano-particles doped with Ni must provoke to develop an understanding of the material response to doping. There are numerous reports presented in the literature on optical properties Zinc oxide nano-particles doped with transition metals (Yan et al., 2003).

The nano-particles and nano-rods have been prepared by various groups using different methods such as sol–gel process, spray paralyses, laser molecular beam epitaxy, metal organic chemical vapor deposition, and pulsed laser deposition. Among these, the solvo-thermal methods have surpassed the others due to possessing better homogeneity, being low-cost, and being simple (Wani et al., 2007). The synthesized nano-particles by above method have high surface area as compared to other methods, low in cost and ease of synthesis. These synthesized nano-particles found applications as antibacterial and antimicrobial agent (Hadjipanayis et al., 1994). The characterization techniques have been used for nano-particles such as Scanning electron microscopy (SEM), Transmission electron microscopy (TEM), X-ray diffraction (XRD), Photoluminescence (PL), Energy Dispersive X-Ray Analysis (EDX), UV-visible Absorption Spectroscopy and Superconducting Quantum Interference Device (SQUID) (Ananda et al., 2014).These techniques are developed due to the recent research in nano-science and nanotechnology is considered most important as well as most growing technologies (Khatoon et al., 2012).

Diseases have been the threat to human being since ancient times and these were mostly caused pathogenic microbes. Bacteria as part of pathogenic microbes' community are responsible spreading different minute and chronic diseases. Although different antibacterials are available in the market but research is going on to develop various chemical agents known as antibacterials to cope this threat in best way. The production of nano-particles their property to kill bacteria has been proving as alternative to conventional medicines. Antibacterial activity of different types of s Therefore it is of great importance to know and to understand antimicrobial effects synthesized nano-particles have been reported in the literature. The effect of doping to nano-particles is the focus of this study.

2. DATA AND METHODOLOGY

The chemicals used were obtained from Merck. The chemicals used were Zn $(CH3COO)_2$ $2H_2O$ (zinc acetate dehydrate) $NiCl_2$ $6H_2O$ (nickel chloride hexahydrate), diammonium oxalate, $ZnSO_4$ $7H_2O$ (zinc sulphate heptahydrate), $Ni(NO_3)_2$ $6H_2O$ (nickel nitrite hexahydrate), NaOH, Nutrient agar, Nutrient broth, ethanol and Methanol.

2.1 Bacterial Strains

These stains were obtained from PCSIR laboratories Lahore, Pakistan. *Escherichia Coli* is responsible for urinary tract infections, food poisoning, neonatal meningitis where as *Bacillus subtilis*, is responsible for sepsis, pneumonia and meningitis in human beings.

Following two methods are employed for the synthesis of Ni-doped ZnO nano-particles.

2.2 Method-I

Nano-particles were synthesized by solvo-thermal method through oxalate precursor route. 0.1M solutions were prepared by mixing (Zn $(CH3COO)_2$ $2H_2O$) 22.277 g in 1000ml DW and $NiCl_2 6H_2O$ 23.769 g in 1000ml DW water. Stoichiometric amount of zinc acetate dihydrate (CDH, 98.5%) and nickel chloride hexahydrate $NiCl_2$ $6H_2O$ (Merck, 97%) were stirred for about 30 minutes. The mixture was precipitated with 175 mL of aqueous solution of diammonium oxalate. 75 ml of ethanol was also added to the reaction mixture. A light green suspension formed. The mixture was refluxed at about 80 °C for 12 hours. The precipitate was recovered by centrifugation and dried in an oven at 55 °C and finally ground to powder. Nano-particles were obtained by thermal decomposition of the precursor at 450 °C for 6 hours in air.

2.3 Method-II

To obtain the desired degree of doping of Ni, $ZnSO_4$ $7H_2O$ was mixed in distilled water with $Ni(NO_3)_2$ $6H_2O$. These solutions are designated as solution A. Solution A was placed in an ultrasonic cleaner operating at 57 kHz for 2 h. Solution B is concentrated solution of NaOH in de-ionized water. After sonication, solution A was stirred with a magnetic stirrer at room temperature. Then solution B was added until a pH of 12 was reached. The solution was magnetically stirred for 0.5 h, and then the solution was allowed to stand at room temperature for 18 h. Subsequently, the solution was centrifuged and washed several times with ethanol and distilled water to remove residual and unwanted impurities. The final product was dried in a vacuum oven at 200 $^{\circ}$C for 1 h to yield Ni-doped ZnO particles.

2.4 Characterization

The structural characterization of synthesized nano-particles was carried out by SEM. For further characterization and average crystallite size determination XRD was also performed. The results were given in Table 1 and 2. Doping of Ni on ZnO was confirmed by EDX spectra. UV-VIS-spectrophotometer was used to check photo activity of nano-particles by energy band gap studies.

2.5 Bacterial Culture Used for Antibacterial Study

The prepared and sterilized agar solution was poured in sterilized petri plates under laminar and allows cooling up to room temperature. 1ml of inoculum was added to each plate immediately into medium plates and was shaken gently. The petri plates then allow standing until bacteria cultured media become semisolid. Wells were created by using stainless steel borer, four wells in each petri plate. Negative and Positive controls in two wells were applied separately in each petri plate along with two wells containing different concentrations of sample solutions by using micro syringe. After that these petri plates were placed for 24 hours at 37°C in an incubator. After allotted time petri plates were taken out from incubator diameters of inhibition zones were measured using millimeters scale. Each experiment was made in triplicate and the inhibition zones are given as the mean ± standard deviation. The results were recorded in Table 3 (Leven et al., 1979).

2.6 Minimum Inhibitory Concentration (MIC)

For determination of MIC recommended methods of NCCLS; 2000 in modified version was employed. In short, aerobically incubated tubes contained 5 mL Muller- Hinton (MH) broth (Difco, USA) with approximate 5×10^9 CFU bacterial cells for 24 hours at 37 °C, without nano-particles (the control group) and various concentrations of synthesized nano-particles. The MIC is determined from the tube without visible growth of the bacterial cells.

2.7 Minimum Bactericidal Concentration (MBC)

In order to find MBC 100 µL of sample from the tube without visible growth transferred into MH agar plate (Difco, USA), and it was then incubated aerobically again for another 24 h (NCCLS, 2000). The concentration of tube without growth was the MBC.

3. FINDINGS AND DISCUSSIONS

3.1 XRD Analysis

XRD diffraction patterns of Ni-doped zinc oxide nano-particles synthesized through co-precipitation method are shown in figure-1 and the XRD diffraction patterns of Ni-doped zinc oxide nano-particles form through solvo-thermal method through oxalate precursor route are shown in figure-2. The nano-particles obtained through co-precipitation method are termed as M_1. On the other hand nano-particles obtained through solvo-thermal method are termed as M_2.

Fig 1: X-ray pattern of the Ni doped ZnO nano-particles M_1

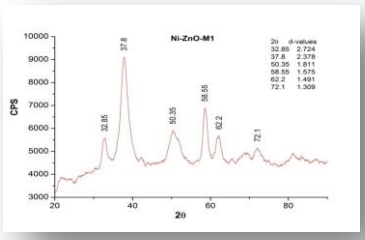
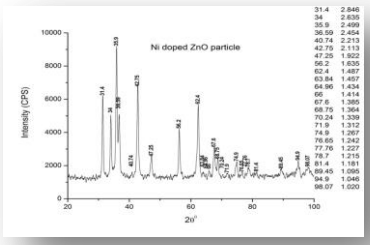


Fig 2: X-ray pattern of the Ni doped ZnO nano-particles M_2



The obtained XRD diffraction patterns having the peaks match to the characteristic peaks of ZnO XRD diffraction pattern (JCPDS file No: 36-1451). Similarly the XRD diffraction patterns of Ni-doped ZnO also match to the characteristic peaks Ni-doped ZnO XRD diffraction pattern (JCPDS card no.78-0643) confirming the doping and synthesis of desired product along with absence of any impurity present in synthesized samples.

3.2 Crystalite Average Size of Nano-Particles

The Crystallite average sizes of all prepared samples of Ni-doped zinc oxide nano-particles were calculated for all the using Scherrer formula.

$$D = \frac{K\lambda}{\beta cos\theta}$$

Here K is constant estimated as 0.9, peak width is represented by β which is calculated at half height of peak in radians, λ represents wavelength of radiation used. Angle of reflection is shown "θ".

Table 1: Crystallite Size of Ni Doped ZnO Nano-Particles M_1

2 θ (deg) of the Intense peak	FWMH Intense peak (θ_2-θ_1)	FWMH Intense peak (β) radians	Grain Size (D) nm	Average Grain Size nm
37.8	1.99	0.035	4.17	
58.5	1.26	0.022	7.20	6.32nm
32.8	1.09	0.019	7.59	

Table 2: Crystallite Size of Ni Doped ZnO Nano-Particles M_2

2 θ (deg) of the Intense peak	FWMH Intense peak (θ_2-θ_1)	FWMH Intense peak (β) radians	Grain Size (D) nm	Average Grain Size nm
36.2	1.63	0.028	5.19	
42.7	1.09	0.019	7.81	7.17nm
62.4	1.09	0.019	8.51	

The average crystallite sizes for sample M_1 is calculated as 6.32nm and for sample M_2 is 7.17nm. This crystallite size is also called particle diameter. From these results it is clear that all the synthesized samples from both the methods having size in nano-scale range.

3.3 SEM Results

The SEM results for all the samples are shown above give detailed view about the crystal structure, particle size and morphology. The SEM results shows that the particles produced through co-precipitation method are of high density, their morphology and shape of crystals shown in figure 3. Nano-particls produced through solvo-thermal method shown in figure 4. Although doping of Ni over ZnO and formation of nano-particles were confirmed by XRD but the samples were not further considered for Energy Dispersive X-ray (EDX) analysis due to SEM results. However samples were considered for optical activity through band gap analysis. The SEM results for particles produced through solvo-thermal were found better in crystal structure and morphology.

Figure 3: SEM of M_1

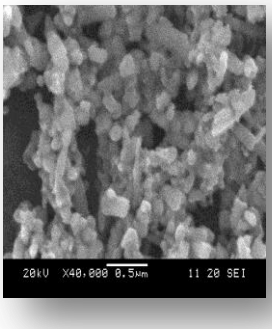
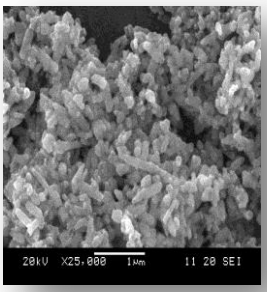
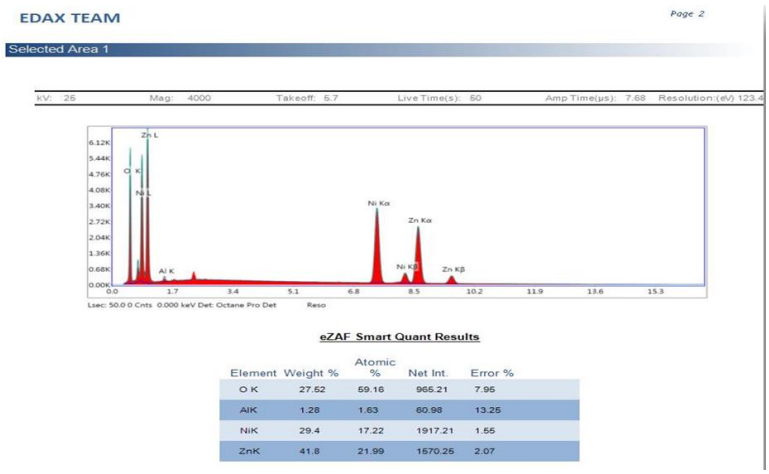


Figure 4: SEM of M_2



3.4 EDX Analysis

The extant of Ni doping in Ni-doped zinc oxide samples was checked using Energy Dispersive X-ray analysis. For this purpose M_2 (Ni-doped ZnO sample) produced from solvo-thermal method is used. The area taken for EDX analysis is given below in figure.



Element	Weight %	Atomic %	Net Int.	Error %
O K	27.52	59.16	965.21	7.95
AlK	1.28	1.63	60.98	13.25
NiK	29.4	17.22	1917.21	1.55
ZnK	41.8	21.99	1570.25	2.07

Figure 5 EDX spectra of Ni-doped ZnO nano-particles synthesized by solvo-thermal method. EDX results shown above were confirmed Ni doping in ZnO material. The element and weight percentage compositions of Ni, Zn and O in the synthesized nano-particles are given in above table. As a result, the accumulation of Ni in the ZnO system dominantly affects the optical, morphological and structural properties of ZnO.

3.5 Band Gap Analysis

Ultra Violet visible spectroscopy is used to study optical activity of Ni-doped zinc oxide nano-particles. Band gap results using Ultra Violet visible spectroscopy were obtained. UV–VIS band gap results of M_1 and M_2 are shown in Figures below.

Figure 6: Band Gap for M_1

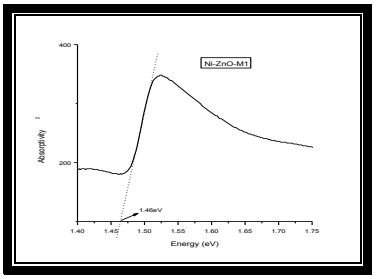
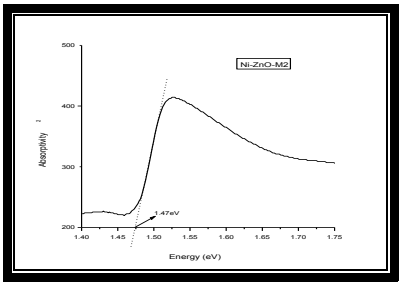


Figure 7: Band Gap for M_2



The optical band gaps are calculated by using Tauc equation as given below:

$$\alpha = \frac{C(h\vartheta - E_g^{bulk})^2}{h\vartheta}$$

Here α is absorption coefficient, C is constant, hυ is photon energy and E_g^{bulk} is band gap, respectively.

Samples which synthesized by both the methods have been successfully studied using UV-Vis. All samples have shown optical activity which is indicated by the band gap values. Sample produced through solvo-thermal method showed slight higher band gap values than the sample produced through co-precipitation method. As a whole we can conclude that Ni doping enhances the optical activity of ZnO system or we can say Ni-doped zinc oxide nano-particles shown higher optical activity than un-doped ZnO nano-particles.

3.6 Antibacterial Activity Results

After the specified time for incubation the petri plates under test were taken out from the incubator and zones of inhibition were observed under Laminar flow hood. Clear zones of inhibition were observed. Antibacterial activity of synthesized Ni-doped zinc oxide nano-particles was dose dependent. It was also clear form antibacterial activity experimental results that antibacterial activity was directly proportional to concentration. The obtained data also showed bacterial sensitivity of Ni-doped zinc oxide nano-particles against two different bacterial classes. The bacterial sensitivity is more against *Gram* positive bacteria with comparison of *Gram* negative bacteria for all synthesized samples. Ni-doped zinc oxide nano-particles were stable and their potent antibacterial properties favors them to be used as antibacterial agent against broad spectrum microorganisms.

Table 3: Zones of Inhibitions of Ni-Doped ZnO Nano-Particles against Different Bacterial Strains

Bioactive Agent		Zone of Inhibition* (Diameter, mm)	
		Escherichia Coli	*Bacillus subtilis*
Ni doped ZnO nano-particles	100µL	11±1.2	14±0.3
	200µL	13±0.9	16±0.2
	300µL	14±0.1	18±0.5
	400µL	15±0.8	20±0.6
	500µL	16±0.6	23±1.2
Cephradinen	100µL	13±0.7	18±0.8
Ampicillin trihydrate	100µL	14±0.9	20±0.5

*Inhibition zone (mm) include the diffusion assay disc diameter (4mm), which carried 50 µL from nano-particles suspension. The diameter of inhibition zones are means triplicate ± standard deviation. $p < 0.05$ when compared with negative control i.e. blank/solvent ($p < 0.05$ is taken as significant).

The MIC is the concentration at which the solution becomes turbid. Small value of MIC for to a material depicts its higher antibacterial effectiveness.

4. CONCLUSION

Ni-doped zinc oxide nano-particles were successfully synthesized by co-precipitation method and solvo-thermal methods. Characterizations of prepared samples were performed using XRD analysis. Particles were found crystalline in nature and their particle size was also found in nanometers. Extant of Ni-doping over zinc oxide was calculated using EDX analysis. The EDX analysis confirmed successful preparation of Ni-doped zinc oxide nano-particles. Furthermore, SEM analysis un-doped and Ni-doped zinc oxide nano-particles were performed which showed that well shaped geometrical nano-particles were prepared from both methods. The optical activity of synthesized samples was checked, band gaps were calculated for using UV-VIS spectroscopy. Band gap results showed that doping of zinc oxide nano-particles with Ni enhances their optical activity. Antibacterial activity for all synthesized samples was checked using ager well diffusion method. The samples had worked as effective antibacterial agents. Their activity was found more promising against *Gram* (+) bacteria as compared with *Gram* (-) bacteria. MIC for the synthesized nano-particles was proving them as future antibacterial agent. The synthesized nano-particles can be used as potential candidate for solar cells aplications. Moreover their use in nano-composite materials may also be investigated in future.

REFERENCES

Al-Harbi, T. (2011). Hydrothermal synthesis and optical properties of Ni doped ZnO hexagonal nanodiscs. *J. Alloy. Compd. 509*(2), 387-390.

Ananda, S. (2014). Synthesis and Characterization of Se-doped ZnO Nanoparticles by Electrochemical Method: Photodegradation Kinetics of Indigo Carmine Dye and Study of Antimicrobial, Antimitotic Activities of Se-doped ZnO Nanoparticles. *J. Am. Chem. Sci. 4*(5), 616-637.

Buzea, C., Pacheco, I. I., & Robbie, K. (2007). Nanomaterials and nanoparticles: sources and toxicity. *Biointerphases*, *2*(4), MR17-MR71.

Cai, A. J., Wang, Y. L., Xing, S. T., Du, L. Q., & Ma, Z. C. (2013). Tuned morphologies of DNA-assisted ZnO struggling against pH. *Ceram. Int. 39*(1), 605-609.

Hewakuruppu, Y. L., Dombrovsky, L. A., Chen, C., Timchenko, V., Jiang, X., Baek, S., & Taylor, R. A. (2013). Plasmonic "pump–probe" method to study semi-transparent nanofluids. *Appl. Opt.*, *52*(24), 6041-6050.

Hua, G., Zhang, Y., Ye, C., Wang, M., & Zhang, L. (2007). Controllable growth of ZnO nanoarrays in aqueous solution and their optical properties.*Nanotechnology*, *18*(14), 145605.

Hsu, H. C., Cheng, H. M., Wu, C. Y., Huang, H. S., Lee, Y. C., & Hsieh, W. F. (2006). Luminescence of selective area growth of epitaxial ZnO nanowires and random-growth-oriented nanobelts. *Nanotechnology*, *17*(5), 1404.

Khatoon, S., & Ahmad, T. (2012). Synthesis, optical and magnetic properties of Ni-doped ZnO nanoparticles. *J. Mat. Sci. Engine. B*, *2*(6), 325-333.

Leven, M., Berghe, D. A. V., Mertens, F., Vlietinck, A., & Lammens, E. (1979). Screening of higher plants for biological activities I. Antimicrobial activity. *Planta Medica*, *36*(08), 311-321.

Liewhiran, C., Seraphin, S., & Phanichphant, S. (2006). Synthesis of nano-sized ZnO powders by thermal decomposition of zinc acetate using Broussonetia papyrifera (L.) Vent pulp as a dispersant. *Curr. Appl. Phy. 6*(3), 499-502.

Moriga, T., Y. Hayashi, Y,. Kondo, K., Nishimura, Y. (2004). Zinc oxide, tin oxide-based transparent conductive thin film of amorphous, *J.* Vac. Sci. *and* Technol *A*, 22, 705–1710.

Raja, K., Ramesh, P. S., & Geetha, D. (2014). Synthesis, structural and optical properties of ZnO and Ni-doped ZnO hexagonal nanorods by Co-precipitation method. *Spectrochimica Acta Part A. 120*, 19-24.

Sahoo, S. K., Parveen, S., & Panda, J. J. (2007). Clinical nanomedicine.The present and future of nanotechnology in human health care. *Nanomed. Nanotechnol. Biol. Med*. *3*(1), 20-31.

Siegel, R. W. (1994). What do we really know about the atomic-scale structures of nanophase materials?. *J. Phy. Chem. Solids*, *55*(10), 1097-1106.

Wang, Z. L. (2004). Zinc oxide nanostructures: growth, properties and applications. *J. Physics: Condensed Matter*, *16*(25), R829.

Xu, M., Fujita, D., Kajiwara, S., Minowa, T., Li, X., Takemura, T., ... & Hanagata, N. (2010). Contribution of physicochemical characteristics of nano-oxides to cytotoxicity. *Biomaterials*, *31*(31), 8022-8031.

Yan, H., Johnson, J., Law, M., He, R., Knutsen, K., McKinney, J. R., ... & Yang, P. (2003). ZnO nanoribbon microcavity lasers. *Adv. Mat. 15*(22), 1907-1911.

Zhang, J., Wang, S., Xu, M., Wang, Y., Zhu, B., Zhang, S., ... & Wu, S. (2009). Hierarchically porous ZnO architectures for gas sensor application. *Crys. Gro. Degn.*, *9*(8), 3532-3537.

Zhong, Q., & Matijević, E. (1996). Preparation of uniform zinc oxide colloids by controlled double-jet precipitation. *J. Mat. Chem. 6*(3), 443-447.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

OBSERVATIONS AND RECOMMENDATIONS FOR THE WOMEN ENTERPRENEURSHIP IN INFORMATION TECHNOLOGY

DOI: 10.17261/Pressacademia.2017.596
PAP-WCTIE-V.5-2017(34)-p.247-252

Ayca Tuzmen Yildirim[1]
[1]Okan University, ayca.tuzmen@okan.edu.tr

ABSTRACT

Women entrepreneurship is a process that is being encouraged, publicized and directed by governmental and non-governmental organizations. There are many forces that allows emergence of the women entrepreneurship in IT. In this paper, I opt to discuss and comment on the journey of the women workforce in IT from corporate organizations to the workforce in women entrepreneurship. This paper introduces observations of the changes in women workforce in corporate USA and Turkey, and introduces the offerings and challenges observed in the designing, launching and running of a new IT business at Incubation centers and universities in Turkey. This paper also offers recommendations for governmental and non-governmental policies of small businesses and recommendations for the strategies of the incubation centers that would help women entrepreneurship.

Keywords: Women entrepreneurship, information technology, incubation center, academia

1. INTRODUCTION

A growing literature in entrepreneurship focus on the theory and definition of entrepreneurship, innovation (Cantillon, 1964); on the forces and motivations for entrepreneurship (Knight, 2005; Sorensen and Nanda, 2013), behaviors of entrepreneurs (Hanssan, 2012), experiences and strategies of entrepreneurs (Navale, 2013). Published research studies of the female entrepreneur have ranged from psychological and demographic studies (Sexton and Bowman-Upton, 1990; Ahl, 2003) to perceived start-up obstacles (Hisrich and Brush, 1984). While these research help us fully understand the motivation of women entrepreneurship, mechanics of entrepreneurship process; we still need to reflect upon the understanding, mechanism and strategies for women entrepreneurship in Information Technology.

This paper makes an inquiry into women entrepreneurship in Information Technology in Turkey. It focuses on the question "What are the motivations and strategies for women entrepreneurship in Information Technology?". Are women in Information Technology being stimulated by a "She can do it, why can't I?" attitude or not? The aim of this paper is to better understand the mechanics that drive the entrepreneurship process in Information Technology undertaken by women entrepreneurs.

This study is important to show the variation of research on women entrepreneurship and make an inquiry into their findings. This paper questions the motivations and demotivatios for women entrepreneurs in IT and discusses the validity of the strategies for improving the mechanics of the entrepreneurship process.

The paper is organized into three sections. In section 2, literature on the observations and research on entrepreneurship are introduced. In section 3, the inquiries into the findings of research on women entrepreneurship are made. The final section concludes the paper with recommendations and suggestions for women entrepreneurs and policy makers in Information Technology in Turkey.

2. INQUIRY INTO WOMEN ENTERPRENEURSHIP

Entrepreneurship has been defined as the willingness to take risk in the new name of an idea, spending time as well as capital on an uncertain venture to make profit (Knight, 2005; Drucker, 2006; Brush et al, 2003). Empirical studies have shown that people are drawn to entrepreneurship due to many driving forces. Sorensen (2013) suggested that significant influences on the decision to become an entrepreneur are workplace peers and social composition. If one entrepreneur is successful in venture, others are also influenced by the demonstration of the possibility of success. After seeing entrepreneurs like Steve Jobs, Mark Zuckerberg, Trabis Kalancik, Nevzat Aydin, social composition stimulates a "He can do it, why can't I?" attitude.

Even though, social composition and workplace peers are to be major driver for entrepreneurship, women entrepreneurs seem to have less such influences since most of the social composition and workplace peers seems to be coming from success stories of men.

Studies by Burton et al (2002), Wagner (2004), Dobrev and Barnett (2005), Gompers et al (2005), Sorensen and Nanda (2013), Elfenbein et al (2010), Parker (2009) provided evidence of the importance of the workplace in the entrepreneurial process. Workplace is particularly important source of entrepreneurial influence in modern societies; therefore, the workplace becomes the setting for unexpected influences, and for the serendipitous flow of information and ideas that may spark entrepreneurial activity (Nanda and Sorensen, 2010)

Even though workplace is the setting for unexpected influences, and for serendipitous flow of information and ideas that may spark entrepreneurial activity; we see less women entrepreneurs in IT sharing workplaces such as in incubation centers, teknoparks or universities in Turkey and therefore having less opportunity to increase the impact of social composition and workplace peer to the mechanics of entrepreneurship process. This further explains why certain regions show variations in rates of women entrepreneurship.

Entrepreneurs are also driven to entrepreneurship by past experience. As Hansson's study suggests, if one has faced multiple work stoppages or have been unemployed in the past, the probability of becoming an entrepreneur has increased (2012). Past experience as well as present experience in workforce is a significant driver for an increase in entrepreneurship in areas where unemployment rate is high, creation of new job opportunities is low (European Commission, 2012). Work stoppers and unemployment in Information Technology are on the rise among women tech makers in Turkey, therefore one would expect an increase in trend in women entrepreneurship in Information Technology.

Knight believed that entrepreneurs are faced with three types of uncertainty: 1) risk which is measurable statically, 2) ambiguity which is hard to measure statically and 3) true uncertainty which is impossible to estimate or predict statically (2005). Risk is seen as the probability of drawing a red ball from a jar containing 5 red balls and 5 white balls. Ambiguity is seen as the probability of drawing a read ball from a jar containing 5 red balls but an unknown number of white balls. On the other hand, true uncertainty is seen as the probability of drawing a red ball from a jar whose contents are entirely unknown.

Given Knight's classification of uncertainty (2005), entrepreneurships are believed to face true uncertainty when it comes to the creation of novel good or service, for a market that did not previously exists rather than when a venture creates an incremental improvement to an existing product or service. Innovations in IT which requires creation of novel good or service, calls for entrepreneurships in IT which are associated with true uncertainty. Therefore, most of the women entrepreneurs in IT are faced with uncertainty in their venture and therefore the question is needs to be examined is whether this is a factor demotivating women tech makers in Turkey.

Empirical studies have looked into similarities and difference in male and female entrepreneurship (Hansson, 2005). It is been observed that there are more similarities than difference between them. The start-up process has been reported similar for men and women, and women appear to have no specific difficulties or information needs (Alsos & Ljunggren, 1998; Birley, Moss, & Saunders, 1987; Dolinsky, 1993; Marlow, 1997; Nelson, 1987). Women seem to be discriminated against by banks in several studies, but the explanations appear to be mainly structural; they own the types of businesses that banks associate with higher risks (Buttner & Rosen, 1988; Buttner & Rosen, 1992).

Literature consists of research which hypothesized the family to be of special significance for women entrepreneurs, either as an obstacle or as a resource (Caputo & Dolinsky, 1998; Dumas, 1992; Holmquist & Sundin, 1990; Nelson, 1987, Ufuk and Ozgen, 2001). Balancing family and work is experienced as a problem for many women entrepreneurs, but the studies do not report if men have similar feelings (Stoner, Hartman, & Arora, 1990).

Empirical studies from Sweden found that the probability of becoming self-employed decreases with age for women, but increases with age for men (Hanssan, 2012). On the other hand, the results of research carried out in European countries, United States of America and Turkey show that women entrepreneurs tend to be older (Ecevit, 2017). This study also

showed that marriage increases the probability of a person becoming an entrepreneur. There are no research which studies the correlation of age and marriage to women entrepreneurship in Information Technology in Turkey; therefore, it is hard to comment on the influence of age and/or marriage on women entrepreneurship in Information Technology in Turkey.

Literature consists of studies which looked into the performance or underperformance of women entrepreneurs (Zapalaska, 1976; Chaganti & Parasuraman, 1996; Chell & Baines, 1998; DuRietz & Henrekson, 2000). They questioned if women entrepreneurs do at all "possess the characteristics required for effective performance as entrepreneurs". These led to the development of polarized measuring instruments (Buttner & Rosen, 1988; and Fagenson & Marcus, 1991) which had a masculine and a feminine scale. The items in the masculine scale were “competitive, active, independent, able to make decisions, does not give up easily, feels very superior, self-confident”, and “stands up well under pressure”. The feminine scale was comprised of “emotional, understanding, warm, able to devote oneself completely to others, gentle, helpful to others, kind”, and “aware of others’ feelings”. The studies that look for psychological differences between men and women entrepreneurs, however, very few, if any, differences between men and women entrepreneurs are found.

3. DRIVING FORCES AND STRATEGIES FOR WOMEN ENTREPRENEUSHIP

Past experience in previous workplaces is the motivator for women entrepreneurship. However, there are few researches documenting whether these experiences are pleasant or unpleasant experiences which are leading women into the venture of creating their new business. There is the hypothesis that most of these experiences are unpleasant therefore women are more inclined towards creating their own business rather than searching for better workplace experience.

There are also few studies that focus on how many of women use their past experience in information technology into the creation of new IT business. Since innovations in IT calls for entrepreneurships which is associated with true uncertainty; one can argue whether women entrepreneurs opt to undertake new business with high uncertainty or one that requires less risk and ambiguity. The practical thinking will suggest that past experience in one area of expertise should be in better use in the creation of a new business in the same area of expertise. However, we still see vast amount of women entrepreneurs coming from information technology but moving on to different areas of business such as merchandising and service type of businesses. The question that comes into concern is "what factors demotivates women going into information technology business even though their past experiences lies in IT?" . Does having family, age, capital, culture, governmental or institutional policies play an important role in this? If so, how can we reshape these influencers so that more women can create new business in Information Technology in Turkey?

Certain regions are more prone to influence innovation in Information Technology. Since Silicon Valley is a great example for such observations, many tech maker aims to start their venture in this regions. All around the worlds and well as in Turkey, similar hubs are planned and are being introduced in order to attract entrepreneurs. As of December 2015, there are 63 teknokent are approved and out of 49 are operational in Turkey. There are 3744 firms in these centers and 39% of them are in software development; 19% in IT, 7% in electronics, %5 machine manufacturing, and 40% in medical, chemistry, food, automotive, defense business (TGBD, 2017). However, there is no research showing how many women tech makers are drawn into such hubs in information technology and how many of them are admitted as business owners.

Many strategies have been taken by entrepreneurs in their venture in tackling uncertainty, risk and ambiguity. Since balancing family and work is experienced as a problem for many women entrepreneurs (Stoner, Hartman, & Arora, 1990) and thus, strategies have been recommended for formulizing the family and new business equation. Caputo and Dolinsky (1998) recommends their government to supply micro-loans for women business starters, since home-based entrepreneurship makes it possible to care for children and run a business at the same time, thereby saving taxpayers money for dependent care. A more Scandinavian approach calls out for public childcare in which case men and women can participate in the labor force and provide childcare on equal terms. The recommendation of this paper would be the combination of both approaches for supporting women tech makers in Turkey.

Institutionalization of support systems for women are common throughout Europe (Wegate, 2017), USA (Women Entrepreneurs Organization, 2017), Canada (Business Women in International Trade, 2017), Turkey (TOBB Istanbul Kadin Girisimciler Kurulu, 2017) and across regions (WeMena, 2017; Women Tech Makers, 2017). Supportive measures include such as:

- promotions of women entrepreneurs
- support networking among female entrepreneurs, potential female entrepreneurs and support organizations
- award women entrepreneurs
- provide focused trainings and workshops
- mentorship by men and women for women entrepreneurs
- dissemination of information on the services provided by other entrepreneurship supporting funds and foundation

- providing free consultations to business in their starting phase,
- scholarships for women entrepreneurs
- providing the starting businesses with a package of the related materials,
- informing businesses of the measure of the support systems to entrepreneurship
- allowing access to capital through micro lending, seed capital, and angel investing
- increasing visibility, scale and reach of existing support programs

4. RECOMMENDATIONS AND CONCLUSION

We see leadership positions taken by women in business and universities in Turkey; however one thing we know from observations is that the numbers are insufficient and that women in some areas are waiting for companies to give them the leadership positions they deserve (Compass, 2015). We need more women techmakers in Turkey. It seems clear that most, if not some of the essential components of an effective support infrastructure for women entrepreneurship in Information Technology needs to be found in Turkey. Therefore, this section focuses on the question "What are the major strategies women entrepreneurs can use in order to be successful?"

We would see more women entrepreneurs in Information Technology by promoting their stories and strategies. Advocating women entrepreneurs such as Coco Chanel, Joy Mangano, Martha Steward, Sara Blakely as well those in IT such as Peng Lei, Gina Bianchini, Cher Wang stimulate a "She can do it, why can't I?" attitude. Through organizations such as Girlsintech, WeMena.org, Womentechmakers among Turkish women in IT and allowing unexpected influences, and serendipitous flow of information, ideas and collaboration in IT may spark more women entrepreneurial activity in Turkey.

In order to increase innovation in information technology, workplaces such as incubation centers, teknoparks and universities should bring together more women and foster their collaboration in a single workplace. By having more women peers in a workplace and their social compositions, we can expect the emergence of further innovations by women in Information Technology.

Construction of an ecosystem which holds the ingredients needed to support women techmakers in IT would require support mechanisms for balancing family life and work. Incubation centers, teknoparks and universities which require spending a certain amount of time at a certain location and measuring their success in terms of how many hours they spend in a certain location does not allow fostering of innovation and construction of work and life balance.

Studies show that innovation, creativity and entrepreneurship does not happen miraculously by spending a certain amount of time at a certain location (Rehman and Roomi, 2012; Walker et al, 2008). Entrepreneurship through home-based business ownership is recommended as a potential solution to the inter-role conflict experienced by women attempting to balance dual work and family roles (Walker et al. 2008). Incubation centers, teknoparks and universities which do not impose time-based access but instead support women entrepreneurs in their business venture to develop, test, market and improve customer and partner relationships are needed by women techmakers in Turkey. However, access to an entrepreneurial ecosystem on a need basis should also be provided to women entrepreneurs by allowing them access to incubation center, teknoparks and university research centers.

We need further studies showing which regions are attracting women entrepreneurs in information technology. There is no research in Turkey, showing how many women tech makers are drawn into teknoparks and incubation center in information technology and how many of them are admitted as business owners. We need to understand whether acceptance criteria for entering women entrepreneurs are seen as obstacles in such hubs. Further, we need to study whether offerings in such hubs are geared towards the needs and expectations of the women entrepreneurs. Improvements on acceptance criteria and offerings of the IT regions, which are not able to attract women tech makers and/or not able to support them in their venture, should be defined and introduced.

A commitment to increase the share of government economic development resources that is made available to women entrepreneurship in Information Technology is needed. A recognition that apart from legislator and regulator, governmental organizations should have the three principal role in the facilitation of entrepreneurship and small business development - as advocate, investor and enabler. In order to play the role of the investor, the governmental and non-governmental agencies will need to increase the scale and scope of the investment pool for women entrepreneurs. The enabler role requires that government help entrepreneurs navigate the regulatory process during the start-up of new business as well as while running it. The advocator role requires that women entrepreneurship in Information Technology is advocated to women in all ages, regions, status and income. Governmental and non-governmental offerings and support to women entrepreneurs should not be limited to only new graduates at certain regions, but should be extended to all women with or without experience, with or without formal training, with or without capital, with or without physical access to a specific region or location.

REFERENCES

Ahl, H. J. 2003, "The Scientific Reproduction of Gender Inequality: A Discourse Analysis of Research Articles on Women's Entrepreneurship".

Alsos, G. A. and Ljunggren, E., 1998, "Does the business start-up process differ by gender? A longitudinal study of nascent entrepreneurs", In P. D. Reynolds, W. D. Bygrave, S. Manigart, C. M. Mason, G. D. Meyer, N.M. Carter, & K. G. Shaver (Eds.), Frontiers of entrepreneurial research, pp. 137-151). Boston, MA: Babson College

Beatty, J. 1998, "The World According to Peter Drucker", pp. 163.

Birley, S., Moss, C., & Saunders, P., 1987, "Do women entrepreneurs require different training?", American Journal of Small Business, vol. 12, no. 1, pp. 27-35

Brush et al, 2003, "Women Entrepreneurs: Moving Front and Center: An Overview of Research and Theory", Retrieved on March 22, 2017, http://www.unm.edu/~asalazar/Kauffman/Entrep_research/e_women.pdf

Burton, M. D., J. B. Sørensen, C. Beckman. 2002. "Coming from good stock: Career histories and new venture formation", Res. Social. Organ. 19, pp. 229–262

Buttner, E. H. 2001, "Examining female entrepreneurs' management style: an application of a relational frame. Journal of Business Ethics, vol. 29, no. 3, pp. 253-269.

Buttner, E. H., & Rosen, B., 1992, "Rejection in the loan application process: Male and female entrepreneurs' perceptions and subsequent intentions", Journal of Small Business Management, vol. 30, no. 1, pp. 58-65.

Cantillon, R. 1964, "Essoi Sur Lo Noture Du Commerce En Gindral", Augustus M. Kelley: pp. 29.

Caputo, R. K., & Dolinsky, A., 1998, "Women's choice to pursue self-employment: The role of financial and human capital of household members", Journal of Small Business Management, vol. 36, no. 3, pp. 8-17

Chaganti, R., & Parasuraman, S., 1996, "A study of the impact of gender on business performance and management patterns in small businesses", Entrepreneurship Theory and Practice, vol. 21, no. 2, pp. 73-75

Chell, E., & Baines, S., 1998, "Does gender affect business' performance'? A study of micro-businesses in business services in the U.K", Entrepreneurship and Regional Development, vol. 10, no. 2, pp. 117-135

Dobrev, S. D., Barnett, W. P., 2005, "Organizational roles and transitions to entrepreneurship", **Acad. Management** J, vol. 48, no. 3, pp. 433–449.

Dolinsky, A. L., Caputo, Pasumarty, R. K., & Quazi, H., 1993, "The effects of education on business ownership: A longitudinal study of women", Entrepreneurship Theory and Practice, vol. 18, no. 1, pp. 43-53.

Druker, P., 2006, "Innovation and Entrepreneurship: Practice and Principle" Routledge Classics.

Dumas, C., 1992, "Integrating the daughter into family business management", Entrepreneurship Theory and Practice, vol. 16, no. 4, pp. 41-55.

DuRietz A. & Henrekson, M., 2000, "Testing the female underperformance hypothesis", Small Business Economics, vol. 14, pp. 1-10

Ecevit, Y. "Research on Women Entrepreneurs", Retrieved March 22, 2017, , https://assets.garanti.com.tr/assets/pdf/tr/diger/K.G.A_eng_dijital_v_1.pdf

Elfenbein, D. W., B. H. Hamilton, W. B. and Zenger, T. R., 2010, "The small firm effect and the entrepreneurial spawning of scientists and engineers", Management Science. 56 (4) 659–681

European Commission, "Statistical Data on Women Entrepreneurs in Europe-2014", Retrieved March 22, 2017, http://ec.europa.eu/DocsRoom/documents/7481/attachments/1/translations

Evaluation of Women's Enterprise Initiative (n.d.). Retrieved March 22, 2017, https://www.wd-deo.gc.ca/eng/18274.asp

Fagenson, E. A., & Marcus, E. A.,1991, "Perceptions of the sex-role stereotypic characteristics of entrepreneurs: Women's evaluations", Entrepreneurship Theory and Practice, vol. 15 , no. 4, pp. 33-47

Girls in Tech (n.d.) Retrieved March 22, 2017, from http://girlsintech.org/#programs

Gompers, P., J. Lerner, D., 2005, "Entrepreneurial spawning: Public corporations and the genesis of new ventures", J. Finance, vol. 60, No. 2, pp. 577–614.

Hanssan, A. 2012, "Tax Policy and Entrepreneurship: Empirical Evidence from Sweden", Small Business Economics.

Hisrich, R. D. and Brush, C., 1984, "The woman entrepreneur: Management skills and business problems", Journal of Small Business Management, vol. 22, no. 1, pp. 30-37

Holmquist, C., and Sundin, E.,1990, "What's special about highly educated women entrepreneurs", Entrepreneurship and Regional Development, vol. 2, pp. 181-193.

Knight, F. H. 2005, "Risk, Uncertainty and Profit", Cosimo: pp. 22-40.

Marlow, S., 1997, "Self employed women: New opportunities, old challenges", Entrepreneurship and Regional Development, vol. 9, no. 3, pp. 199-210.

National Association of Women Business Owners 2007 Survey of Women Business Owners. http://businesscommunicationnetwork.com/wp-content/plugins/BNet/cache/1312925017917.pdf

Navale, A. B. 2013, "Developing Entrepreneur Skills for Corporate Work", Research Directions, vol 1, no. 4, pp.

Nelson, G. W., 1987, "Information needs of female entrepreneurs", Journal of Small Business Management, vol. 25, no. 1, pp. 38-44.

Parker, S. C. 2009. "Why do small firms produce the entrepreneurs?", J. Socio-Economic. vol. 38, vol. 3, pp. 484–494.

Rehman, S. and Roomi, M. A. 2012, "Gender and Work-life Balance: A Phenomenological Study of Women Entrepreneurs in Pakistan", Journal of Small Business and Enterprise Development, vol. 19, no. 2, pp.209 - 228.

Sexton D. L. and Bowman-Upton, N., 1990 "Female and Male Entrepreneurs: Psychological Characteristics and Their Role in Gender-related Discrimination". Journal of Business Venturing, vol 5, no 1, pp 29-36.

Sorensen, J., Nanda, R. 2013, "Workplace Peers and Entrepreneurship", Management Science, vol 56, no . 7, pp. 1116-1126.

Stoner, C. R., Hartman, R. I., & Arora, R.,1990, "Work-home role conflict in female owners of small businesses: an exploratory study", Journal of Small Business Management, vol. 29, no. 1, pp. 30-38

TGBD, 2017, "Turkiyede Teknoparklar", Retrieved March 22, 2017, http://www.tgbd.org.tr/WebContent/WebContent/4707

TOBB Istanbul Kadin Girismciler Kurulu, Retrieved March 22, 2017, http://www.ikgk.org.tr/

Ufuk, H., & Ozgen, O., 2001, "Interaction between the business and family lives of women entrepreneurs in Turkey" , Journal of Business Ethics, vol. 31, no. 2, pp. 95-106.

Wagner, J. 2004, "Are young and small firms hothouses for nascent entrepreneurs? Evidence from German micro data", Appl. Econom. Quart. vol. 50, no. 4, pp. 379–391.

Walker, W. et al., 2008, "Women and worklife Balance: Is Homebased Business Ownership the Solution?", Equal Opportunities International, vol. 27, no. 3, pp.258 - 275.

WeGate, (n.d.). Retrieved March 22, 2017, https://www.wegate.eu/who-we-are

Women Entrepreneurs MENA (n.d) Retrieved March 22, 2017, from http://www.we-mena.org/

Women TechMakers, (n.d.). Retrieved March 22, 2017, https://www.womentechmakers.com

Zapalska, A.M., 1997, "A profile of women entrepreneurs and enterprises in Poland", Journal of Small Business Management, vol. 35 , no. 4, pp. 76-82.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

TURKEY ON THE PATH OF ESTABLISHING KNOWLEDGE ECONOMY: INNOVATION, ICTs AND EDUCATION

DOI: 10.17261/Pressacademia.2017.597
PAP-WCTIE-V.5-2017(35)-p.253-261

Burcu Sakiz[1]
[1]Istanbul Aydin University, burcubaydar78@gmail.com

ABSTRACT

Knowledge-based economy is re-defining enterprises, re-formatting working styles, empowering individuals and re-shaping the links between education and business tasks in today's globalized world. Emerged communication and information technologies especially invented in six decades ago is still changing the nature of business and the economy dramatically. It is obvious that transition to knowledge-intensive business models become popular since late 1950s. Also it is clear that advent of the knowledge economy shapes and changes the rules for enterprises that organize their activities in the global market places. Additionally intellectual capabilities become more important than physical inputs which was very important for the first machine age. The application of knowledge is one of the main sources of growth in the global economy and it becomes the strategic factor for the success. Having sufficient national production, sustainable growth, strong economy and modern technologies to create a welfare society is crucial for any modern nation. This study presents the concept of the knowledge economy and it's framework in this respect. Developed by World Bank "Knowledge Assessment Methodology" is one of the important methodologies in order to measure the progress of countries towards having a knowledge based economy. It has 4 pillars including 83 structural and qualitative variables and 12 knowledge indicators. World Bank's four Knowledge Economy pillars are economic and institutional regime, education, innovation, and Information and Communication Technologies (ICTs) developed for countries to do fundamental assessment of countries' readiness for the knowledge economy and help them the transition to a Knowledge-based Economy. Also in this paper the analysis of knowledge economy for Turkey especially the role of innovation, ICTs and education pillars are examined.

Keywords: Knowledge economy, knowledge assesment methodology, innovation, ICT, education.

1. INTRODUCTION

In the knowledge society that we are in, service/knowledge technologies have come into prominence. Knowledge became most valuable asset when it is compared to land, physical labor, machines and factories. Thomas A. (1997) asserts that wealth is production of knowledge in today's world. According to him, "knowledge and intellectual capital which is organized knowledge that can be used to produce wealth is becoming corporate America's most valuable asset and can be its sharpest competitive weapon. The challenge is to find what you have-and use it".He demonstrated how knowledge has become the most important factor in economic life instead of natural resources, machinery, or financial capital. In the knowledge society that we are in service/knowledge technologies have come into prominence. Use of computers and communication technologies especially in academic and business world and in personal life has increased year by year. Therefore major changes have been observed in social and economic structure especially in developed countries. Knowledge Society has witnessed to alteration and transformation of populations. Knowledge brings speed and effectiveness for production and it plays the key role for the world's economy. In the 21st century, the knowledge-based economy has become the major trend for international societies (George, et al., 2008).

The change towards a knowledge-based economy is happening on a global scale, a transformation is taking place in all industrialized economies and many developing economies are also aspiring to reach this target. Many studies in the literature focus on the knowledge-based economy. Gür's (2001) study about constructing a new framework which is to be utilised for the assessment of possible strategies which can be applied during the transition to a knowledge-based economy. He developed mixed integer programming model in order to determine the required levels of human resources and information and communications technology investments for given levels of R&D investment of the country that

determines most significant characteristic of the phenomenon. The results indicate that Turkish government should increase R&D to considerable levels in order to trigger the transition to a knowledge-based economy. Second important finding is transformation towards a knowledge-based economy with an inefficient innovation system may require considerable amounts of additional resources compared to transforming with a more efficient system.

Umut's (2001) study about constructing a new framework that utilizes assessment of possible strategies can be applied during the transition to a knowledge based economy. A mixed integer programming model developed by him to determine the required levels of human resources and information and communications technology investments for given levels of R&D investment of the country. Results indicated Turkish government should increase R&D to considerable levels in order to trigger the transition to a knowledge-based economy.

Salduz (2005) studied that examine Turkey's position as a candidate country in adaptation to the Lisbon Strategy. She gives general information concerning knowledge economy and EU's activities aiming to reach Lisbon strategy and EU-Turkey comparison with using knowledge economy indicators also she examined Turkey's adaptation process. Study reveals that Turkish economy has many disadvantages both for knowledge production and using information technologies and economy needs reconstruction to close this gap and also for knowledge economy that reconstruction requires renovations in science and technology policies, education policy and industrial policy.

Arıkan (2008) examines the dynamics of the innovationy with 122 Turkish firms from a wide variety of sectors, through exploring the characteristics of organizations in terms of their openness to innovations. Relating this innovative culture both to the systems organizations utilize to enhance innovations and to the actual innovations implemented in his thesis. The results show that innovation leads to better business performance. An organization's market orientation, learning orientation, and entrepreneuiral orientation lead to more openness to innovations.

Tuncay (2008) studied the place and importance of knowledge economy in the economic growth process. an augmented Cobb-Douglas production function, which is an approach based on neo-classical theory is used in investigating the effect of knowledge on the economic growth of Turkey and thus the direction and size of the relation between investments in knowledge communication technology and economic growth in Turkey during years between 1980 to 2006. The results showed that the investments in communication technology knowledge have a positive effect on economic growth for Turkey.

Senem (2011) studied discussions on clarifying of relationship with knowledge economy and growth and productivity paradox by linking the presence of threshold effects. "being over of a specific level of some factors that countries have, leads to differences in the relationship with knowledge economy and growth" is the hypothesis. The panel data set is analyzed consisting of 39 countries with different level of development with fixed effects method for years 1995, 2000, 2007, 2008, 2009. In growth model which is per capita income is dependent variable; Knowledge Economy Index (KEI), R&D, capital, employment ratio are explanatory variables examined for the threshold effects for KEI and R&D variables as well as the numerical and statistical chancing of KEI coefficient is examined for indexes containing human capital, technological knowledge and innovation capacity, information and communication technologies (ICT) .

Memişoglu (2012) studied major factors of knowledge-based economies between years 2000-2010 based on World Bank Knowledge Assessment Framework, on economic performance indicators such as Gross Domestic Product(GDP), GDP per capita and economic growth rate in Brazil, Russia, India, China, South Africa and Turkey (BRICST). Secondary education and ICT infrastructure are found to be important infrastructure factors affect GDP per capita positively for BRICST countries. Country's ability to innovate, benefit from ICT and enhance economic performance can be affected by accumulation of educated people in Research and Development. Also personnel number of R&D which is an indicator for innovation potential has positive influence on the GDP. The major things affect economic performance found as ICT infrastructure expansion together with educated R&D personnel in the BRICST countries.

Vuslat(2014) studied reviews institutional transition in particular universities around the world by means of knowledge-based economy. The aim of her study is to comprehend the institutional transition by means of knowledge-based economy in general and transition of universities is investigated deeply among institutions as an important part for innovation and knowledge in 4 Turkish Universities as Istanbul Technical University, Boğaziçi University, Sabancı University, and Istanbul Bilgi University. Also some universities from the world are examined about their management, productivity, and academic change.

Işık (2012) studied effects of R&D, Innovation, Patent, Information Technologies for Turkey in years between 1990-2010. The competitive advantage factors in knowledge economies are analyzed by various methods. As a result of theoretical and empirical analyses, it is determined that the factors of competitive advantage in knowledge economies, contribute to the economic growth and the development of our country's position in the world in respect to knowledge economy.

2. KNOWLEDGE BASED ECONOMY

The knowledge economy can be defined as "in the use of knowledge to generate tangible and intangible values. Technology helps to transform a part of human knowledge to machines and tools. This knowledge could be used by decision support systems and generate economic values in various fields". Besides, emerging technologies and changes in the production structure have changed the demand structure significantly. (Sadık, 2000). As Charles Leadbeater (2000) states, more of the value of manufactured products will come from the software and intelligence that they embody, and more of what we consume will be in the form of services in the new economy. The knowledge across all sectors, content of products and processes is ascending. Everything is getting smarter from computers and photocopiers to cars and corn.

It is emerged as a result of intensive knowledge in economic activities, globalization of economic activities and improvements in information and communication Technologies (ICTs). OECD(Organisation for Economic Cooperation and Development) defines knowledge economy as: "Economy in which knowledge is being used in distribution and production."(OECD, 1996). According to OECD Knowledge-Based Economy Report (1996), the new driver of productivity and economic growth was defined as knowledge. The new global economy is very distinctive in terms of its strategies of flexible production organized around principles of knowledge based economy that knowledge is accepted as the main drive of the economic growth.

Susan(2005) sets the essence of the knowledge economy as:

- the balance between knowledge and resources (labor and capital) has shifted toward knowledge;
- securing long term economic growth will be much more dependent on knowledge;
- education will play a critical role in economic growth;
- education systems will need to respond in new ways to the demands of knowledge economy.

Economic incentives and institutional regime, policies and information and communications technologies (ICTs) are pillars of the knowledge-based economy. Access to networking is essential in acquiring and disseminating knowledge and the Internet is one of the key drivers of ICT, resulting in new approaches to doing things. We are witness to arising of new important definitions and drivers for economic performance. Such as information society with increasing communication and computer networks, learning economy with the need of workers to acquire a range of skills and to steady adapt these skills, and national innovation systems with the raising importance of knowledge and technology diffusion requires better understanding of knowledge networks are defined in knowledge-based economy (OECD, 1996). There has been a knowledge explosion because of new inventions especially in the last five decades by growth in technological researches and in the usage of the technical tools and equipments.

Knowledge-based economy also leads collaboration opportunities for information and communication technologies to conceive high qualified products with low costs (European Commission, 2001). Economic model in which knowledge comes into prominence, drives existing business processes and way of doing works to be more qualified (Oytun, 2013). As a matter of fact, any kind of production is knowledge-based. Economy's main elements which are production, consumption, distribution and their relationships also new market structure have been restructured based on knowledge. Consumers began taking goods and services faster, without being limited by time and location. Researchers can analyze consumer behaviors better and more accurate with more datum in digital environments. Barriers to market entry and exit reduce and information becomes a function of competition (Oğuz, 2005). Since barriers have been removed, producing innovative gods and services have become a must. The economic effects of physical distances, geographical differences and the cost of access to information have been decreased result of using ICTs. Moreover, new startups costs are declining and the advantage to compete in new markets is increasing. Despite financial capital has accepted as a formerly scarce resource, in today's world qualified human resources become scarce resources (Kamil, 2007).

Since knowledge has no fixed capacity, shows more rapid change in that sense and tends to be obsolete in a short period of time, it is difficult to calculate knowledge's market value. Additionally, knowledge is difficult component to quantify. The first attempt for the assesment of knowledge and development of a framework to interpretation of data related to science, information, technology, communication and innovation has been made by OECD. The OECD report notes that —OECD countries continue to evidence a shift from industrial to post-industrial knowledge-based economies. Here productivity and growth are largely determined by the rate of technical progress and the accumulation of knowledge. Of key importance are networks or systems which can efficiently distribute knowledge and information. Learning on the part of individuals and firms is crucial for realizing the productivity potential of new technologies and longer-term economic growth. Many other institutions like World Bank, Eurostat Eurostat, UNESCO, International Telecommunication Union and United Nations have introduced a range of indicators in order to analyze country's potential for knowledge and knowledge based economy. They have been built several indexes. Well-known institutions which develop these indexes are The World Bank, Harvard

University International Development Center, McConnell International, The Economist Intelligent Unit, UNCTAD, The United Nations Development Program, World Economic Forum and The Mosaic Group (Leila, et.al, 2007).

2.1. Pillars of Knowledge Economy

In 1999 the World Bank developed Knowledge Assessment Methodology (KAM) which is a method aims to show opportunities to countries and identify problems may them face towards transition to knowledge economy. KAM is designed to supply basic assessment of countries' readiness for the knowledge economy, and identifies sectors or specific areas where policymakers may need to focus more attention or future investments. It has an online interactive tool that produces the Knowledge Economy Index (KEI)–an aggregate index representing a country's or region's overall preparedness to compete in the Knowledge Economy (KE). Knowledge Economy Index indicates a country's or a region's overall level of development in terms of knowledge economy. KAM has 69 structural & qualitative variables on 4 pillars and normalized from 0 (worst) to 10 (best) for 100 countries. The unique strength of the KAM lies in its cross-sectoral approach that allows a holistic view of the wide spectrum of factors relevant to the knowledge economy. These factors are briefly summed up as follows: Information and Communication Technologies: Appropriate business environment that the knowledge economy can be developed in; R&D and Innovation; Education

3. ON THE PATH OF ESTABLISHING KNOWLEDGE ECONOMY AND TURKEY

3.1. Turkey and World Knowledge Economy Comparisons

The underlying reasons behind differences in countries' long-term growth performances are grouped under four main headings such as investment (fixed capital, and information and communication technology sector), non-formal education, innovation and structural change. According to our study, relationship between the Knowledge Economy Index (KEI) and GDP examined and results are positive. According to the our regression test depicted below, the correlation between KEI and economical performance is about 85%. GDP and KEI relations, Knowledge Economy Index of Turkey pillars figures and KEI table are given below.

Figure 1: GDP per Capita & KEI Relation

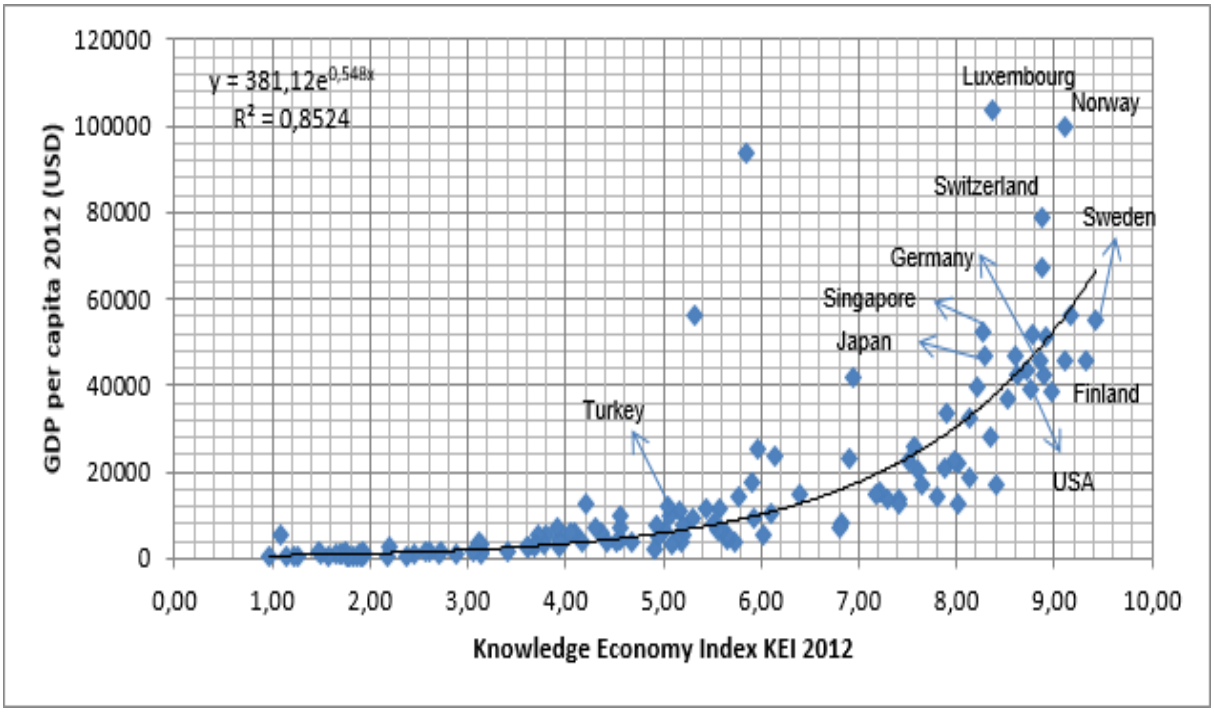
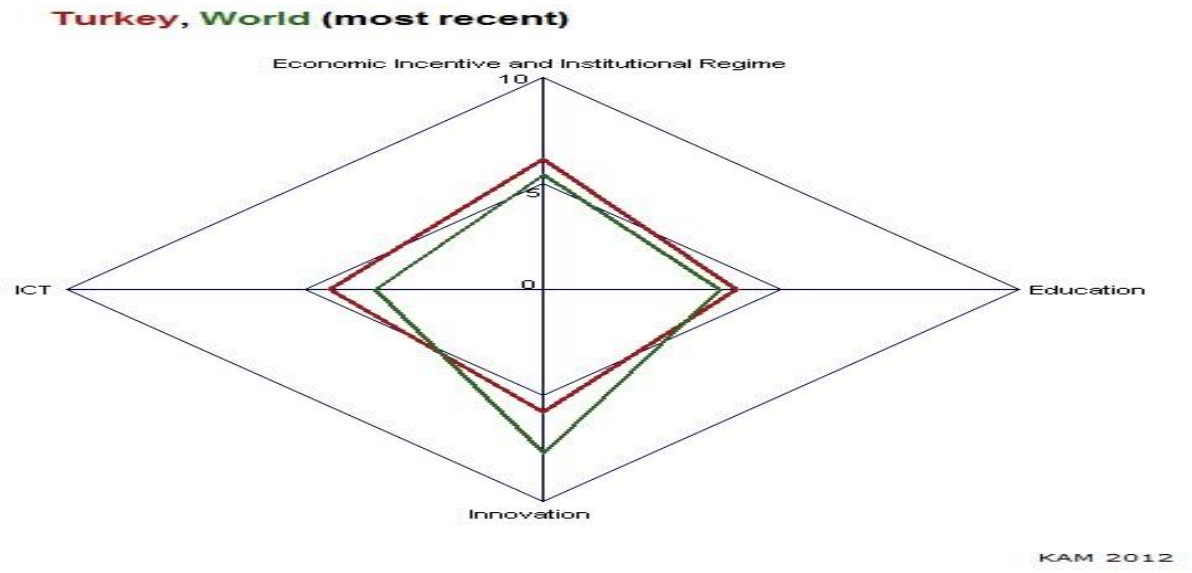


Figure 2: Turkey and World KEI



According to "Informatics for Breakthrough, Turkey for Economics Information and Communication Technology Sector-Breakthrough Strategy 2023" report, while Turkey's competitiveness in the information and communication sector (ICS) is poor, some other developing countries' competitiveness is high. While South Korea, Singapore and Brazil don't show any alteration, in China and India there has been a significant increase in competitiveness in the last three years. ICS's productivity in Turkey is four times of total private sector productivity. Despite this importance, it is seen that contribution that ICSs should make in increasing productivity in Turkey hasn't occurred. A conclusion shouldn't be drawn like there is no affect of ICS in increase in productivity. The underlying reason of this is ICS' small share in whole economy and this small is decreased rather than increase in the examined period. Taking into account all variables of Turkey's KI and KEI ranks, results are: fixed capital investment and information and communication sector investments are not enough; Research & Development expenditures fall behind the fast developing countries, but increase in production efficiency by innovating could be achieved; develop the educational level of human capital not achieved, tendency to high-tech products in the structure of production remained limited. In below table knowledge related indexes are given for World and Turkey.

Table 1: Turkey and World knowledge related indexes

Index	World	Turkey
Knowledge Economy Index	5,12	5,16
Knowledge Index	5,01	4,81
Economic Incentives and Institutional Regime	5,45	6,19
ICT	3,58	4,5
Innovation	7,72	5,83
Education	3,72	4,11

3.2. ICTs, Innovation and Education Pillars

In the 21st century, we are in the midst of "knowledge revolution" that increased importance of education, innovations and also updated skills for sustainable economic performance and improvement. The knowledge economy is transforming the demands of the labor market in economies throughout the world. In industrial countries, where knowledge based industries are expanding rapidly, labor market demands are changing accordingly (World Bank, 2003: 1).

ICT includes a dynamic information infrastructure-ranging from radio to the internet-is required to facilitate the effective communication, dissemination and processing of information. A dynamic information infrastructure is needed to facilitate the effective communication, dissemination, and processing of information. information infrastructure that facilitates the communication, dissemination, and processing of information and technology. The increased flow of information and

knowledge worldwide reduces transactions costs, leading to greater communication, productivity and output. Today's global competitive environment, becoming successful and increasing the productivity will be possible not only by structure-process and technological changes but also by using knowledge and utilizing of ICTs(Information and Communication Technologies) especially teaching and developing of high quality human sources and by preparing them for the new concepts. ICTs lower the costs of various aspects of knowledge activities. How intense countries use their knowledge and ICTs and contributions do these technologies becomes very important. In this context, Turkey has been reconstructing certain processes and policies to keep pace with alterations and to be competitive benefiting from advantages of the process (Eda, 2011).

Turkey would be count as to be a young industrial country by the courtesy of policies such as protectionism and incentives. Turkey owes partial change to the law of diminishing returns. Law of diminishing returns, gives developing countries a chance to learn producing goods with high added value. However, law of increasing returns is valid in knowledge society. Thus, specialization in the production of goods with high added value and making production more efficient through knowledge are required (Sadık, 2000). According to the "Information and Communication Technologies toward 2023 Goals" report of International Investors Association, it is expected that impact of 1 unit of growth in information and communication sector will be 1.8 units of growth in whole economy. In 2011, ICT sector size at the global level has reached $4.1 trillion. And the size of the sector in Turkey is estimated to be $ 30, 3 billion. Turkey's country's share in global information and communication sector remaining at 0.75% indicates the sector's growth potential. When the sub-sectors of the market in Turkey are analyzed, it is seen that the growth potential of ICT industry whose share is 0.4% in global market, is much higher. In case of reaching targeted 8% share of ICTs in $2 billions GDP, merely to be provided contribution through total factor productivity growth is estimated to be over $71 billion (YASED International Investors Association, 2014).

An educated and skilled population that can create and use knowledge and an effective innovation system consisting of research centers, universities, think tanks, and other organizations that can create knowledge are two drivers for development. Educated, creative and skilled people and effective national innovation system are two important factors. For innovation system, KAM uses three variables: R&D per million population, patent applications granted by the US Patent and Trademark Office (USPTO) per million population, and scientific and technical journal articles per million population. In summary KAM's innovation pillar aims to measure country's innovation system—firms, research centers, universities, think tanks, consultants, and other organizations—must be capable of tapping the growing stock of global knowledge, assimilating and adapting it to local needs, and creating new technology.

According to Innovation Union Score Board which is used to measure innovativeness by Europe Union Commission and is compiled through 25 indicators, even Turkey displays an increasing performance since 2006; she is still between countries that categorized as low in terms of innovation. In below figure , rankings of Turkey and other countries can be seen. Also we get application data from European Patent Office to compare european countries with Turkey. Turkey still has a lot of potential to grow its performance compared to its population. Moreover according to Turkish Patent Institute's data, Turkey's patent application numbers are increasing since 20 years.

Figure 3: Innovation Union Scoreboard 2015 (European Commission).

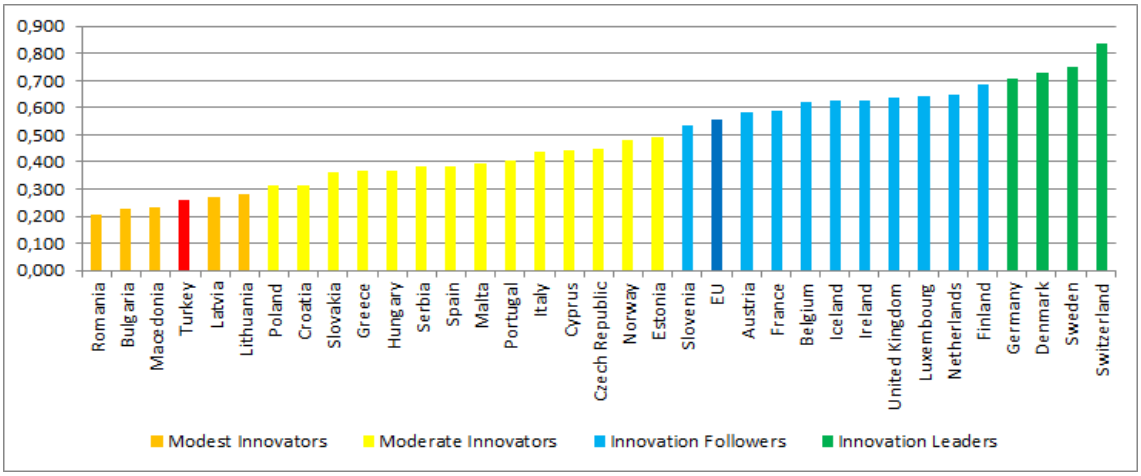
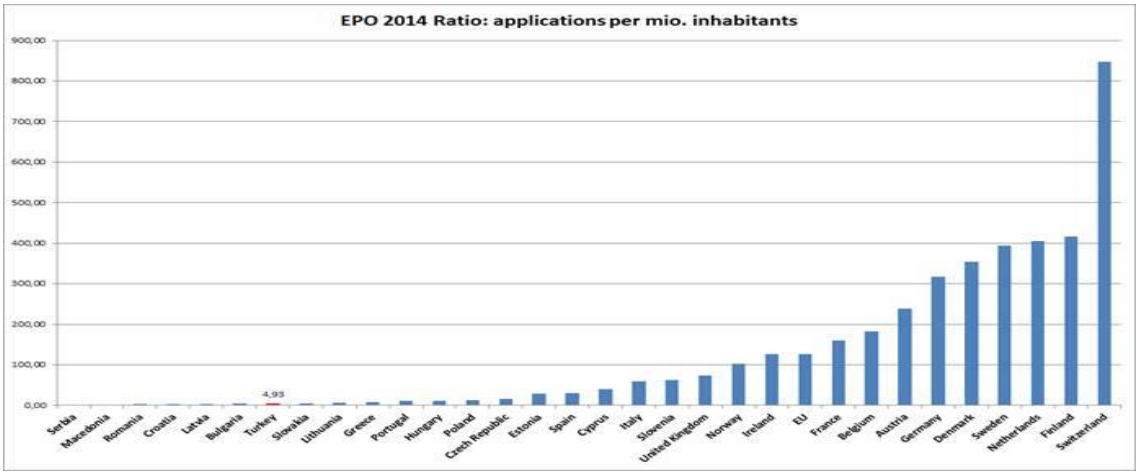


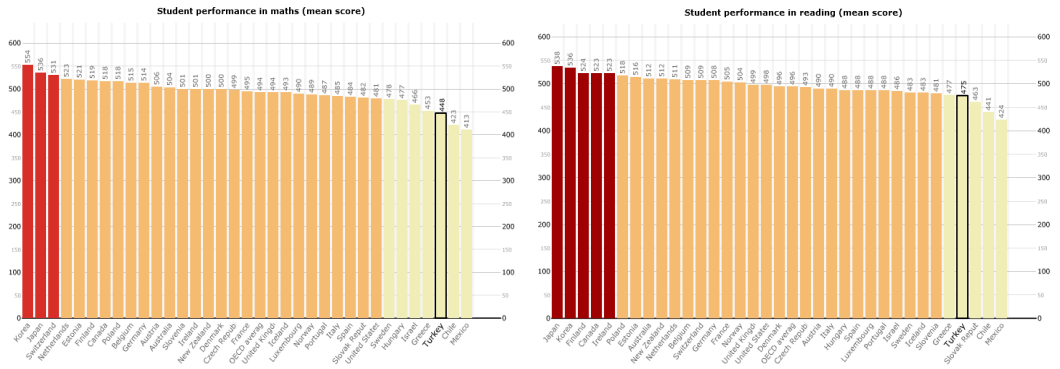
Figure 4: European Patent Office applications per mio. inhabitats

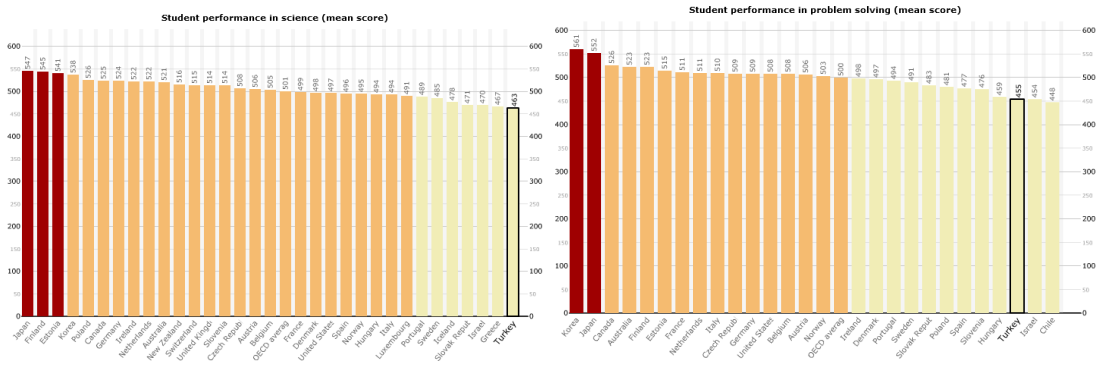


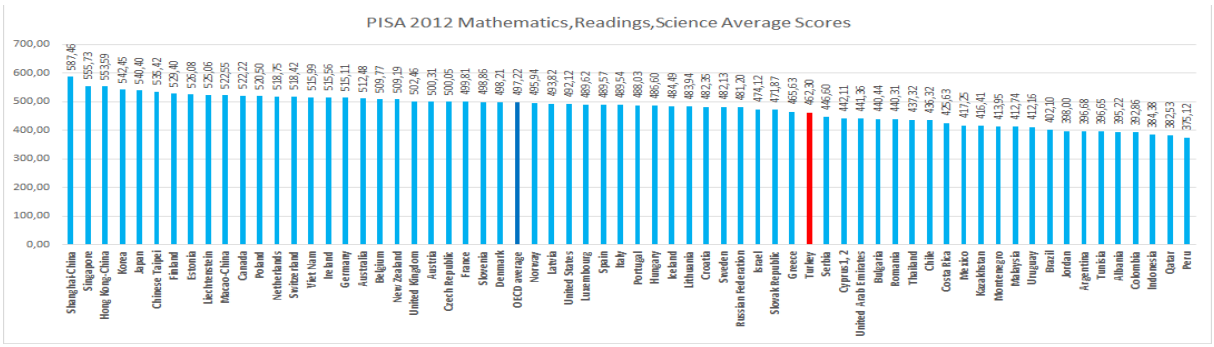
In industrial countries, where knowledge based industries are expanding rapidly, labor market demands are changing accordingly (World Bank, 2003: 1).World Bank's KAM methodology "education and skills" pillar aims to measue country's people needs education and skills that enables them to create and share, and to use it well. Education has a significant role to play that goes beyond input/ investment models to a finessed output set of specifications. For education pillar, three variables are used: the adult literacy rate (percentage of population aged 15 and above) gives a very broad stock measure of educated population, gross secondary enrollment rate and gross tertiary enrollment rates provide a flow rate. Experiences of developed countries shows that national innovation system, human development, effective education system and information-delivery technologies and business environments should affect each other satisfyingly and they should develop parallel to each other in order to have a successful knowledge economy (World Bank, 2004). Literacy rate, adult total (% of people ages 15 and above) in Turkey was 94.92 as of 2012 and 95,5 for 2015 as an estimation.

Over the past decade, the OECD Programme for International Student Assessment, PISA, has become the world's premier yardstick for evaluating the quality, equity and efficiency of school systems. PISA evaluates education systems worldwide by testing in key subjects critical thinking in math, science, and reading to 15 year olds(OECD,2012) It conducts research on the 65 countries that make up 90 percent of the world's economies. The OECD Directorate for Education has found that student achievement in math and science are a sound indicator for future economic health. Turkey has involved this test since 2003. As of 2015, they revealed 2012 PISA test results. According to PISA 2012 results, Turkey is ranked 42nd for maths, 41st for reading skills and 45th for science skills. Turkey had low PISA scores in 2003 and recorded an average four-point improvement in all three skills in 2012.

Figure 5-6-7-8-9: Student performance in maths, reading, science, problem solving means by countries







4. CONCLUSION

Knowledge economy relies primarily on the use of ideas rather than physical abilities and on the application of technology rather than the transformation of raw materials or the exploitation of cheap labor. It is an economy in which knowledge is created, acquired, transmitted and used more effectively by organizations, individuals, enterprises, and communities and that's why it promotes economic and social development. Knowledge economy discourse borrows heavily from work developed by a group of 1960s intellectuals, futurologists and information economists, like Peter Drucker (1969), Fritz Machlup (1962) and Daniel Bell (1973)–that industrial societies were in transition to becoming variously knowledge economies, post-capitalist and post-industrial societies. It is an economy in which knowledge is created, acquired, transmitted and used more efficiently and effectively by governments, enterprises, organizations, individuals and communities and it promotes economic and social development. In 1999 the World Bank Institute launched a project entitled "Knowledge for Development". It aims were to raise awareness among national policymakers about the powerful growth effects of knowledge and to encourage economists to combine global and local knowledge in order to accentuate comparative advantages (World Bank, 2008). According to findings, Turkey should increase investments in information and communication Technologies, also increase R&D expenditures, increase level of productivity in production by innovations, upgrade the level of education to have required necessary qualified human resources and should be directed to high-tech products in the production structure on the path of establishing knowledge economy.

REFERENCES

Arıkan, 2008. Evaluating the dynamics of innovation in Turkey: The impact of innovation on business performance, Boğaziçi University, Business Administration, PhD Thesis, İstanbul.

Işık, 2012. The economic analysis of R&D, innovation, patent and information technologies which influence the establishment of competitive advantage in knowledge economies: An application on Turkey's Economy, Atatürk University Social Sciences, Economics, Phd Thesis, Erzurum.

Leila, vd, 2007. "Adoption A Proper Tool For E-Readiness Assessment In Developıng Countrıes; Case Studies: Iran, Turkey And Malaysia", Journal of Knowledge Economy & Knowledge Management Volume **II** Spring, s.54.

Vuslat, 2014. "The transition of institutions by means of knowledge based Economy: case of Turkish universities", ITU, Department of Urban and Regional Planning Regional Planning Programme, M.Sc. Thesis, Istanbul

Eda, 2011. “Enformasyon Toplumu ve Bilgi Çağında Türkiye'nin Gelişim Süreci", Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İktisat Politikası Bilim Dalı, Master Thesis, İstanbul.

European Comission, 2001. The Impact of The E-Economy on European Enterprises: Economic Analysis and Policy Implications, Brussels.

European Comission, 2014. Innovation Union Scoreboard, http://ec.europa.eu/enterprise/policies/innovation/files/ius/ius-2014_en.pdf

Paul & Dominique, 2002. “Economic Fundamentals of the Knowledge Society”, Stanford University SIEPR Discussion Paper No.**01-14** .

Kaynak, 2011. "Avrupa Birliği Yolunda Bilim Ve Teknoloji Bağlamında Türkiye’nin Ab-27 Ülkeleri Karşısındaki Mevcut Durumu", Bilgi Ekonomisi ve Yönetimi Dergisi, Cilt: VI Sayı: **II**

Umut, 2001. Assessment of possible effective strategies in the transition process to a knowledge-based economy: The case of Turkey, Middle East Technical University Industrial Engineering Department, Master’s Thesis, Ankara.

George & Yi-Hsing & Zheng-Yi, 2008. “Competition policy for technological innovation in an era of knowledge-based economy.” Knowledge-Based Systems, Vol. 21, pp. 826-832.

Gür, 2001. Assessment of possible effective strategies in the transition process to a knowledge-based economy: The case of Turkey , Middle East Technical University Industrial Engineering Department, Master’s Thesis, Ankara.

Oğuz, 2005. “Bilgi Ekonomisi’nin Olası Mikro Ekonomik Etkilerinin Teorik Analizi”, http://oguzkara.com/Bilgi%20Ekonomisi'nin%20Olasi%20Mikro%20Ekonomik%20Etkilerinin%20Teorik%20Analizi.pdf

Charles, 2000. The weightless society: Living in the new economy bubble. NewYork: Texere.

European Comission, 2001. The Impact of The E-Economy on European Enterprises: Economic Analysis and Policy Implications, Brussels.

European Comission, 2014. Innovation Union Scoreboard, http://ec.europa.eu/enterprise/policies/innovation/files/ius/ius-2014_en.pdf

European Patent Office, 2015. http://www.epo.org/about-us/annual-reports-statistics/annual-report/2014/statistics/patent-applications.html

Oytun, 2013. "Türkiye İçin 2010-2012 Dönemi Karşılaştırmalı Bilgi Ekonomisi Analizi", Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,2013, 8(2), s. 115-139.

OECD, 1996. OECD, The Knowledge Based Economy, No. General Distribution OECD/GD(96)102, Paris.

OECD, 2015. http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf

Senem, 2011. “Analyzing of threshold effect in relationship with knowledge economy and growth”, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü / İktisat Anabilim Dalı, Ekonomi, Yüksek Lisans Tezi

Salduz, 2005. “Turkey's adaption to the transformation process of the European Union through competitive knowledge economy”, Dokuz Eylül Üniversitesi, Social Science Institute, Master Thesis, İzmir

Susan, 2005. Re-imagining and rescripting the future of education: global knowledge economy discourses and the challenge to education systems, Comparative Education, 41 (2), 151-170.

Tuncay, 2008. “Knowledge economy and economic growth: An econometric analysis on Turkey (1980-2006)”, Atatürk Üniversitesi / Sosyal Bilimler Enstitüsü / İktisat Anabilim Dalı, Ekonomi, Erzurum

Thomas A., 1997. Intellectual Capital: The new wealth of organization, Crown Business; 1st edition, San Francisco, CA

Kamil, 2007. “Teorik Çerçevesi ve Uygulama Örnekleriyle Dünyada ve Türkiye’deYazılım Endüstrisi”, Devlet Planlama Teşkilatı Müsteşarlığı, Planlama Uzmanlığı Tezi (yayımlanmamış eser), Ankara.

World Bank, 2003. Lifelong Learning for a Global Knowledge Economy, Washington, DC

World Bank, 2004. "Turkey Knowledge Economy Assessment Study", http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2007/04/06/000310607_20070406113501/Rendered/PDF/393650TR0Knowledge0economy01PUBLIC1.pdf .

World Bank, 2010. Knowledge Assessment Methodology , http://www.worldbank.org/kam .

World Bank, 2015. http://data.worldbank.org/indicator/SE.SEC.NENR/countries

YASED International Investors Association,2014, http://www.yased.org.tr/webportal/Turkish/Yayinlar/Pages/YASED2023HedefleriYolundaBIT.aspx

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

A WIFI BASED RESTAURANT ORDER AUTOMATION

DOI: 10.17261/Pressacademia.2017.598
PAP-WCTIE-V.5-2017(36)-p.262-269

Kerem Izzet Atam[1] Tansal Gucluoglu[2]
[1]Yildiz Technical University, keremizzet.atam@stu.khas.edu.tr
[2]Yildiz Technical University, tansal@yildiz.edu.tr

ABSTRACT

In traditional way of meal ordering process at a restaurant, firstly customers browse through the restaurant menu and then they wait for a waiter to come and take their orders. This procedure can be slow and decrease customer satisfaction during busy hours. Moreover, it is prone to human made mistakes. Automated systems can be used to optimize procedures by means of minimizing human made mistakes, decreasing waste of paper and thus making ordering procedures more time and cost efficient. There have been numerous attempts to automate ordering process in restaurants in recent years. For example, having a tablet PC on each restaurant table is one of the most popular solutions in the market while it is not cost effective since it requires many tablets. In this paper, we have designed and implemented an automation solution which benefits from WiFi's captive portal feature on mobile device of restaurant customers. In the proposed system, customers just need to connect restaurant's open WiFi from their table without waiting for a waiter to come and take their orders. Once their connection succeeds, customers will be directly forwarded to the web page on cloud server where they can browse through a menu and submit their orders without installing any application or any change in their mobile device. Each customer is identified by their WiFi card's MAC address which is reported to cloud server by lightweight software that runs on WiFi access point (AP) together with their received signal strength (RSSI) level and restaurant identification code. System is designed to scale easily to multiple restaurants placing configured WiFi AP on each restaurant using same cloud server. Our proposed system provides an easy to use solution for restaurant customers and a cost effective solution for restaurant owners as they can utilize their human resources more efficiently.

Keywords: Digital food ordering system, wireless food ordering system, restaurants, Wi-Fi, network

1. INTRODUCTION

Advancements in numerous fields of technology especially in wireless networking, cloud systems, and mobile devices have been deeply affecting our daily life routines and even changing centuries of traditions within less than quarter century. There have been developed numerous solutions in so many fields to optimize various procedures by substituting of humans tasks with automation systems. Hospitality industry is one of those areas where engineers studied to combine and utilize these technologies to develop automation mechanisms to have optimized and efficient systems in recent years.

Hence most of systems in hospitality industry are used by staffs but not customers even though customers are in the most critical part of their process. So staffs use those automation systems to work more efficiently, even though they interact with their customer in conventional way. For example think of ordering scenario described by steps in a restaurant:

- Waiters visit each one of tables to have orders from customers in traditional way.
- Then they use their hand terminal gadget to input orders to automation system.
- Automation system handles billing and also updates screens in kitchen.

This is one of most common scenario because staffs are obligated and also trained to use these automation systems but customers cannot be forced to be trained to use such new systems since they are the ones who must be satisfied, especially

in hospitality industry. In our proposed system, restaurant customer who will happen to use our system for the first time will be having ordering scenario as follows:

- Customer pick one of table and settles himself
- He starts to wait for the waiter to come to take his order.
- He notices paper holder with instructions simply telling him to connect WiFi point with mobile device to submit his order.
- Once he connects to WiFi AP, splash page with menu of restaurant welcomes with browser under URL of "http://menu.wf".
- He checks for the meals and adds desired ones into his basket.
- He submits his order with table code that is written on paper holder.
- Internet usage will be open to customer as he correctly enters 4 digits table code.

Customer is not required to install any application or make any change on his mobile device to use this system due to web based menu page, he also do not need to type domain name of website thanks to captive portal feature of WiFi AP which automatically directs customers to website. This system is advantageous compared to existing conventional procedures and other variants in the market by means of manpower need, likelihood of human made mistakes, capital, operational expenses and most importantly customer satisfaction. Since all it is really is for customer to submit his order via this automation and all it needs for restaurant is a WiFi AP to install system. The rest of this paper is organized as follows. Section 2 reviews the existing and new approaches that are followed in recent years. Section 3 describes system design. Section 4 shows and discusses experimental results of our implementation and finally in section V we draw conclusion and future scope.

2. LITERATURE REVIEW

2.1. Waiter Based Systems

Waiter based approaches are most widely used methods where waiters have full authority to manage communication between customer and kitchen staff. It is favorable for places where social interaction between customer and restaurant staff is important. But it has so many drawbacks since it is bulky and open to human made mistakes.

2.1.1. Pen and Paper Model

This method and its variants are most primitive and most common procedure used in restaurants. In this system waiter uses small notepaper to write down orders. Menu cards are printed into paper. Ordering operation starts as customer sits to their table and waits for the waiter to come. Then waiter delivers menu card and waits for some time for customer to decide what to order. After waiter notes the orders he goes to kitchen part to inform kitchen staff for meals.

This traditional method has so many drawbacks in terms of efficiency and cost. System is highly depended on paper which is not suitable material for environments where foods are prepared and served. Paper can easily be affected by water, fire and oil. It is open to human made mistakes because waiters listen orders, they note them and they deliver it to kitchen. There are three people in this communication chain which makes it really likely to happen mistakes. It is common for restaurants to update price lists or making various modifications on menu cards. But even for minor edits whole menu cards need to be updated which results in paper waste and money. Waiters should check every table for customers if they are ready to order or waiting for check out. It can be even more troubling for places where there are more than few floors and architecture doesn't let staff to see every point in restaurant. It is also not easy to manage paper notes for further investigation to collect and evaluate statistics about restaurant.

2.1.2. Kiosk Screen

These systems are early examples hence most common systems in an attempt of digitalizing ordering process and also maintain communication between waiters and kitchen staffs. There are mostly one or few screen terminals in central positions in restaurant depends of size of place. As waiters collect orders with their notepaper, they come and input those data into screen terminals. Order information will reflected to billing and screens in the kitchen. Each screen terminals are connected to a PC running special software.

This system makes it possible to evaluate detailed statistics about orders and understand customer tendencies for different time periods. It decreases service time as waiter doesn't need to go to kitchen to inform kitchen staff for each order. It is also less likely to have human made mistakes because it is easier to follow orders reflected into screens compared to reading from paper or keeping in mind the orders. Even though it is still bulky since waiters still need to note orders into papers and system still prone to mistakes. This system can't bring any solutions for menu cards to be easily modifiable as they are printed into papers.

2.1.3. Wireless Mobile Terminals

They are basically mobile versions of kiosk systems, namely personal digital assistant (PDA). They are useful as waiters do not need to note every order to paper or memorize them to re-input these information into kiosk. Waiter with PDAs can easily visit tables with it by carrying this device like a notebook. They can directly input orders and details into PDA once customers verbally inform customers.

Even though it is one of most improved technology among common automation systems for waiters in market it has some drawbacks too. Especially large restaurants with high amount of waiters require more of those devices which will lead to more expenditure. Each one of those devices continuously should function properly with enough battery level. So need for maintenance for all devices will cost time and manpower to restaurant as many pieces of hardware used in system.

2.2. Customer Based Systems

Customer based systems are advance version to the existing widely used automation technologies in market. In these systems customers take action first and perform operations via touch screens to submit their orders. They optimize communication chain between customer, waiter and kitchen staff. Since these systems are used by the customers they must provide easy to use and understandable interfaces because it is not practical to train each one of customer to use and teach such technologies.

2.2.1. Mobile Application Approach

Mobile application approach is low cost solution by means of necessary hardware as it does not require any additional device for restaurant. This method can be applied for customers to order before they come to restaurant to save time [1] or for customers inside restaurant [2]. In these systems mobile device of customers is used to submit their order. These systems expect customers to have mobile device and also expect them to install special application for restaurant. Sadly this scenario less likely to happen as long as customer is not obligated to use to order from mobile device. Because it is already a big challenge to expect customers to submit their order by themselves but in this method they also need to find this application in appstore, probably download it from internet that charged on their ISP subscription and submit order with table number which is mostly encoded into QR code. This method requires more developer work as customers might have various kinds of device beside commonly used Android and iOS based devices. There might be customers having Windows phone or even Laptop with Linux based operating system. Most of variant of such systems limit itself with Android and iOS because even struggling between versions of Operating Systems is difficult enough to maintain.

2.2.2. Touch Screen Assembled on Table

In this system all tables in restaurants have touch screens assembled or special casing that connects screen to table. Most of implementations use tablet computers with special application always running on it [3][4][5]. This method is favorable compared to mobile application approach as customer can easily spot those screens on table and start to use it. But installing these systems in restaurant will cost so much money. Additionally increased number of hardware will require more maintenance manpower considering all screens must work properly under condition of public use. Moreover screens under public use on food table will likely to have fingerprints of previous customer. It will either lead to customers with hygiene concerns abstain from using it or will require for waiters to clean those screen for ever service. It is likely to observe restaurants with these systems have low cost resistive touch screens which result in low quality of user experience. Because such screens have no multi touch support and screen reaction is slow.

3. PROPOSED SYSTEM

As it is described above waiter based ordering systems have more of workload into one person in communication chain from customer to kitchen staffs. New generation systems are trying to solve this with giving authority to customer on their order with various approach which are either not user friendly or expensive to install. In our study we designed and implemented a solution considering both user experience and cost of installing in a restaurant along with operating it with no additional effort or expenses like providing maintenance to so many pieces of hardware. Objective of our proposed

system is to use WiFi, web and mobile technologies to address three main problem in restaurants that new generation solutions that could not able to solve :

- Decrease work load on waiters in process of ordering by optimizing communication chain between customer, waiter and kitchen staff with more involvement of customers.
- Provide restaurants dynamic menu card system where they can modify whenever they want and even put daily advantages instead of using printed media.
- Customers no longer need to ask for WiFi password to restaurant staff as they will able to use internet once they submit their order with correct 4 digits of table code that provided to customer with Plexiglas stands on each table.

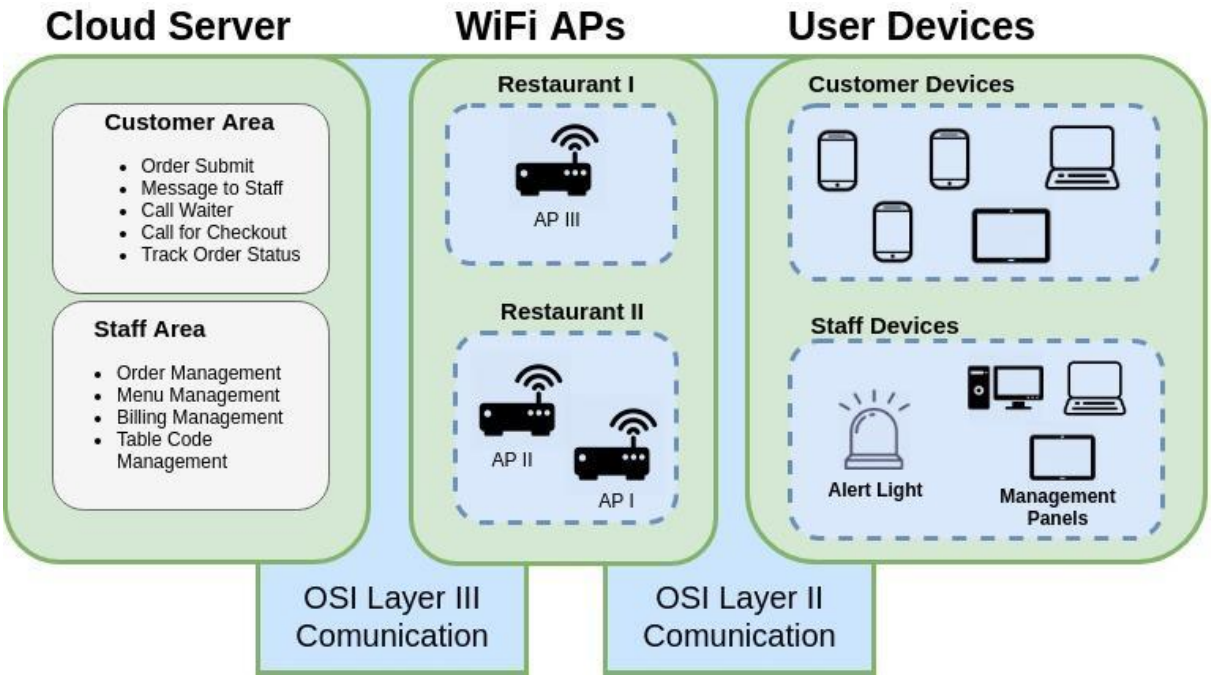
Our system is beneficial to mobile application based approach as there no need for customers to install any additional application or to do any change on their mobile device. Furthermore this system even motivates customer to use it by providing direct access to open WiFi. It is also easier to maintain our system for wide range of devices as it is developed for browsers as web site instead of specific operating system.

This solution is more cost effective solution compared to system with touch screens assembled on each tables. All it needs for our system to install in restaurant is to place WiFi AP with internet connection and provide tablet to waiter or kitchen staff to receive orders. There is no need to maintain a server computer inside restaurant as this system is cloud oriented. This will decrease maintenance work and make it easier to manage more than one restaurant from one center.

3.1. System Design

System has three main parts as it is depicted in Figure I; which are user devices, WiFi APs and cloud server.

Figure I: System Design



3.1.1. Cloud System

This is Linux based Ubuntu 14.04 PC running Apache HTTP server. This machine serves both customers and staffs for multiple restaurants. Following technologies are used in our cloud server to manage and process orders, customers, menu list and table codes for all restaurants. MongoDb is used as database software which is classified as open source document

oriented NoSQL database program. We have prefered it as NoSQL database systems like MongoDb fits agile approach very well. Because it gives space to software to be more elastic for changes according to different business models that are used in different restaurants. PHP is used in backend of web development to process and manage all kinds of data in automation and store them into database. AngularJS and JQuery are used in frontend of web development. AngularJS was our prior choice because we consider it easier way to manipulate HTML pages compared to other Javascript frameworks. This server contains two different UIs one for restaurant customers and other one for restaurant staffs which are separated as customer and staff area respectively.

Customer Area - When customers are connected to WiFi point they will automatically be forwarded to this area of server. Then server will detect customer's restaurant by the report of AP that customer device connected. Customer UI of related restaurant will welcome afterward. Customers can browse through menu page, submit their order, track their order status, leave message to restaurant staff, call waiter and make call for checkout within this area.

Staff Area - This section will serve to restaurant staffs including managers, waiters, kitchen staff and whoever in charge of menu lists and ordering process. Only devices with pre-configured MAC addresses have access to this part of server. There is also login page to ensure security of management panel against MAC spoofing attacks. Restaurant staff can access management UIs and control menu lists to be shown to customer area and manage ordering process.

3.1.2. WiFi APs

Those are TP-Link devices providing wireless connection within IEEE 802.11 b/g/n protocols which operates in 2.4GHz carrier frequency band. Restaurant can have more than one WiFi AP to maintain full signal coverage to customers within dining area. WiFi APs carry following tasks with our software running on them:

- Redirect customers to menu page of restaurant under domain name "menu.wf" once they connect to WiFi point with captive portal feature. It will let system to catch customers really quickly without need of installing any application or doing any change on mobile device.
- As our system is designed to give support multiple restaurant. WiFi AP sends Restaurant ID to server so that our software that runs on cloud will know which restaurant's customer is accessing. Customers will see the menu page of related restaurant that they are in.
- MAC address of WiFi card of customer's mobile device is reported to cloud server to identify each customer. MAC address is unique identifier that every WiFi card that have. Our system can identify each customer with this address without need of customer enrolling or signing in to our system.
- RSSI level of each customer device reported to server with the order submission by WiFi AP. RSSI is a digital value of measurement of the power present in a received radio signal. This information can be seen by management panel for security reason and also monitoring coverage performance of WiFi AP. This system can be used anywhere WiFi AP coverage reaches. So that some people with abusive intents might submit orders outside of restaurant. These suspicious orders can be detected by the RSSI performance as they will have weaker signal powers.

3.1.3. User Devices

These can be various devices from mobile phone to laptop PC that able to connect WiFi APs to access either customer area or staff area on cloud server. These devices should have WiFi card supporting IEEE 802.11 b/g/n protocols which operates in 2.4GHz carrier frequency band and able to run any kind of web browser. Today all kinds of smart device even old fashioned ones meet with such requirements.

Customer Devices - These are mobile devices that are used to access customer area on cloud server. Devices are directed to menu splash page in customer area by their connection to WiFi AP. They can view meals, submit their orders, track their order, send call to waiter with message and ask for bill to checkout within customer UI.

Staff Devices - These are equipments that are used by restaurant staff to manage order process and modify menu lists. MAC address of network interfaces that are used for this purpose are defined into cloud server for security control with the help of MAC address reporting mechanism that WiFi APs have.

- Management Panels: These devices are used to access staff area of cloud server to manage ordering process and modify meal lists. These can be small tablet to desktop PC having connection to cloud server via WiFi AP.

- Alert Light: We have developed a light system in our automation in a request of restaurant staffs. As our system is alternative to their existing automation, continuously checking our system's order management panel requires extra work for them. We have designed order alert module that notifies staffs with light for the any customer call that needs action or approval by restaurant staff. We have used an open source IoT platform named NodeMCU having built in ESP8266 WiFi chip with full stack TCP/IP support along with MCU.

4. RESULTS AND DISCUSSION

In our development we have aimed to apply agile approach which leads us to work with our customers during continuing development process to shape the system in optimum way addressing practical solutions on place. Sarayli restaurant welcomed us to install our system which is located in close vicinity of our development environment. So we could easily observe and test our system. This is big restaurant with capacity to serve 3000 people in 4000 m2 area. They like to test this system in garden part where their waiters having difficulty to manage that area with waiters. We have chosen 8 tables for our test setup and placed them our paper stands which contain table codes and instructions to use system. One tablet PC is placed just next to kiosk screen to receive orders as management panel.

During our tests we have surveyed with restaurant staffs and managers and ask for any kind of problem to solve and any additional feature to develop. Their first demand was full adaptation of our system to their existing automation software which was waiter based kiosk system described above in literature review section. Sadly it is not possible for us to maintain our software to work with any type of automation as we have developed this system as a generic design which should able to be applied to any restaurant. It is not scalable to develop systems that need efforts to adapt each one of restaurant having different existing automation software. Their another demand came from waiters. They have complained that they need to check management panel continuously all the time to not miss any received orders. To solve this issue we come up with wireless alert light which warns waiters for incoming orders or demands of customer. So they no longer felt to need for check this screen all the time. One more improvement request came for billing summary. Existing billing system was showing receipt per customers. But sometime it might also be asked per table total including every customers in table. We have developed detailed view in billing management section where they can view order totals per table and per customers.

We have observed our system performance by number of orders and statistics of menu page traffic during one week period. Statistics for users are obtained both within our system and by Yandex Metrica, a free service designed to measure websites visits and analysis of user behavior. In our tests; restaurant menu page viewed by 112 different devices with the bounce rate of 26.6% which means 73.4% of customers browsed through menu page and checked meals during their attempt to connect to internet via our WiFi AP. Our system received 17 of 83 orders in 8 tables during one week period which makes 20% of orders. Considering this restaurant have mainly customers from middle aged people who visit this place with their families, we name this outcome as acceptable result as it has even higher potential for places with younger customers who are more motivated to use new technology.

Figure II: Devices by Number of Sessions

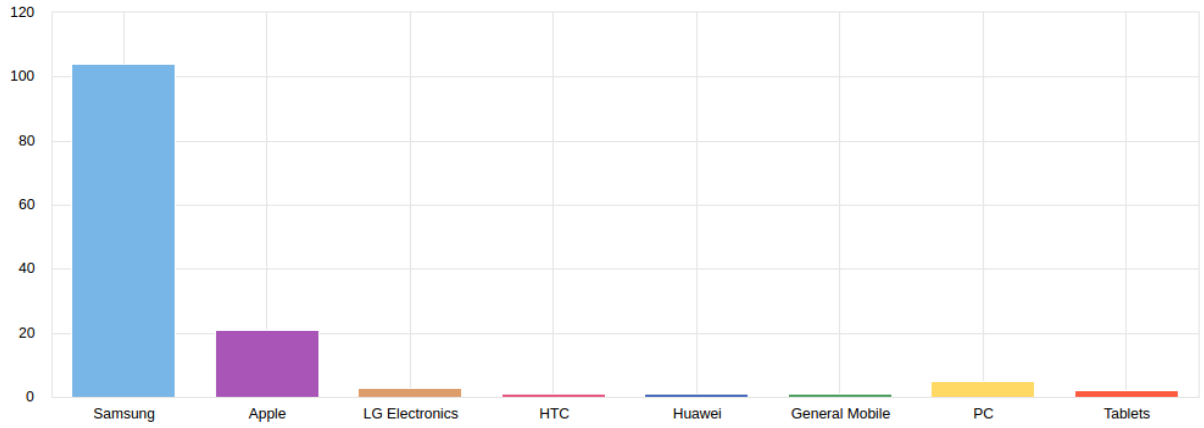


Figure III: Age Groups

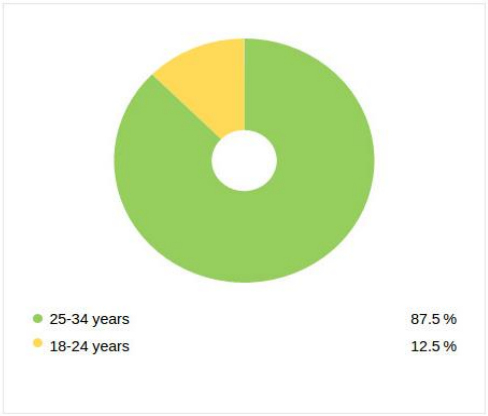


Figure IV: Device Shares

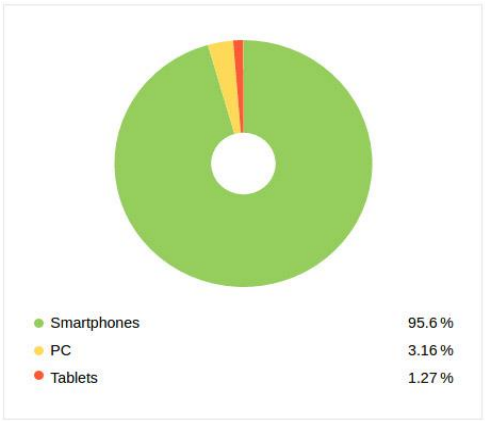


Figure III and figure IV shows share of age groups and device types respectively. Figure II depicts detail statistics of customer devices by number of sessions. Yandex Metrica reports support our claims that people who are in over aged almost never tried to use system even though majority of customers are from this age group according to our observations. Statistics also shows superiority of our system over mobile application approach explained in literature section where it is hardly possible to give support to large spectrum of devices from various vendors.

5. CONCLUSION

In this study, we have proposed our design and implementation of restaurant automation with different approach unlike systems in use in the market and proposed designs in literature. We have discussed advantages of our system as it attracts customers with WiFi connection and let them view menu list, submit their order, track order process, call waiter and message to restaurant staff with provided UI option on their mobile device. We aimed to have less man power need of restaurant staffs as we give more authority to customers into their ordering process.

As our project will be used by more and more in restaurants, our design is open to development for other features and ideas. This system has unique network model as this system exists in both ends of WAN communication from AP to cloud server. It brings different kinds of security threats and also makes it possible for different approaches to secure. We have used MAC based security protection but it can also be exploited by MAC spoofing attacks. To prevent this kind of abusive penetration attempts into our system, smarter methods can be used to verify staff devices using machine learning algorithms by taking usage behaviors into account to detect suspicious login attempts.

REFERENCES

[1] Varsha Chavan, Priya Jadhav, Snehal Korade and Priyanka Teli, "Implementing Customizable Online Food Ordering System Using Web Based Application", in IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 4, April 2015.

[2] Shweta Shashikant Tanpure, Priyanka R. Shidankar, Madhura M. Joshi, "Automated Food Ordering System with Real-Time Customer Feedback", in International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 2, February 2013

[3] Nibras Othman Abdul Wahid, "Improve the Performance of the Work of the Restaurant Using PC Touch Screen", in Comput Sci Syst Biol 7: 103-107. doi:10.4172/jcsb.1000144

[4] Kunal Gupta, Stuti Saxena, "DESIGN AND IMPLEMENTATION OF WIRELESS MENU CARD", in Confluence The Next Generation Information Technology Summit (Confluence), 2014 5th International Conference

[5] Bhaskar Kumar Mishra, Tanmay Bakshi, Bhawani Singh Choudhary, "Touch Based Digital Ordering System on Android using GSM and Bluetooth for Restaurants", in India Conference (INDICON), 2015 Annual IEEE

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

SAFETY SHOES WEARER'S COMFORT PERCEPTION AND EFFECTS AMONG MANUFACTURING EMPLOYEES

DOI: 10.17261/Pressacademia.2017.599
PAP-WCTIE-V.5-2017(37)-p.270-273

Baba Md Deros[1], Mohd Iezalman Ab Rahman[1], Nurul Huda Baba[1] & Ahmed Rithauddeen Yusof[1]
[1]Universiti Kebangsaan Malaysia

ABSTRACT

This study investigates ergonomic problems faced by manufacturing employees due to wearing inappropriate safety shoes. A total of 30 survey respondents was recruited based on the shoes they wear. Pedar-X was used in the experiment to measure the pressure that acts on the wearer' feet. Survey results showed the wearer' experienced the highest pain with the frequency of 80% for right and 83.33% for the left heel for wearing safety shoes two to three times a week. Meanwhile, Pedar-X recorded average peak pressure, force and surface area: Amporlo, 90.53 kPa, 259.57 N, and 67.74 cm2; Land Master, 75.07 kPa, 273.79 N and 72.65 cm2; Jack Parcell ordinary flat, 67.85 kPa, 170.68 N and 52.08 cm2 respectively. It was found strain on the front and middle of the feet were lowest. Meanwhile, the back foot, which locates the heel recorded the highest pressure value of 185 kPa as a result of wearing inappropriate safety shoes.

Keywords: Peak pressure, force, surface area, foot, comfort

1. INTRODUCTION

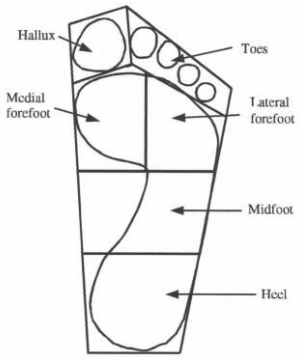
There is a wide range of safety shoes designed and used in the workplace depending on the existing hazard. It can be divided into three categories, such as safety, protective and occupational shoes. Safety shoes are made of various materials and are equipped with a special internal protection element, such as toecap, metatarsal anti-stress, ankle protector and anti-penetration are installed in the safety shoes (Irzmańska, 2014; ISO 20344, 2012) for foot protection from impact and pressure. These protective elements are made from different materials, which includes steel, aluminum, plastic and composite (Koradecka and Konarska, 2002).

2. LITERATURE REVIEW

The foot is primarily supported by the plantar surface. Figure 1 shows the definition of the six plantar surface regions. For example, when a person is standing (i.e. the inside of a shoe), there will be interaction between the plantar surface of the foot and the surface on which the person stands. The interaction between the plantar surface, dorsal surface of the foot and the load-bearing surface, the whole body posture contributes to foot and surface deformations and hence to perceived comfort, discomfort, pain or fatigue (Witana et al. 2009; Cham and Redfern, 2001; Kelaher et al. 2000; King, 2002). High plantar pressures have been linked to foot pain and discomfort (Gardner et al. 1988; Hodge et al. 1999). Other factors that could contribute towards comfort are firmness, anthropometric features and pressure distributions (Na et al. 2005).

In manufacturing industries, safety shoes are specially designed for employees to wear to protect their foot from workplace hazards such as rolling, falling or sharp objects. The basic functions of safety shoes are to provide protection the foot against hazards present at the workplace (Irzmańska, 2014; Koradecka (2010). There are various types of safety shoes have been specially designed for use in factories and in other industries. Safety shoes have a wide variety of designs and each type of safety shoes manufactured has its own advantages and disadvantages. In addition, the design of appropriate safety shoes for various industrial sectors, professions, and work positions need to meet all requirements with respect to ergonomic and protective properties, durability, hygiene, comfort for use and functionality to avoid the occurrence of foot pain and injury in the workplace (Dobson et al. 2017; Koradecka, 2012).

Figure 1: Definition of the Six Plantar Surface Regions (Hong et al. 2005).



On average, the majority of employees working in the manufacturing industries spent a total of eight hours per day or more. While working, for example, an operator in semiconductors or in the automotive manufacturing factory, employees usually work in prolonged sitting and standing postures. Wearing inappropriate safety shoes will negatively affect the comfort level of an employee. Thus, the two main objectives of this study are to review and evaluate the overall effects of wearing of safety shoes to employees.

3. DATA AND METHODOLOGY

3.1 Respondents

The study population is employees from manufacturing industry itself. They were selected due to their positions, such as the factory managers, executives, engineers, safety officers and supervisors from the plant.

3.2 Data Collection

The survey questionnaire comprised of three parts; socioeconomic status, risk factors, and prevalence of pain in the lower part of the body. A survey questionnaire was given to selected respondents in the manufacturing industries. Every respondent was given the explanation and guidance on how to complete the survey questionnaire.

3.3 Pedar-X System

Pedar-X system data collection method was used to determine the shoe pressure. This method used an insole or a liner attached with a sensor that can measure the amount of pressure in a shoe. In this study, data was collected while the subject is standing in the upright position only (Ramanathan et al. 2010). During the experiment, subjects were asked to wear three different types of shoes. The main purpose of this experiment was to measure and determine data distribution experienced by the wearer of the shoes.

Parameters considered in this method are contacted area, contact time, maximum force, pressure-time integral, force-time integral, peak pressure, mean force and mean area. However, to simplify the data collection in this study, only a few parameters were taken, which include peak pressure, force and mean area. Having done that, comparisons were made for all the three types of shoes selected. Figure 2 shows the experimental set-up preparation for Pedar-X System.

Figure 2: Preparation for Pedar-X System Device

4. FINDINGS AND DISCUSSIONS

4.1 Survey Questionnaire Results

Based on Cornell analysis found in the survey questionnaire, the impact of safety shoe application showed pain commonly occurs at the left and right heel with a total average prevalence of 81.67% and pain occurs as a result of their standing work posture. Part of the calf showed the relatively high average prevalence of pain at 51.67%. Meanwhile, the hips showed the lower average prevalence of pain at 26.67% compared to other body parts. Other parts of the body, such as the neck, shoulders, back, arms, hips, wrists, buttocks, hips, and knees are not affected with the wearing of safety shoes.

The survey result shows that when the work is done while standing, the pressure of the shoes will be higher than the pressure on the shoes while sitting. This is the main cause why parts of the heel have a high percentage of the pain after wearing safety shoes. Employees at this automotive car body assembler work at different workstations. Each worker has to install different car components, where he or she has to stand upright while installing each component in the vehicle.

4.2 Pedar-X System Analysis

Data analysis using the Pedar-X system was conducted on three different types of shoes, two of them were safety shoes (i.e. Amporlo and Land Master) and the third one is ordinary flat shoes (i.e. Jack Parcell). Normal data distribution from the Pedar-X system for Amporlo safety shoes shows that the peak pressure was 82.5 kPa, the force of 243.77 N and surface area 67.33 cm^2. For Land Master safety shoes, the peak pressure was 67.5 kPa, the force of 224.363 N and surface area 72.17 cm^2. Jack Parcell the flat shoes indicate the peak pressure of 60kPa, force 163.193 N and surface area 48.67 cm^2.

Experimental results as shown in Figure 3, when the subject was wearing Amporlo safety shoes, the pressure distribution data recorded at foot front was 88kPa, at the middle part of the foot was 15kPa, and at the back of the foot (i.e. heel) was 185kPa. Meanwhile, for Land Master safety shoes as shown in Figure 4, the pressure distribution data recorded at the foot front was 88kPa, at the middle part of the foot was 15kPa, and at the back foot (i.e. heel) was 160 kPa. For Jack Parcell normal flat shoes as shown in Figure 5 the pressure distribution data readings recorded at the front of the foot was 50 kPa; at the middle part of the foot was 15kPa, and at the back foot (i.e. heel) was 83 kPa.

Figure 3: Pressure Distribution Recorded on the Subject's Foot When Wearing Amporlo Safety Shoes

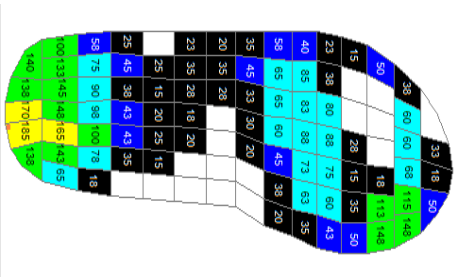


Figure 4: Pressure Distribution Recorded on the Subject's Foot When Wearing Land Master Safey Shoes

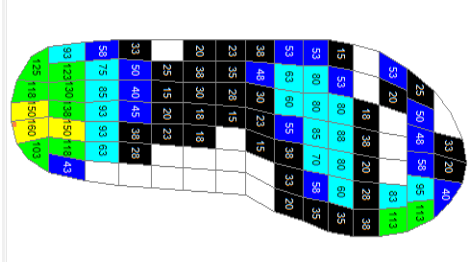
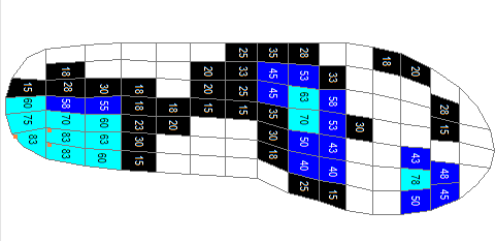


Figure 5: Pressure Distribution Recorded on the Subject's Foot When Wearing Ordinary Flat Shoes



5. CONCLUSION

Based on the data found in the survey conducted, the impact of wearing safety shoes resulted in pain at left and right heel with a total average prevalence of 81.67% due to their standing work posture. The heel recorded the highest pressure distribution data when wearing all the three shoe types: Amporlo safety shoes, 185kPa; Land Master safety shoes, 160 kPa; and Jack Parcell normal flat shoes, 83 kPa. Meanwhile, the pressure distribution data recorded at the foot front when wearing: Amporlo safety shoes, 88kPa; Land Master safety shoes, 88 kPa; and Jack Parcell normal flat shoes, 50 kPa. The pressure distribution recorded at middle of the foot when wearing both safety shoes and ordinary flat shoes were the lowest and all three different shoes recorded the same value, which is at 15kPa. When compared with data from the Pedar-x system, proved that the strain on the front and middle of the feet are the lowest and when compared with data from the Pedar-x system, also proved that the pressure on the back of the legs which is the heel, highest with the value of 185kPa.

REFERENCES

Cham, R., Redfern, M.S., 2001, "Effect of flooring on standing comfort and fatigue, Human Factors", vol. 43, no. 3, pp. 381–391.

DeVocht, J.W., Wilder, D.G., Bandstra, E.R., Spratt, K.F., 2006, "Biomechanical evaluation of four different mattresses", Appl. Ergon., vol. 37, no. 3, pp. 297–304.

Dobson, J.A., Riddiford-Harland, D. L., Bell, A.F. & Steele, J.R. 2017, "Effect of work boot type on work footwear habits, lower limb pain and perceptions of work boot fit and comfort in underground coal miners", Applied Ergonomics, vol. 60, pp. 146-153.

Gardner, L.I., Dziados, J.E., Jones, B.H., 1988, "Prevention of lower extremity stress fractures: a controlled trial of a shock absorbent insole", Am. J. Public Health, vol. 78, pp. 1563–1567.

Hodge, M.C., Bach, T.M., Carter, G.M., 1999. "Orthotic management of plantar pressure and pain in rheumatoid arthritis, Clin. Biomech, vol. 14, pp. 567–575.

Hong, W. H., Lee, Y. H., Chen, H. C., Pei, Y. C., & Wu, C. Y. 2005, "Influence of heel height and shoe insert on comfort perception and biomechanical performance of young female adults during walking. Foot & ankle international, vol. 26, no. 12, pp. 1042-1048.

Irzmańska, E. 2014, "Footware use at workplace and recomendations for the improvement of its functionality and hygiene", AUTEX Research Journal, vol. 14, no. 2, pp. 89-94.

ISO 20344, 2012. Personal Protective Equipment. Test Methods for Footwear.

Kelaher, D., Mirka, G.A., Dudziak, K.Q., 2000. Effects of semi-rigid arch-support orthotics: an investigation with potential ergonomic implications. Appl. Ergon. 31 (5), 515–522.

King, P.M., 2002, "A comparison of the effects of floor mats and shoe in-soles on standing fatigue", Appl. Ergon., vol. 33, no. 5, pp. 477–484.

Koradecka D. 2012, "Use of Personal Protective Equipment in the Workplace. In: Handbook of Human Factors and Ergonomics", USA: John Wiley &Sons Press 2012, pp. 895-910.

Koradecka, D., 2010. Handbook of Occupational Safety and Health. CRP Press Taylor & Francis Group, USA.

Koradecka, D., Konarska, M., 2002, "Advance in physiology and ergonomics and their impact on th hygiene of work", Medycyna Pracy (Occup. Med. – in Polish), vol. 53, no. 1, pp. 15-21.

Na, S., Lim, S., Choi, H.S., Chung, M.K., 2005, "Evaluation of driver's discomfort and postural change using dynamic body pressure distribution", Int. J. Ind. Ergon, vol. 35, no. 12, pp. 1085–1096.

Ramanathan, A. K., Kiran, P., Arnold, G. P., Wang, W., & Abboud, R. J., 2010, "Repeatability of the Pedar-X® in-shoe pressure measuring system, Foot and Ankle Surgery, vol. 16, no. 2, pp. 70-73.

Witana, C. P., Goonetilleke, R. S., Xiong, S., & Au, E. Y., 2009, "Effects of surface characteristics on the plantar shape of feet and subjects' perceived sensations", Applied Ergonomics, vo. 40, no. 2, pp. 267-279.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

OVERVIEW AND COMPARISON OF THREE CLASSIFIERS: ARABIC DOCUMENTS AS A CASE STUDY

DOI: 10.17261/Pressacademia.2017.600
PAP-WCTIE-V.5-2017(38)-p.274-277

Essam Hanandeh [1]
[1]Zarqa University, hanandeh@zu.edu.jo

ABSTRACT

Nowadays, text classification is used in various fields of research and applications, such as information retrieval, text mining, and data mining. This study tests the Naïve Bayes, K-Nearest Neighbors, and Support Vector Machine algorithms on a relatively large dataset of Arabic documents. This dataset comprise 1,000 Arabic documents that are distributed across 10 classes. This comparison is based on recall and precision measures. The evaluation results show that the Support Vector Machine algorithms classifier outperforms the other two.
Keywords: Arabic text categorization, KNN, NB, SVM, text mining

1. INTRODUCTION

Classification constitutes a significant part of data mining, text mining, and machine Learning. In classification, a machine or human attempts to identify to which class from a set of classes a new instance belongs (Agirre et al., 2009). Machines are configured to classify different instances by referring to instances whose classes are known. Therefore, classification algorithms adopt supervised learning, whereas clustering algorithms adopt unsupervised procedures. Usually, classification algorithms use different features to determine the class of each instance under consideration. In our study, each instance represents a text document (Syiam et al., 2006). Therefore, classification of documents requires a highly dimensional feature space with scarce data. As dimensionality increases, the space of the scarce data increases. The increase in dimensionality and scarcity makes the classification problem harder to solve.

This study explores the effectiveness of three popular classification algorithms to classify Arabic text documents. The three classifiers under study are Support Vector Machine (SVM), Naïve Bayes (NB), and *K*-Nearest Neighbors (KNN).

This study also identifies the effects of reducing the advantage of these classification algorithms. Data mining and text mining are very important because they can handle the rapid growth of data that are collected and stored into large and numerous databases. These databases exceed human ability for comprehension, classification, and organization without the aid of powerful tools. Data mining is necessary for turning data that is stored in these databases into useful information which may help in decision making (Karima et al., 2005). One of the significant applications of data mining is text classification. Text classification aims to automatically assign

2. RELATED WORKS

This section presents some studies that are relevant to the present study. Therefore, only the studies that tested the effectiveness of different classifiers on a collection of Arabic documents are presented.

(Al-Kabi et al., 2007) investigated the effectiveness of six classification methods (i.e., inner product, cosine, Jaccard, Dice, NB, and Euclidean). They computed inner product, cosine, Jaccard, and Dice as associative coefficients of the vector space model (*VSM*). Their findings show that cosine is the most effective method. Furthermore, they concluded that NB is better than the other five methods tested in their study.

(Gharib et al., 2009) used SVM to classify 1,132 Arabic documents. They compared the results that they obtained from SVM with those from NB, KNN, and Rocchio. Their comparisons show that Rocchio is the best classifier for small feature sets, whereas *SVM* is the best classifier for large feature sets.

(El-halees, 2011) adopted a combined approach to extract opinions from Arabic documents. To enhance the performance of algorithms in classifying Arabic documents, he adopted a combined approach that consists of three methods. The lexicon-based method was used first. The resultant categorized documents were used as a training set for maximum entropy method, which subsequently classified other documents. Lastly, KNN was used to classify documents, which underwent both the lexicon-based method and the maximum entropy method. The results of his experiment show significant improvements in the performance of KNN.

(Wahbeh et al., 2012) conducted a comparative study on four free data mining tools (i.e., WEKA, Orange, KNIME, and Tanagra) for text classification. They concluded that WEKA is the best method.

(Khorsheed et al., 2013) conducted a study that involves a survey of Arabic text classification and a comparison of the effectiveness of different methods for Arabic text classification. They concluded that SVM is the best method, followed by the decision tree algorithm (C4.5), and then NB.

(Hanandeh and Mamoun., 2014) Conducted a study that aims at investigating different variations of vector space models (VSMs) using KNN algorithm. The Experimental results against the Saudi data sets reveal that Cosine outperformed over of the Dice and Jaccard coefficients.

3. EXPERIMENT RESULTS

The analysis of the results presented in the literature show that:

1. No standard Arabic corpuses are available that can be used easily. Almost all of the authors justified this finding by the lack of Arabic corpus in general (Duwairi, 2011) (Karima r al., 2005).

2. The technique of removing stop words, digits, numbers, punctuation marks, and non-Arabic words was used to prepare the text for classification. Some of the authors extracted root words (Duwairi, 2011)(Gharib et al., 2009) (Mesleh, 2008). Whereas the others preferred not to do so (Ababneh et al., 2014) (Alsaleem, 2011) (Al-Harbi et al., 2008) (Bawaneh et al., 2008). Because of the problem on the conflation of numerous terms to the same root word (Khreisat, 2009)

3. Almost all of the authors used recall, precision, and F1 (Ababneh et al., 2014) (Gharib et al., 2009) (Bawaneh et al., 2008).

4. Differences were observed among classifiers in terms of accuracy, error rate, and time taken to build the classification (Wahbeh et al., 2012).

5. SVM outperformed KNN and NB (Agirre et al., 2009) . To assess the accuracy of the proposed classifiers, Arabic text corpus was collected from online magazines and newspapers. A total of 1,000 document with varying lengths and writing styles were collected. These documents fall into 10 pre-defined categories. Every category contains 100 documents.

The set of pre-defined categories include sports, economy, Internet, art, animals, technology, plants, religion, politics, and medicine. Two authors manually categorized the collected documents. Every document was assigned to only one category. Whenever a document was found to belong to more than one category, it was assigned to the category with the maximum likelihood according to human categorizer's judgment. The accuracy of a classifier is expressed in terms of recall and precision. Figure 1 shows the recall and precision values for different values of *k*.

Figure 1: Recall & Precision for k10, k20, k50, ad k100

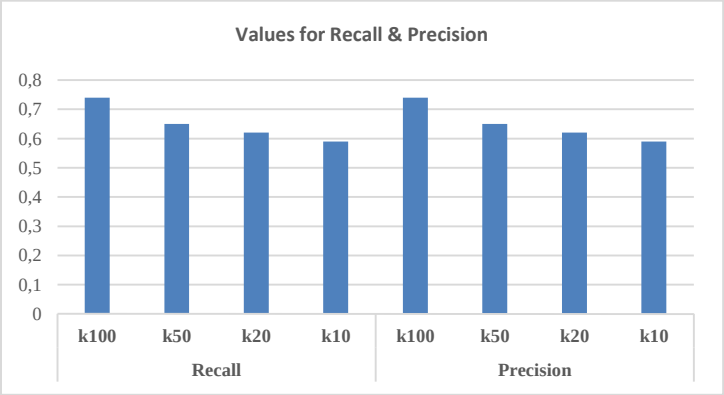


Figure 2 shows the recall and precision for all categories by all classifiers. The figure shows that the recall and precision values of SVM are better compared with NB and KNN.

Figure 2: Recall & Precision for SVM, NB, and KNN

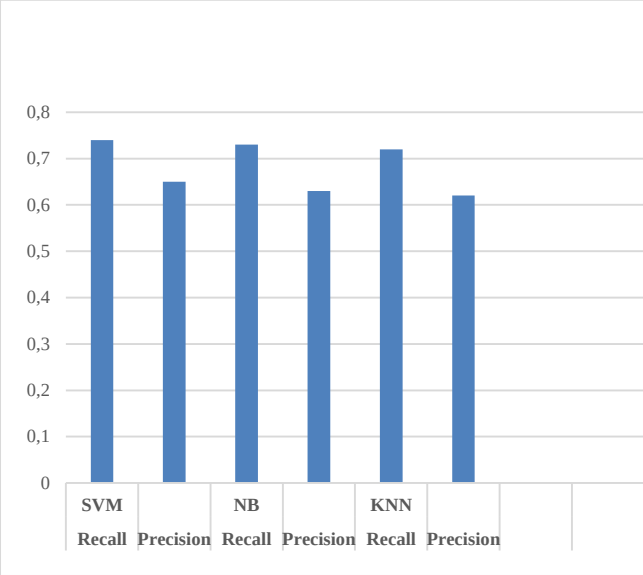


Table 1 shows the recall and precision values for all categories on the dataset for the three classifiers. The table shows that SVM has the best recall and precision values, followed by NB, and then KNN.

TABLE 1: VALUES OF RECALL AND PRECISION FOR THREE CLASSIFIER

Category Name	SVM		NB		KNN	
	Precision	Recall	Precision	Recall	Precision	Recall
Sports	0.97	0.891	0.962	0.884	0.95	0.871
Internet	0.662	0.241	0. 643	0.231	0.634	0.21
Art	0.412	0.862	0.402	0.842	0.386	0.811
Economy	0.401	0.978	0.392	0.964	0.378	0.943
Animals	0.921	0.674	0.903	0.654	0.896	0.642
Plants	0.941	0.593	0.933	0.587	0.912	0.567
Technology	0.492	0.398	0.478	0.379	0.463	0.362
Politics	0.994	0.454	0.983	0.441	0.976	0.434
Religion	0.873	0.601	0.865	0.599	0.851	0.583
Medicine	0.795	0.697	0.787	0.682	0.772	0.671

Another measure that was obtained from the experiments is the amount of time taken to build the models, which are used for testing the accuracy of the classifiers. This measure illustrates that NB takes the shortest amount of time to build the model, followed by KNN, and then SVM.

4. CONCLUSIONS

This study aims to compare three classification techniques using Arabic text documents which fall under four classes. The comparison is based on two main aspects of the classifiers: accuracy and time. In terms of time, results show that NB takes the shortest time to build the model, followed by KNN, and then SVM. On the other hand, SVM achieves the highest accuracy, followed by NB, and then KNN.

REFERENCES

Ababneh, J., Almomani, O., Hadi, W., El-Omari, N.K.T., and Al-Ibrahim, A., "Vector Space Models to Classify Arabic Text," International Journal of Computer Trends and Technology (IJCTT), vol 7, 2014

Agirre E., Lacalle O., and Soroa A., "Knowledge-Based WSD on Specific Domains: Performing Better than Generic Supervised WSD," *in Proceedings of the 21st International Joint Conference on Artificial Intelligence*, San Francisco, USA, pp. 1501-1506, 2009.

Alsaleem, S., " Automated Arabic Text Categorization Using SVM and NB," International Arab Journal of e-Technology, Vol. 2, 2011

Al-Harbi, S., Almuhareb, A., Al-Thubaity, A., Khorsheed, M. S. and Al-Rajeh, A. "Automatic Arabic Text Classification," Proceedings of The 9th International Conference on the Statistical Analysis of Textual Data, Lyon-France, 2008

Al-Kabi, M. N., & Al-Sinjilawi, S. I. (2007). a Comparative Study of the Efficiency of Different Measures To Classify Arabic Text. *University of Sharjah Journal of Pure & Applied Sciences*, *4*(2), 13–26.

Bawaneh, M.J., Alkoffash, M.S., and Al Rabea A.I."ArabicText Classification using K-NN and Naive Bayes". Journal of Computer Science, vol. 4, 2008.

Duwairi, R. "Arabic Text Categorization," The International Arab Journal of Information Technology, Vol. 4, 2007.
El-halees, A. (2011). Arabic Opinion Mining Using Combined Classification Approach. *Proceeding The International Arab Conference On Information Technology, Azrqa, Jordan.*

Gharib, T. F., Habib, M. B., & Fayed, Z. T. (2009). Arabic Text Classification Using Support Vector Machines. *International Journal of Computers and Their Applications*, *16*(4), 192–199. Retrieved from http://purl.utwente.nl/publications/75679

Hanandeh E., Mamoun S.The Automated VSMs to Categorize Arabic Text Data Sets,INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY,VOL 13, NO 1 (2014)MARCH-2014.PP.4047-4081

Khorsheed, M. S., & Al-Thubaity, A. O. (2013). Comparative evaluation of text classification techniques using a large diverse Arabic dataset. *Language Resources and Evaluation*, *47*(2), 513–538. http://doi.org/10.1007/s10579-013-9221-8
Khreisat, L. "A machine learning approach for Arabic text classification using N-gram frequency statistics," Journal of Informatics, Volume 3, 2009.

Karima, A, Zakaria, E and Yamina, T.G. "Arabic Text Categorization: A Comparative Study of different Representation Model, " Journal of Theoretical and Applied Information Technology, Vol. 38, 2005.

Mesleh, A.M.A. Support Vector Machine text Classifier for Arabic Articles: Ant Colony Optimization-based Feature Subset Selection., The Arab Academy for banking and financial Science, PHD. Thesis, 2008.

Syiam. M. M., Z. T. Fayed & M. B. Habib. An intelligent system for Arabic text categorization. IJICIS, Vol.6, No. 1 JANUARY 2006.

Wahbeh, A. H., Al-Radaideh, Q. A., Al-Kabi, , M. N., & Al-Shawakfa, E. M. (2012). A Comparison Study between Data Mining Tools over some Classification Methods. *International Journal of Advanced Computer Science and Applications, Special Issue on Artificial Intelligence, , 2*(8), 19–26.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

REMOVAL OF COPPER (II) AND IRON (III) MIXTURE BY PILOT NANOFILTRATION

DOI: 10.17261/Pressacademia.2017.601
PAP-WCTIE-V.5-2017(39)-p.278-285

Aoufi Boutheyna, Mohamed Amine Didi*
Laboratory of Separation and Purification Technologies, Department of Chemistry, Tlemcen University, Algeria. madidi13@yahoo.fr

ABSTRACT

In the current study, the effect of operating conditions such as pH value, feed flow, concentration of the solution and the applied pressure for the removal of copper(II) and iron(III) mixtures for the production of drinking water by nanofiltration membrane was investigated. The results show that it is possible to extract all of the iron (III) and copper (II) at the same time to a salt mixture of Fe 50% - salt Cu 50% for concentration 4 ppm, pH = 4.5 and pressure = 6 bars. The best results for the copper (II) were obtained for the various mixtures at the pressure of 6 bars at varying pH.

Keywords: Extraction, nanofiltration, copper (II), iron (III), mixture, synergism.

1. INTRODUCTION

The use of membrane technology has gained momentum in recent times and is preferred over conventional techniques like chemical precipitation or solvent extraction [1]. Advances over the last 10 years have shown a significant growth of papers published on nanofiltration (NF) membranes in many different areas. NF membranes in contact with aqueous solution are slightly charged due to the dissociation of surface functional groups or adsorption of charge solute. These properties have allowed NF to be used in niche applications in several areas especially for water and wastewater treatment, pharmaceutical and biotechnology, and food engineering [2]. in recent years, the use of nanofiltration membranes (NF) has increased rapidly in the chemical, petrochemical, biotech and desalination industries, since the NF technology over comes operational problems that are associated with conventional techniques. Several studies have been reported in which NF membranes have been used as tools for heavy metal removal [3].

Nanofiltration has some advantages over other membrane techniques, for example it has higher rejection of divalent ions and lower rejection of monovalent ions, lower operating pressure, higher flux and lower energy consumption compared with RO[4].

The aim of this work was to study the efficiency of copper (II) and iron (III) mixtures retention by using the SNTE NF270-2540. The effects of pressure and initial feed concentration on the membrane performance were studied.

2. EXPERIMENTAL

2.1 Apparatus

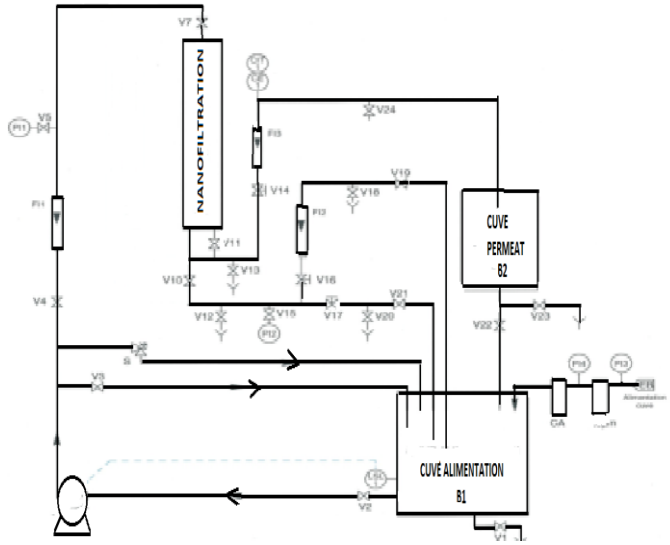
All chemicals used in this research were of analytical grade. All metal salt solutions were prepared by dissolving the appropriate weight of the salt of each metal in distilled water and made to a total volume of 50 L. The pH of solutions was measured using a pH-meter (Adwa), provided with a glass combine electrode. The conductivity measurements were carried by using MC126 Conductivity meter provided with an electrode.

The metal ion concentrations were determined by atomic absorption spectrophotometer (PINA cle 900 H - Perkin Elmer), using an air acetylene flame, two wavelengths were used 327.40 nm (linearity: 0.17 - 8 ppm) for copper and 302.06 nm (linearity: 0.4-20 ppm) for iron a range of standards solutions for various concentrations were prepared from a standard solution of 10 ppm for iron and 8 ppm for copper.

2.2 Pilot Equipment

Filtration was performed with a tangential filtration, capacity100 L (Figure 1). All the experiments were carried out in a closed system, where the permeate does not return to the tank whereas the retained liquid returns to the tank.

Figure 1: Schematic Diagram of Nanofiltration Pilot



2.3 Membrane Description

The module membrane spiral used was 1016 mm long and has a width of 61 mm. The nanofiltration membrane is a thin film composite membrane. All is established by three layers: A layer support in polyester (120 µm); a micro porous intercalary layer Polysulfone (40 µm) and a layer barrier (active layer) ultrathin of polyamide on the superior surface (0.2 µm).

Table 1: Characteristique of Membrane

Data from manufacturer		
Reference	Material	Cut-off (Dalton)
AFC 30	Polyamide on polyethersulfone	100
MRT 10	Polysulfone on polypropylene	200
MPT 08	Polyamide on polyethersulfone	200
MPT 34	Polysulfone on polypropylene	300
MPT 31	Polysulfone on polypropylene	400
MPT 04	polyethersulfone	>1000
SNTE NF270-2540 (used in present work)	**polyamide**	**200-400**

2.4 Extraction Procedure

The experiments were carried out at the Laboratory of Separation and Purification Technologies. After each experiment the membrane is cleaned by a hydrochloric acid solution for 10 min, and then it is rinsed with distilled water.

2.5 Analytical Methods

The volumetric flux was determined by measuring the permeate volume collected in given times interval. Owing to electro neutrality conditions, it was observed that both cation and anion rejection rates were the same, that is to say $R_{cation} = R_{anion} = R$. Consequently, the rejection rate can be calculated by Eq. (1):

$$R = 100\left(1 - \frac{C_p}{C_0}\right) \quad (1)$$

C_p concentration of salt in the permeate (ppm),

C_0 concentration of salt in the feed solution (ppm).

2.6 Results and Discussion

2.7 Determination of the Hydraulic Membrane Permeation

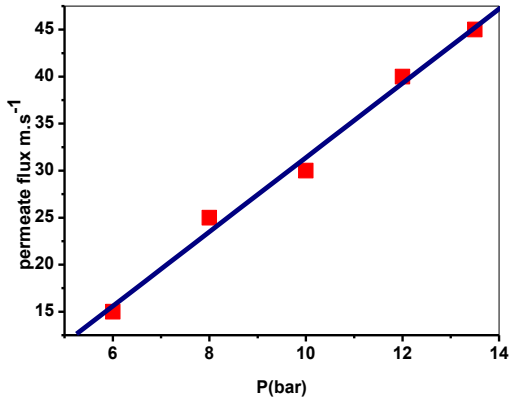
The permeability of the membrane is given by the slope of the Figure 3 which is equal with L_P= 3.95 m s^{-1} bar^{-1} and the resistance, R_m = 0.253 bar m^{-1} s, Eqn. (2).

$$R_m = \frac{1}{A} = \frac{S\Delta P_m}{Q_P} \qquad (2)$$

A permeability of the membrane
S membrane area
ΔPm the effective transmembrane pressure
Q_P permeation volume flow rate.

The value of the L_p obtained on the used membrane (SNTE NF270-2540) was proved to be 1.457; 106 times as large as that obtained on the membrane Nanomax-50 (Millipore USA)[5]. and on the Duramem MWCO 900, where L_P is 0.028 x 10^{-6} m s^{-1} bar what shows that our membrane is very successful, and can be used in the industrial scale[6] .

Figure 2: Permeate Flux Variation as a Function of Pressure for Distilled Water

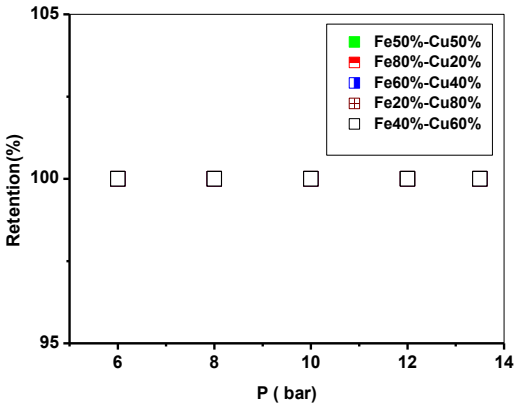


2.8 Effect of Pressure and Concentration for Mixtures

2.9 Effect of Pressure and Concentration for Iron

Figure 3 shows that, for variations of pressure from 6 - 13.5 bars, the retention is quantitative (100 %) for iron and copper at different proportions, while maintaining a total concentration of 4 ppm for the mixture. A pressure of 6 bars is enough for a full purification of iron. The difference between the iron and the copper is not observable, because the presence of the copper in the mixture does not influence the retention of the iron.

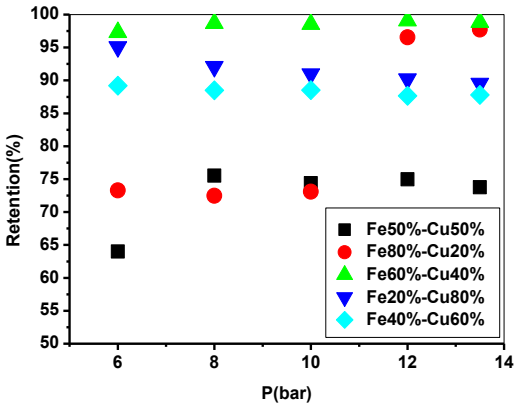
Figure 3: Variation of Iron Retention as a Function of the Pressure for Different Mixtures



3. Effect of Pressure and Concentration for Copper

The presence of iron influences the retention of the copper as seen from the results presented in Figure 4. The best retentions were obtained for the mixture Fe 60 % + Cu 40 % about pressure from 12 bars. For the mixture Fe 80 % + Cu 20 %, pressure up to 10 bars, the retention is constant (around 74 %) then increases exponentially until 97 % beyond this pressure. The effect of pressure was important on the retention, whatever the proportions of the mixture; the total concentration being always maintained at 4 ppm.

Figure 4: Variation of Copper Retention as a Function of the Pressure for Different Mixtures



Effect of pH

Effect of pH for the retention of iron and copper in the different mixtures

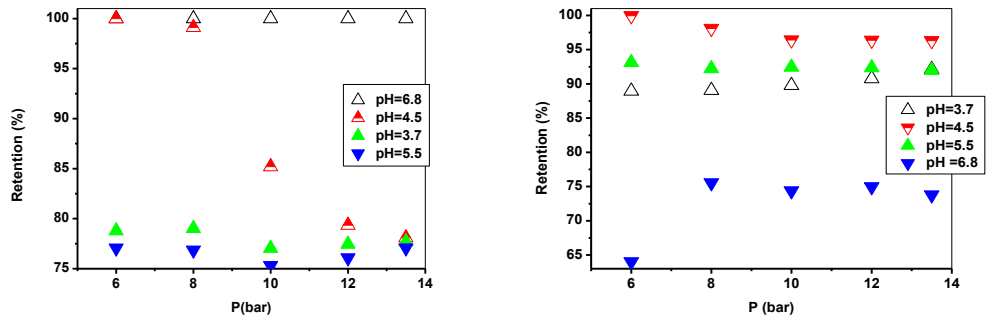
Mixture iron salt 50% - copper salt 50%

In view of the figure 5 (A), the obtained results show that the accepted pressure is about (6 - 13.5 bars) and the pH = 6.8; the retention of the iron is total (100 %). With pressure 6 bars, the retention of the iron is also total, in pH = 4.5. The difference at pH = 6.8 and pH = 3.7 then 5.5 becomes important. The best conditions of extraction of the iron correspond to pH = 4.5 and the pressure of 6 bars, with the addition of some mL of HCl. Without the addition of HCl, the best conditions are pH = 6.8 and the pressure of 6 bars.

For pH 3.7, 4.5 and 5.5, the retention of the iron at the pressure of 13.5 bars is almost the same (78 %) We can conclude that there is an interaction between these two parameters. A study of plan of experience would allow quantifying this interaction. The purification is total for the mixture in the pressure of 6 bars and in pH = 4.5.Whereas the best selectivity is obtained for a pH = 6.8 (without addition of HCl) and the accepted pressure of 6 bars.

Figure 5: Variation of iron (A) and copper (B) retention as a function of the pressure for mixture salt iron 50%-salt copper 50%

(A) (B)

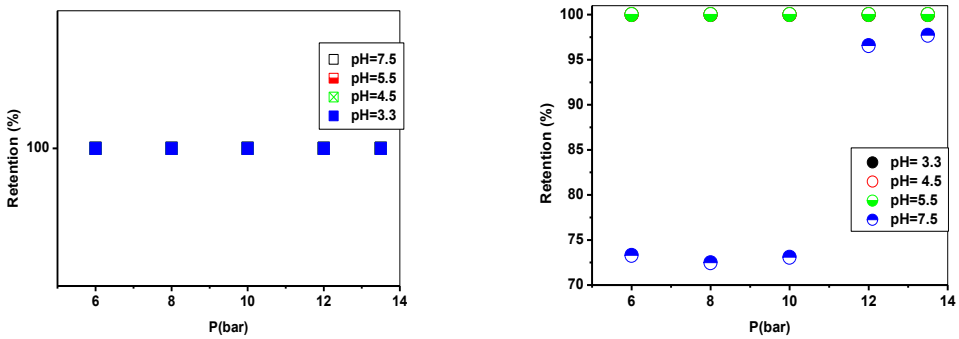


Mixture iron salt 80%-copper salt 20%

In view of the figure 6 (A), the obtained results show that the retention of iron is total (100 %) at pressure 6 - 13.5 bars and pH 3.3 - 7.5. In view of the figure 6 (B), for pH 3.3; 4.5 and 5.5 (with the addition of some mL of HCl) and a pressure from 6 - 13.5 bars the retention of copper is total (100 %). The membrane extracts the mixture without distinction between the iron and the copper, although these two metals have different physical-chemical properties; the iron Macke left some Ferromagnetic metals. The best selectivity is obtained for a pH = 7.5 (without addition of HCl) and the accepted pressure of 8 bars.

Figure 6: Variation of iron (A) and copper (B) retention as a function of the pressure for mixture salt iron 80%-salt copper 20%

(A) (B)

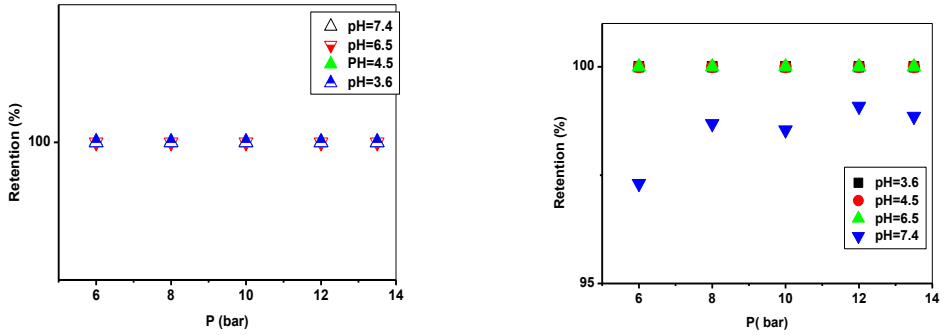


Mixture iron salt 60%-copper salt 40%

In view of the Figure7 (A), the obtained results show that the retention of the iron is total (100%) at pressures from 6 to 13.5 bars, and pH from 3.6 to 7.4. In view of the figure 14 (B), pH 3.6, 4.5, and 6.5 (with the addition of some mL of HCl) and a pressure from 6 to 13.5 bars the retention of the copper is total (100%). The separation takes place at P = 6 bars and pH = 7.4 which is a neutral pH. On these conditions of separation the speeds of diffusion through the membrane have different values.

Figure 7: Variation of iron (A) and copper (B) retention as a function of the pressure for mixture salt iron 60% - salt copper 40%.

(A) (B)



Mixture iron salt 20%-copper salt 80%.

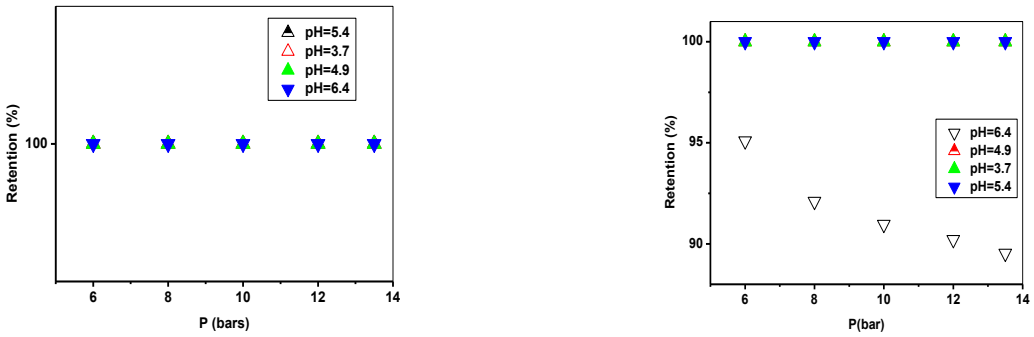
In view of the Figure 8(A), the results show that at pressures from 6 to 13.5 bars and pH from 3.7 to 6.4, the retention of the iron is total (100%). In view of the figure 8 (B), for pH 3.7, 4.9, and 5.4 (with the addition of some ml of HCl) and at pressures from 6 to 13.5 bars the retention of the copper is total (100 %). In these pH, the process does not make a difference between the iron and the copper.

In pH = 6.4 and at P =13.5 bars, the separation between both metals is the most important.

The retention decreases at pH = 6.4 with the increase of the pressure.

Figure 8: Variation of iron (A) and copper (B) retention as a function of the pressure for mixture salt iron 20%-salt copper 80%.

(A) (B)

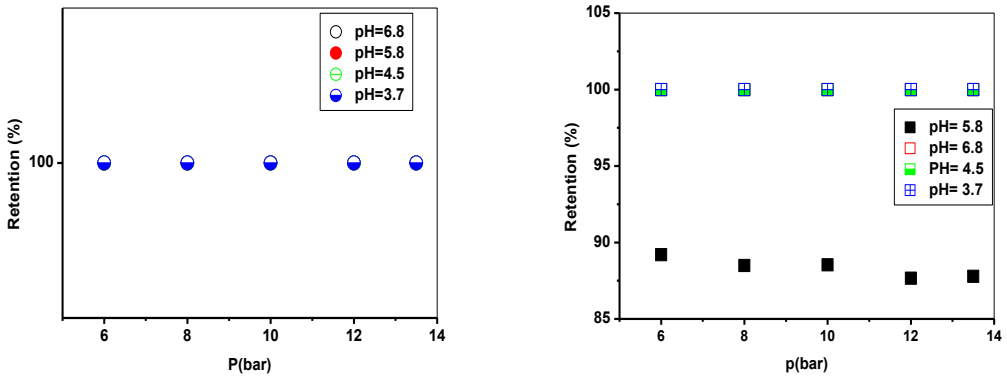


Mixture iron salt 40%-copper salt 60%

In view of the Figure 9 (A), the results show that the retention of the iron is total (100%) at pressure from 6 to 13.5 bars and at pH from 3.7 to 6.8. In view of the Figure 9 (B), the retention of the copper is total (100 %) for pH 3.7, 4.5, (with the addition of some ml of HCl) and 6.8 (with the addition of some ml of NaOH) at pressures from 6 to 13.5 bars. The difference of retention is obtained at pH = 5.8. In this pH, the effect of increase in pressure on the retention was weak.

Figure 9: Variation of iron (A) and copper (B) retention as a function of the pressure for mixture salt iron 40% - salt copper 60%.

(A) (B)



Effect of synergism.

- $\frac{Cp_m}{Cp_s} > 1$ **synergism** (the effect of the mixture is greater each of the ion in the mixture).
- $\frac{Cp_m}{Cp_s} < 1$ **antagonism** (the effect of the mixture is less than that each of the ion in the mixture).
- $\frac{Cp_m}{Cp_s} = 1$ **non interaction** (interaction the mixture has no effect on the adsorption of each of the adsorbates in the mixture) [7,8].

.

Table 2: Effect of Copper in the Mixture (*C*Ps = 0.22 ppm)

Mixture	$C_{P,m}$	$\frac{C_{P.m}}{C_{P,s}}$	Effect
Salt of Fe 50%- salt of Cu 50%	0.95	4.32	synergism
salt of Fe 80%- salt of Cu 20%	0.07	0.32	antagonism
salt of Fe 60%- salt of Cu 40%	0.03	0.14	antagonism
salt of Fe 20%- salt of Cu 80%	0.19	0.86	antagonism
salt of Fe 40%- salt of Cu 60%	0.40	1.82	synergism

Table 3: Effect of Iron in the Mixture (C_{Ps} = 0.21 ppm)

Mixture	$C_{P,m}$	$\frac{C_{P.m}}{C_{P,s}}$	Effect
salt of Fe 50%- salt of Cu 50%	0.00	0.00	antagonism
salt of Fe 80%- salt of Cu 20%	0.00	0.00	antagonism
salt of Fe 60%- salt of Cu 40%	0.00	0.00	antagonism
salt of Fe 20%- salt of Cu 80%	0.00	0.00	antagonism
salt of Fe 40%- salt of Cu 60%	0.00	0.00	antagonism

Several factors are considered to correlate metal ion uptake and metal ion properties. Factors like:

(i) electronagetivity of the metal ion,

(ii) electrostatic attrition due to charge to radius ratio,

(iii) ability to form metal hydroxide complex and

(İV) Suitable site for adsorption on adsorbent are responsible for competitive adsorption of one metal ion over another [9,10].

4. CONCLUSION

The results of the previous experiment show that NF270-2540 membrane is more efficient for copper and iron extraction due to its higher hydraulic permeability to distilled water (Lp = 3.95m s^{-1} bar^{-1}), A pressure of 6 bars is sufficient for such purification with respect to the iron, whatever the proportions Of the mixture and whatever the pH allowed.

The best results for copper (II) are obtained for the various mixtures of Pressure of 6 bar and variable pH . It is possible to extract all iron (III), copper (II) and at the same time for one

Mixture of 50% Fe salt - Cu 50% salt with a total concentration of 4 ppm, pH = 4.5 and the Pressure = 6 bar.

During the term of this study, the consideration of the difference (pH_r - pH_n), the retention of H^+ or of OH^- according to its sign, establishes a good descriptor of the imbalance of Ionic partitions realized between the solution and the pores of the membrane. The mechanisms of transfer of ions to be proposed in this study should allow a better understanding the selectivity observed during the nanofiltration.

Symbols

pH_r pH of permeate

pH_n pH of retentate.

Cp_m: concentration of an ion in the mixture ($Fe^{3+,}$ Cu^{2+}).

Cp_S: concentration of a single ion in solution.

REFERENCES

[1] V.V. Goncharuk, A.A. Kavitskaya, M.D. Skil'skaya, Nanofiltration in drinking water supply, J. Water Chem. Technol. **2011**, *33 (1)*, 37.

[2] Mohammad, A. W., Teow, Y. H., Ang, W. L., Chung, Y. T., Oatley-Radcliffe, D.L., Hilal, N., *Desalination,* **2015**, *356*, 226.

[3] Al-Rashdi, B., Somerfield, C., Hilal, N., *Sep. Purif. Rev.*, **2011**, *40*, 209.

[4] Mulder, M., *Basic Principles of Membrane Technology*, Kluwer Academic Publisher, Dordrecht, **1996**.

[5] Belkhouche, N., Didi, M. A., Taha, S., Benfarés, N., *Desalination,* **2009**, *239*, 58.

[6] Tsibranska, I., Saykova, I., *J. Chem. Technol. Met.*, **2013**, *48,* 333.

[7] Roy, A., Bhattacharya, J., *Sep. Purif. Technol.*, **2013**, *115*, 172.

[8] Mahamadi, C., Nharingo, T., *Bioresour. Technol.* **2010**, *101*, 859.

[9] Ben Frarès, N., Taha, S., Dorange, G., *Desalination,* **2005**, *185*, 245.

[10] Pontié, M., Lhassani, A., Diawara, C.K., Elana, A., Innocent, C., Aureaua, D., Rumeau, M., Croue, J.P., Buisson, H., Hemery, P., *Desalination,* **2004**, *167*, 347.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

COMPARATIVE STUDY ON THORIUM (IV) SORPTION ONTO SODIC BENTONITE AND MAGNETIC BENTONITE

DOI: 10.17261/Pressacademia.2017.602
PAP-WCTIE-V.5-2017(40)-p.286-292

Mohamed A. Didi, Abdelkader Miraoui
Laboratory of Separation and Purification Technologies, Department of Chemistry, Tlemcen University, Algeria. madidi13@yahoo.fr

ABSTRACT

In this paper, the liquid-solid extraction of Thorium (IV) is made by Sodic bentonite and Magnetic bentonite. Magnetic adsorbent can be quickly separated from a medium by a simple magnetic process, in view of these properties; some parameters were studied to assess the performance of maghemite nanocomposite clay for the removal of Thorium ions. The operating variables studied are initial La(III) concentration, pH, ionic strength, temperature and contact time. The time needed for magnetic bentonite to adsorb the maximum of Th (IV) is 45 minutes and 60 minutes for sodic bentonite. For magnetic bentonite, optimal extraction yield was achieved in a initial pH equal at 6.2 and for sodic bentonite, the variation of initial pH has no influence on the extraction yield. The sorption capacities of sodic bentonite and magnetic bentonite are 41.24 and 31.34 $mg.g^{-1}$ respectively. Adsorption equilibrium data were calculated for Langmuir and Freundlich isotherms. It was found that the sorption of Th (IV) on sodic and magnetic bentonite was better suited to the Langmuir adsorption model. Thermodynamics data leads to endothermic and spontaneous process for magnetic bentonite and exothermal for sodic bentonite. The quantitative elution study of Thorium can be realized with acetic acid for sodic bentonite and sulfuric acid for magnetic bentonite.

Keywords: Bentonite, nanoparticle, isother, thorium, process

1. INTRODUCTION

Like all heavy metal, thorium exhibits harmful character for nature and human health. Thorium toxicity is similar as that of uranium [1–3].

The presence of thorium in wastewaters could be hazardous for human health. That is why thorium removal from aqueous effluents still remains a major issue to be addressed. A brief overview of the wide literature produced in this regard reveals the use of solvent extraction with various commercial extractants, liquid membrane and ion exchange resins [4, 5].

The extraction of Thorium is performed by γ-Fe_2O_3 enriched with bentonite clay and sodic bentonite, the objectives of this study are to assess the performance of maghemite nanocomposite clay for the removal of Thorium ions. Effects of pH and temperature on the adsorption process are also investigated. The adsorption capacity has been studied using the adsorption isotherm technique. Various isotherm equations are used to determine the best isotherm equation which represents the experimental data adequately and satisfactorily. The kinetic adsorption is also studied and tested using different models. Further studies can explore the possible regeneration of nanocomposite clay for reuse.

2. EXPERIMENTAL

2.1 Reagents

Thorium solution at 10^{-2} M was prepared by dissolving of Th $(NO_3)_4.4\ H_20$ (from FLUKA) (0.552 g) in 100 mL of distilled water. The initial pH of the sample solutions were adjusted by using dilutes HNO_3 or NaOH (from Sigma-Aldrich). $NaNO_3$ (from Merck), $C_2H_3NaO_2$ and NaS_2O_3 were used in the salt effect. Arsenazo III 10^{-3} M (from Fluka) was prepared by dissolving 0.0820 g in absolute ethanol. Hydrochloric acid (from Organics), sulfuric acid (from Fluka), nitric acid (from Cheminova), and acetic acid (from Riedel Dehaen) were used from elution study.

The natural bentonite used in this study was obtained from deposits in the area of Maghnia, Algeria. For magnetic nanoparticles synthetized, $FeCl_2.4H_2O$, $FeCl_3.6H_2O$, NH_4OH, HNO_3 and Fe $(NO_3)_2$ were used.

2.2 Apparatus

The extraction of Th (IV) was studied by the batch process using a stirring vibrator (Haier model). pH measurements were performed with a pH meter using a combined electrode mark (Adwa). A magnet and centrifugation for the recovery of the magnetic particles and sodic bentonite, respectively, in the aqueous phase were used. Samples containing Th (IV) were analyzed by spectrophotometer (Analytik Jena Specord 210 Plus) with ArsenazoIII as ligand.

2.3 Preparation of Magnetic Particles

The ferrofluid magnetic used was the maghemite (γ-Fe_2O_3) nanoparticles dispersed in an aqueous solution. Particles were synthesized by coprecipitation of a stoichiometric mixture of ferrous and ferric chloride in an ammonium hydroxide solution. The precipitate magnetite (Fe_3O_4) obtained was acidified by nitric acid and oxidized into maghemite (γ -Fe_2O_3) at 90°C with iron (III) nitrate. The maghemite particles obtained were precipitated by the acetone, then dispersed into water leading to an ionic ferrofluid acid (pH=2.0). After these step, nanoparticles were positively charged, with nitrate as counter ions.

2.4 Preparation of Magnetic Bentonite

The composites can be prepared by the following procedure:

Dissolve $FeCl_3$ (7.8 g, 28 mmol) and $FeSO_4$ (3.9 g, 14 mmol) in 400 mL solution at 70°C. Add the clay bentonite at 3.3, 6.6 or 9.9 g in order to obtain the following adsorbent: iron oxide weight ratios 1:1, 1.5:1 and 2:1. To this suspension add a solution of NaOH (100 mL, 5 mol L^{-1}) drop wise to precipitate the iron oxides. Wash the obtained solid materials with distilled water and dry in an oven at 100°C for 2 h. After the preparation a simple test with a magnet (0.3 T) can be carried out showing that the whole material is completely attracted to the magnet.

2.5 Extraction and Analysis Procedure

The method of extraction used for this study, was carried out by a mixture of 4 mL of Th(IV) solution of known concentration, and 0.01 g of our solid adsorbent (bentonite mineral and magnetic bentonite) in Erlenmeyer with stopper, under vigorous stirring. Both liquid and solid phases were separated by centrifugation for bentonite mineral and magnet for magnetic bentonite, the solid phase will be regenerated for other applications and the liquid phase was measured by the UV-visible spectrometer. The sample of Th (IV) was analyzed by a mixture of 100 µL ArsenazoIII and 100 µL of Th (IV) in a medium HCl 9M (2 mL). The interaction of Arsenazo (III) with Th (IV) can be estimated at λ_{max}= 660 nm.

The percentage of Thorium ions that was extracted by solids extractant was determined as (Eq. 1):

$$Extraction\,\% = \frac{Ci - Ce}{Ci} * 100 \qquad (1)$$

The amount of Thorium uptakes at time t, qt (mg g^{-1}), was calculated by Eq. 2:

$$qt\left(\frac{\text{mg}}{\text{g}}\right) = \frac{(Ci - Ct)}{W}.V.M \qquad (2)$$

where C_i, C_t and C_e are the initial, time t and equilibrium Th (IV) concentration (mol L^{-1}), respectively, V (4 mL) is the volume of the solution, M molecular weight (g mol-1), and w is the mass of the solids adsorbents (0.01 g).

3. RESULTS AND DISCUSSION

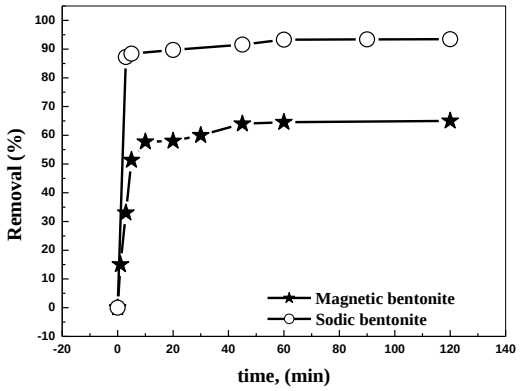
In this study different parameters are examined: contact time, effect of initial pH, initial concentration of Thorium, the effect of ionic strength.

3.1 Effect of Time Contact

The equilibrium time is the necessary and sufficient time to reach the equilibrium exchange of Th (IV) between the aqueous phase and the adsorbents. To perform this effect, we will fix the concentration initial of Thorium (10^{-4} mol L^{-1}), values of initial pH, and adsorbents amount (0.01g). Figure 1 illustrate the extraction yield as a function of time, where the extraction efficiency increases rapidly with increasing of time. The time needed for magnetic bentonite to adsorb the maximum of Th (IV) is 45 minutes (58%, 8.26 mg.g^{-1}) and 60 minutes for sodic bentonite (89%, 8.65 mg.g^{-1}).

Fig. 1: Removal of thorium by magnetic and sodic bentonites as a function of time

$[Th(IV)]_0=10^{-4}$ mol L^{-1}, w = 0.01 g, V = 4 mL, Ø = 250 rpm.

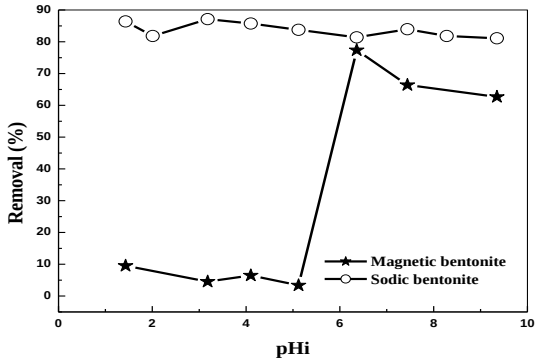


3.2 Effect of pH

Sorption of thorium by magnetic bentonite and sodic bentonite were studied at different pH value ranged from 1,43 to 9,35 and the results are show in Figure 2, it is seen that the extraction yield is very low and stable in acidic region and the removal of thorium begins to increase with increase in pH and remains maximum in the pH 6,36. With further increase in pH, the sorption efficiency decreases. This observation could be explained as follows: At pH in the range of 1.4–5.1, hydrogen ions compete strongly with metal thorium (IV) ions for the active sites, which results in less metal sorption. When solution pH is from 2.0 to 4.5, the competing effect of hydrogen ions decreases, instead, thorium (IV) ions sorption increases. When pH is 6.2, a large number of active adsorption sites are released, and the maximum number of adsorption sites are obtained, which provide lots of opportunities for thorium (IV) ions attack, so there are the maximum thorium(IV) sorption efficiency at the pH values 6.2. Beyond pH 6.2, insoluble thorium hydroxide started precipitating, which led to low thorium (IV) ions sorption efficiency at equilibrium. By against the variation initial pH of the stock solution of thorium, has no influence on the extraction yield for sodic bentonite, which is between 80% and 85%.

Fig. 2: Removal of thorium by magnetic and sodic bentonites as a function of initial pH.

$[Th(IV)]_0 = 10^{-4}$ mol L^{-1}, w = 0.01 g, V = 4 mL, Ø = 250 rpm.



3.3 Effect of Initial Metal Concentration

Several experiments were also undertaken to study the effect of varying the initial thorium concentration on uptake (q) from the solution by 0.01 g of our adsorbent. The amount of Th(IV) sorbed per unit mass of the particles magnetic increased with the initial metal concentration.

Fig. 3: Removal of thorium by magnetic and sodic bentonites as a function of [Th(IV)].

w = 0.01 g, V = 4 mL, Ø = 250 rpm.

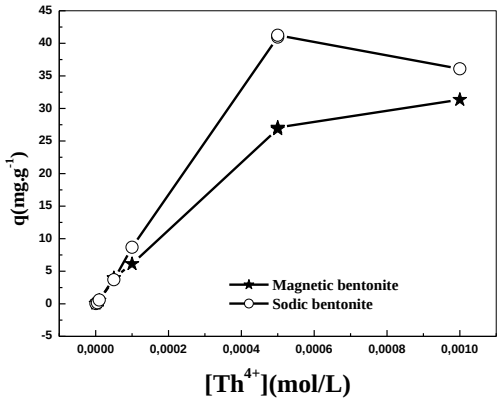


Figure 3 show that the maximum sorption capacities for the metal ions were 41.24 $mg.g^{-1}$ (0.177 $mmolg^{-1}$) for sodic bentonite and 31.34 $mg.g^{-1}$ (0.135 mmolg-1) for magnetic bentonite, this values indicate that sodic bentonite and magnetic bentonite were effectives sorbents in treatment of diluted thorium solutions.

3.4 Isotherm Adsorption

The sorption data, commonly known as adsorption isotherms, are basic requirements for the design of adsorption systems. Classical adsorption models, Langmuir and Freundlich, were used to describe the equilibrium between adsorbed Th (IV) ions on the sodic and magnetic bentonite site.

In this concentration range, the Langmuir isotherm correlated better than Freundlich isotherm (Table 1) with the experimental data from adsorption equilibrium of thorium ions by magnetic bentonite and sodic bentonite, suggested a monolayer adsorption. The maximum adsorption values were in accordance with the values obtained experimentally (Table 1).

Fig. 4: Langmuir isotherm plot for the sorption of Th(IV) onto sodic

and magnetic bentonite, w = 0.01 g, V = 4 mL, Ø = 250 rpm.

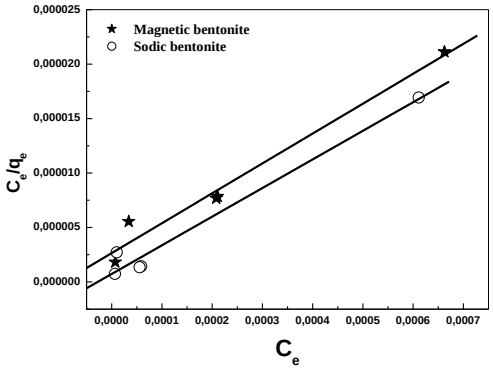


Fig. 5: Freundlich isotherm plot for the sorption of Th(IV)

onto sodic and magnetic bentonite, w = 0.01 g, V = 4 mL, Ø = 250 rpm.

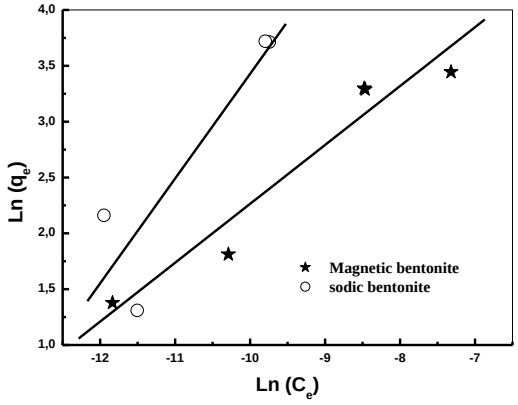


Table 1: Isotherm models parameters for the adsorption of Th(IV)

on magnetic bentonite and sodic bentonite

	q_m (exp.), $mg.g^{-1}$	Langmuir isotherm	Freundlich isotherm
Sodic bentonite	41.24	R = 0.9883 q_m(calc.)= *38,03* $mg.g^{-1}$ K_L=9927.7	R =*0,8983* K_F=362217,44 n=0,9380
Magnetic bentonite	31.34	R =*0.9863* q_m(calc.)= *36.44* $mg.g^{-1}$ K_L=*10362.69*	R =*0,9663* K_F=1881,83 n=0,527

3.5 Effect of Temperature and Thermodynamic Parameters

The effect of temperature on the sorption of thorium from nitrate solution by sodic bentonite and magnetic benonite at pH 6.2, V=4 mL, w=0.010 g, Ø=250 rpm and concentration Th(IV) 10^{-4} mol L^{-1} is studied. This study is efficient for the determination of thermodynamic data such as, the Gibbs free energy change (ΔG), enthalpy change (ΔH) and entropy change (ΔS). ΔG is calculated using the following equations:

$$\Delta G = \Delta H - T\,\Delta S \quad (3)$$

$$\Delta G = - RT\,LnD \quad (4)$$

where R is the gas constant (8.314 J $mol^{-1}K^{-1}$), and T the temperature (K).

Figure 6 shows removal percent (%) of thorium ion onto sodic bentonite and magnetic bentonite as a function of the temperature. It can be seen that the yield extraction of Th (IV) increases with increasing temperature for magnetic bentonite, this behaviour indicates that thorium sorption onto magnetic bentonite is an endothermic and spontaneous process, as supported by the positive values of ΔH and ΔS (Table 2); decrease in ΔG values with increase in temperature showed that the sorption was most favourable at higher temperature. In contrast, increasing temperature from 25 to 50° C was found to have a detrimental effect on the extraction process for sodic bentonite, the negative value of ΔH indicate that the extraction process is exothermal.

Fig. 6: Removal of thorium by magnetic and sodic bentonites as a function of temperature. $[Th(IV)]_0=10^{-4}$ mol L^{-1}, w = 0.01 g, V = 4 mL, Ø = 250 rpm.

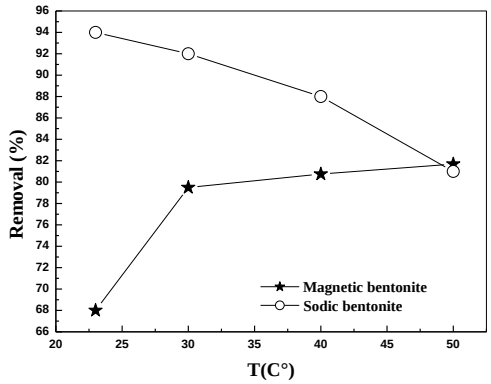


Fig. 7: Plot of Eq.6 for the thorium sorption on by magnetic and sodic bentonites, $[Th(IV)]_0=10^{-4}$ mol L^{-1}, w = 0.01 g, V = 4 mL, Ø = 250 rpm.

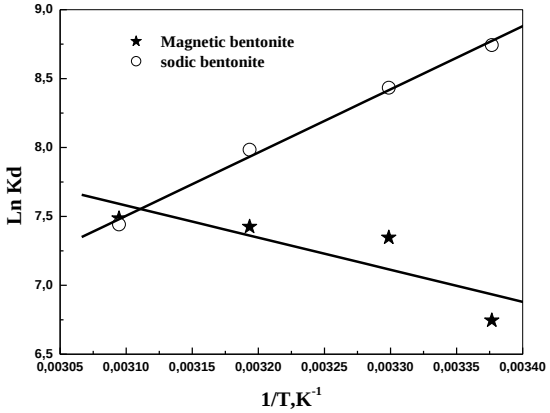


Table 2: Thermodynamic parameters for the sorption of Th(IV) on sodic and magnetic bentonite

Bentonite	ΔH (kJ.mol^{-1})	ΔS (J/K)	ΔG(kJ/mol)			
			T=296 K	T=303 K	T=313 K	T=323K
Sodic bentonite	-38,138	-55,78	-21,62	-21,23	-20,67	-20,11
Magnetic bentonite	5,699	79,897	-17,962	-18,52	-19,32	-20,12

4. CONCLUSION

In this paper, Liquid-solid extraction of Thorium (IV) is made by sodic bentonite and magnetic bentonite. The extraction efficiency was determined as a function of various parameters such as time, initial pH, Thorium concentration and temperature. The experimental capacity obtained was 31.34 mg.g^{-1} for magnetic bentonite and 41.24 mg.g^{-1} for sodic bentonite. The sorption of Th (IV) achieves equilibration at 45 minutes and 60 minutes for magnetic bentonite and sodic bentonite, respectively. Optimal extraction yield was achieved in a initial pH equal at 6.2 for magnetic bentonite, By against, the variation of initial pH has no influence on the extraction yield.

The results obtained in this study are very close and there is not a big difference, but magnetic bentonite have a unique superiority in separation, it can be quickly separated from a medium by a simple magnetic process, in view of this properties and by the results obtained in this study, magnetic bentonite can be a promising material for sorption,

immobilization and pre-concentration of rare earth elements, radioactive metal and heavy metal ions from large volume of solutions.

REFERENCES

1. Ishikawa Y, Wada I, Fukumoto M (2001). J Environ Pathol Toxicol Oncol 20(4): 311–315.

2. Martling U, Mattsson A,Travis L B, Holm L E, Hall P (1999). Radiat Res 151: 293–299.

3. VanKaick G, Bahner M L, Liebermann Deta l (1999). Radiologe 39(8): 643–651.

4. Tavlarides L L, Bae J H, Lee C K (1987). Separation Science and Technology. 22(2–3): 581–617.

5. Singh H, Gupta C K (2000). Miner Proc Extr Metal lRev. 21(1–5): 307–349.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Sener

ENHANCED WEB CACHE REPLACEMENT POLICY BASED ON DATA MINING AND RFSD SCORING

DOI: 10.17261/Pressacademia.2017.603
PAP-WCTIE-V.5-2017(41)-p.293-298

Mohammad Edris Rajaby, Tuğrul Çavdar
Computer Engineering Department, Karadeniz Technical University, Trabzon 61080, Turkey. edris.rajaby@gmail.com, ulduz@ktu.edu.tr

ABSTRACT

By rapid growth of the Internet users and devices, the number of servers also increase simultaneously which causes the exponential increment of the internet traffic and static data. Handling these huge amounts of user requests and efficiently responding to them require high bandwidth links, powerful servers and robust equipment, which despite the availability of these requirements getting the full user satisfaction is extremely difficult and a tough challenge. In order to overcome the mentioned problem, the cache servers are being used as a suitable solution. The performance of web cache server directly depends on its replacement policies. Several cache replacement policies have been proposed in literature each having varied hit rate (HR) and byte hit rate (BHR) performances on different networks. The replacement policy proposed in this paper is a dynamic cache replacement policy which trains itself utilizing previous network logs and by exploiting the data mining clustering algorithm. Once the training step is completed, the proposed policy utilizes the normalization formulas to score each metric of the enquiries including recency, frequency, size and delay. Simulation results showed that the proposed policy has the optimum performance on different networks and it not only improved the performance of web cache server in term of HR and BHR, but also decreases the data retrieval time (Delay Ratio (DR)) of the cache servers.

Keywords: cache replacement policy, cache server, proxy server, scoring

1. INTRODUCTION

The recent tremendous developments in the technology field had negative effects on the web performance in terms of network congestion, server load, data retrieval time and bandwidth consumption. To tackle the above issues, different solutions are available such as incrementing the link and server capacities and deploying backup servers in multiple locations. In addition to being very costly, none of the mentioned enhancements can provide a solid and optimum resolution. Cache servers have been implemented as an efficient solution that is capable of decreasing the data retrieval time, reducing the network congestion, lowering the server loads, preventing the need for high bandwidth links and eventually resulting in customer satisfaction (Davison, 2001) and (Podlipnig & Böszörmenyi, 2003).

The cache server deployment is shown in figure 1. The performance of web cache server directly depends on its replacement policies (Zhang, 2015).

Various web cache replacement policies are proposed by (Zhang, 2015), (Sulaiman, Mariyam Shamsuddin, Forkan, & Abraham), (Ali, Sulaiman, & Ahmad, 2014), (Jeon , Lee, Cho, & Ahn, 2003), (Jarukasemratana & Murata, 2013) and (Abdalla, Sulaiman, & Ali, 2015). Besides the cache HR and BHR that play important roles in web cache server performance, DR also has a significant impact on the performance of web cache servers. Furthermore, most of the traditional and proposed algorithms are based on the Recency such as LRU (Least Recently Used), Frequency such as LFU (Least Frequently Used) and Size factor. However, traditional policies being used currently are not sufficient as their decision depends on one factor while ignoring most others (Abdalla, Sulaiman, & Ali, 2015).

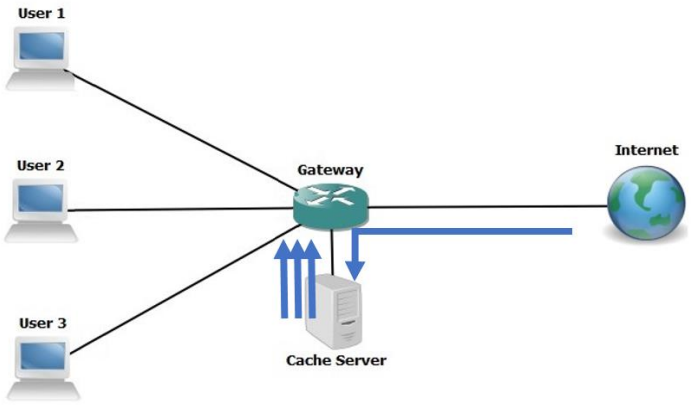


Figure 1: Deployment of cache server

Wong (2006), categorized the cache replacement policies into five categories and provided the design rationale along with the advantage of each category as bellow:

1- Recency-based: Recently requested objects will be requested again in the near future. It is good when the same object is requested by many users at the same time.
2- Frequency-based: The most popular objects should be cached. It is efficient when the web sites having objects with quite steady popularities are accessed by users.
3- Size-based: Keeps smaller objects by removing larger one. It is good when information-based web sites are accessed by users.
4- Function-based: Higher hit ratio can be achieved by considering more parameters. It has better performance when the system has high processing and memory resources.
5- Randomized: The simplest structure with no need for high computation overhead. It is good for systems with limited processing and memory resources.

According to the simulation results of the proposed policies in the literature, the performance of different policies varied on different datasets and can't guarantee the optimum performance on all networks. Furthermore, the reconfigurability of the replacement policy is irrecusable criteria. The replacement policy proposed in this paper is a dynamic cache replacement policy that analyses the network according to previous logs and afterward based on the analyzation result weighs and scores the RFSD metrics. Subsequently, using the metric weights and scores, a general score will be generated for each unique object; and based on this general score the object with the lowest score will be deleted from the cache server memory. The paper is organized as follow: section 1 provides overall information about this paper, presents related works done in literature and introduces the proposed solution. Section 2 is Methodology and explains the proposed methods in detail. Section 3 presents the simulation results and finally section 4 summarizes the paper.

2. METHODOLOGY

Data Preproccessing: The data collected from online resources usually has many errors and cannot be used for data analysis purposes directly due to its incompleteness, inconsistency and lack of specific behaviors and trends. Data preprocessing is a verified method for resolving such issues and it is a crucial step in the data mining procedure. Analyzing data without preprocessing can produce ineligible results. Data pre-processing consists of cleaning, normalization, transformation, feature extraction and selection, etc. The outcome of data pre-processing is the final training data set that could be used in data mining process (Kumar & Reddy, 2014).

In order to be accepted as a cacheable URL, we selected the URLs which not contain substrings such as 'cgi-bin' or '?', a file extension such as '.cgi', and if the response produced from origin server has a proper status code (e.g., 200 Success) (Arlitt, Cherkasova, Dilley, Friedrich, & Jin, 2000).

The flowchart of the proposed cache replacement policy is shown in figure 2. After preprocessing step the proposed policy perform bellow phases:

2.1. RFSD Scoring

The bellow formula (1) is used to score the Recency, Frequency and Delay (data retrieval time) metrics.

$$Score_{rfd} = 8 \times \left(\frac{(Val_{cur} - Val_{min})}{(Val_{max} - Val_{min})}\right) + 1 \text{ ... (1)}$$

The above formula (1) is modified to score the Size metric of the objects. Unlike the RFD, the size metric is needed to be scored in inverse order. The smaller objects take the higher scores and the bigger objects take the lower scores.

$$Score_{s} = 8 \times \left(\frac{(Val_{max} - Val_{cur})}{(Val_{max} - Val_{min})}\right) + 1 \text{ ... (2)}$$

In the formula (1) and (2) shown above, $Score_{rfd}$ is the score of the recency, frequency and delay metrics, $Score_s$ is the score of the size metric, Val_{max} is the maximum value of the corresponding metrics, Val_{cur} is the current objects metric value and the Val_{min} is the minimum value of the corresponding metrics. 8 and 1 are the constant numbers which were used to score the metrics between 1 and 9.

2.2. Policy Adjustment Using K-Means Algorithm

One of the common and popular clustering algorithms in Data Mining is the K-Means algorithm which uses the distance between each data point and the centroid of the cluster to partition the data points. In this method, the data points are assigned to the nearest centroid according to a specifically selected proximity measurement. Afterward, the centroids of all clusters are updated and these two steps are repeated until the centroids are stable. There are varied measurements in K-Means algorithm for computing the cluster centroids such as Manhattan distance, Euclidean distance, and Cosine similarity. Overall, the Euclidean measure is commonly used in K-means algorithm (Aggarwal & Reddy, 2013). Euclidean distance was used in this paper according to formula (3).

$$d(m,n) = \sqrt{|m_1 - n_1|^2 + |m_2 - n_2|^2 + \cdots + |m_i - n_i|^2} \text{ ... (3)}$$

Where *m* and *n* are the data points and *d* is the distance between them.

To train the proposed algorithm based on the previous logs, the number of clusters K was defined as one cluster. The algorithm should calculate the mean of the cluster by calculating the mean for each variable (RFSD) as illustrated in table 1.

Table 1: Calculated means for cluster

Variables	Dataset (a)	Dataset (b)
Recency	6.453	7.054
Frequency	5.002	6.765
Size	7.012	4.343
Delay	4.234	3.243

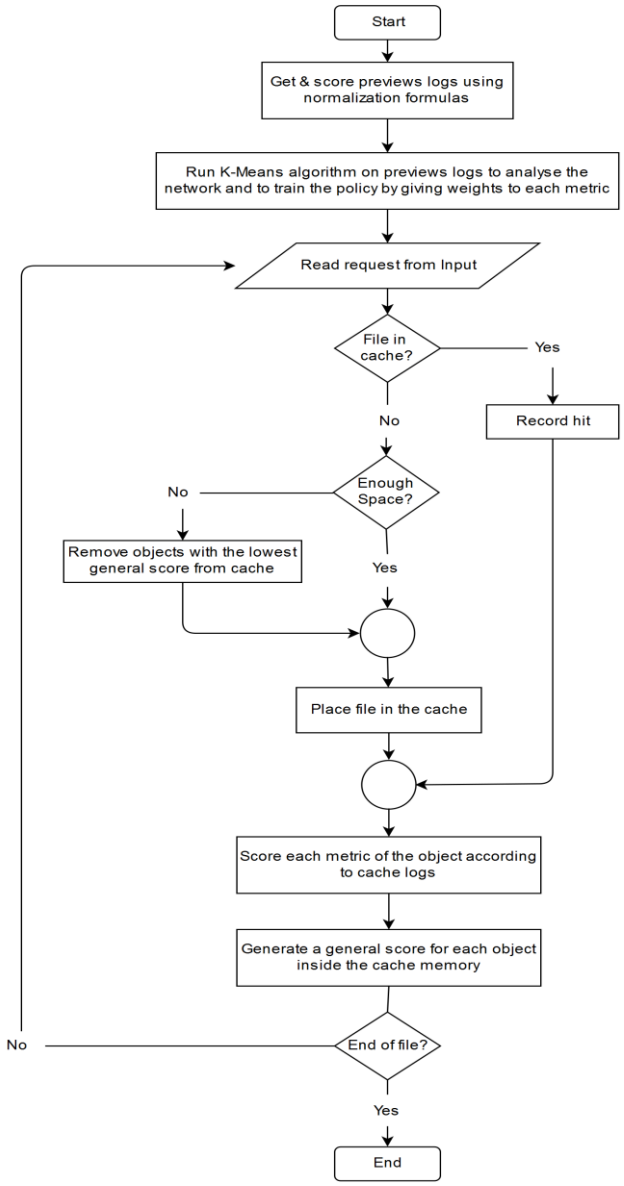


Figure 2: Flowchart of the proposed policy

The proposed algorithm weighs the metrics according to the means calculated by the K-means clustering algorithm. For example based on the above results in table 1 for the dataset (a) the algorithm weighs the size as 10^4, the recency as 10^3, the frequency as 10^2 and the delay as 10. If an object has the scores of recency= 8, frequency= 7, size= 9 and delay equal=5, the general score generated by policy for this object should be 9875.

3. SIMULATION RESULTS

The dataset used in this experiment was the BU Web traces (Boston University, 1995) dataset provided by Cunha of Boston University. This dataset is contained of 9633 files, and recorded 1,143,839 web requests from different clients during six months. BU traces consist of 37 client machines divided into two sets: undergraduate students set (called 272 set) and graduate students set (called B19 set). In this work, the dataset 272 is called as the dataset (a) and the dataset B19 as the dataset (b). The dataset (b) has 32 machines and the dataset (a) has 5 machines (Ali & Siti, 2009). In this experiment, one day logs (13-December) of the dataset (a) was used to train the policy and one day logs (14-December) of the dataset (a) to test the performance of the policy. Furthermore, to prove that the proposed policy has the optimum performance on all networks the policy has been tested on the second dataset (database (b)). In the second dataset, one day traces (17-January) was used to train the policy and one day traces (18-January) to test the performance of the policy.

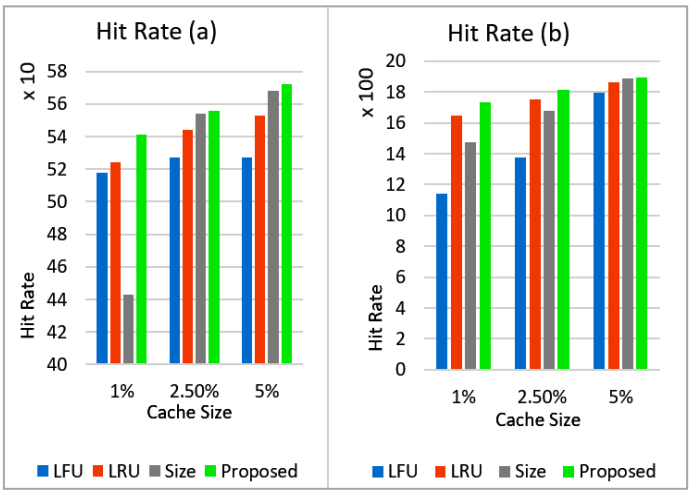


Figure 3: Hit rate of the policies for dataset (a) and dataset (b)

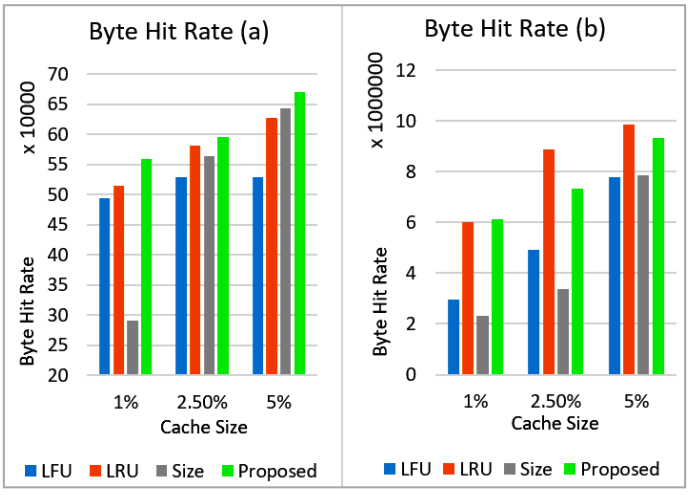


Figure 4: Byte hit rate of the policies for dataset (a) and (b)

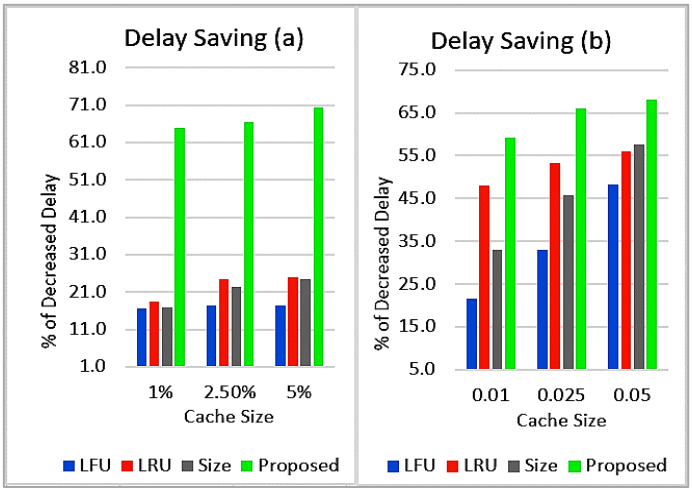


Figure 5: Delay saving of the policies for dataset (a) and dataset (b)

In cases where the size of the cache is very large, almost all policies have good performance. The best policy is the one that has the best performance in case of limited and small cache size. According to the simulation results as demonstrated in figurs. 3, 4 and 5, the proposed cache replacement policy is compared with the conventional algorithms such as LRU (Least Recently Used), LFU (Least Frequently Used) and Size. As shown in figure 3. the proposed policy has the best performance in datasets (a) and (b) in term of HR, and based on figure 4. the proposed policy provides the best performance for BHR as well in cases that the cache size is small. Moreover, as proved in figure 5. one of the significant competencies that the proposed algorithm has is its terrific increment of cache server performance in terms of delay ratio compared to traditional algorithms such as LFU, LRU and Size.

5. CONCLUSION

Cache servers play a significant role to improve the performance of the Internet, and the performance of cache servers directly depends on the cache replacement policies used in them. In this work, a new cache replacement policy has been proposed which is based on the data mining k-means clustering algorithm and RFSD scoring method. Numerous replacement policies have been proposed in literature, but most of them used one or two metrics of the requested objects, their performance on various networks differs and they do not provide solid and optimum solution on all networks. Using the k-means algorithm and previous network logs, the proposed policy is capable of adjusting itself to any network structure. It scores the RFSD metrics by modifying the normalization formulas and generates a general score for each object. Subsequently, the object with the lowest general score will be deleted when the cache is full and the new requested object will be cached.

The performance analysis shows that one of the significant competencies that the proposed algorithm has is its terrific increment of cache server performance in terms of delay ratio compared to traditional algorithms such as LFU, SIZE and LRU. Concurrently, it is demonstrated that the proposed policy remarkably improves the performance of the cache servers with regard to HR and BHR on any network.

REFERENCES

Abdalla, A., Sulaiman, S., & Ali, W. (2015). Intelligent Web Objects Prediction Approach in Web Proxy Cache Using Supervised Machine Learning and Feature Selection. Int. J. Advance Soft Compu. Appl, 7(3).

Aggarwal, C. C., & Reddy, C. K. (2013). Data clustering: algorithms and applications. Boca Raton: CRC Press.

Ali, W., & Siti, M. (2009). Intelligent client-side web caching scheme based on least recently used algorithm and neuro-fuzzy system. International Symposium on Neural Networks. Berlin.

Ali, W., Sulaiman, S., & Ahmad, N. (2014). Performance Improvement of Least-Recently-Used Policy in Web Proxy Cache Replacement Using Supervised Machine Learning. International Journal of Advances in Soft Computing & Its Applications, 6(1).

Arlitt, M., Cherkasova, L., Dilley, J., Friedrich, R., & Jin, T. (2000). Evaluating content management techniques for web proxy caches. ACM SIGMETRICS Performance Evaluation Review, 27(4), 3 - 11.

Boston University, (1995). BU Web Trace, http://ita.ee.lbl.gov/html/contrib/BU-Web-Client.html.

Davison, B. (2001). A Web caching primer. IEEE internet computing, 38 - 45.

Jarukasemratana, S., & Murata, T. (2013). Web Caching Replacement Algorithm Based on Web Usage Data. New Generation Computing, 31(4), 311–329.

Jeon , J., Lee, G., Cho, H., & Ahn, B. (2003). A prefetching Web caching method using adaptive search patterns. 2003 IEEE Pacific Rim Conference on Communications Computers and Signal Processing (PACRIM 2003) (Cat. No.03CH37490) . IEEE.

Kumar, P., & Reddy, D. (2014). Novel Web Proxy Cache Replacement Algorithms using Machine Learning Techniques for Performance Enhancement. INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, 3(1), 339-346.

Podlipnig, S., & Böszörmenyi, L. (2003). A survey of Web cache replacement strategies. ACM Computing Surveys (CSUR), 374-398.

Sulaiman, S., Mariyam Shamsuddin, S., Forkan, F., & Abraham, A. (n.d.). Intelligent Web caching using neurocomputing and particle swarm optimization algorithm. 2008 Second Asia International Conference on Modelling & Simulation (AMS). IEEE.

Wong, K.-Y. (2006). Web cache replacement policies: a pragmatic approach. IEEE Network, 20(1), 28-34.

Zhang, J. (2015). Replacement Strategy of Web Cache Based on Data Mining . 2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC).

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

AN INTEGRATED SURVEY IN AFFILIATE MARKETING NETWORK

DOI: 10.17261/Pressacademia.2017.604
PAP-WCTIE-V.5-2017(42)-p.299-309

Ali Norouzi
Department of Business Management, Istanbul Aydin University, Florya, Istanbul, Turkey. norouzi@ieee.org

ABSTRACT

As the Internet plays more colorful role in our everyday lives, the consumer's shopping habits is changing toward increase in share of online purchase. Along with expansion of the Internet in people's everyday life the recent years have been featured with growth of online marketing among businesses. Firms have come to understanding that online marketing is a vital element for increasing brand awareness and grab attention of the modern consumers. Moreover, there has been a continuous increase in engagement of the consumer in social media and blogs, which makes them a key marketing channel for the businesses. Businesses currently have to admit the necessity of joining online marketing as a measure to improve their brand awareness and communicate with the target customers. Affiliate marketing functions based on marketing channel and brings less risk comparing with other online marketing channels. As an almost novel phenomenon, online marketing has experiences a continuous growth over the recent years. The present article examines affiliate marketing network, which is an online marketing channel that gives businesses the opportunity to achieve more visibility with relatively low costs. Many popular media including blogs, voucher code sites, and price comparison sites cooperate using affiliate marketing model. Different online marketing challenges are first introduced to give a theoretical background of online marketing in general. Afterward, more detailed analyses of affiliate marketing, as an online marketing channel and the main area of focus of the study is given. Affiliate marketing and all its features from theoretical aspect are discussed in this article, also commissioner and marketing plans in practice, the requirements, and the environment are discussed.

Keywords: Affiliate marketing network, affiliate program, online advertising,

1. INTRODUCTION

Along with fast growth of the Internet, many held that expansion of the Internet would result in disintermediation, which means elimination of intermediaries in the value chain between manufacturers and end costumers. Jacobs argued that following development of new communication ways in the new economy, distribution channel was reduces from two/three intermediaries to zero or one intermediary, which means more cost savings and customization.

At any rate, beside the fact that the Internet actually can decrease the number of intermediaries in some fields, companies still have the chance to use their expertise in adjusting their activities so that they could remain competitive at the market.

Companies were very fast to spot the marketing capacities of the Internet after introduction of the World Wide Web . By using advantages of the new technologies, companies can improve their competitive features, while the technologies bring the threat of new start-up companies that in some cases are completely virtual. Companies need to adjust their business with the trends of changes or otherwise face losing their markets .

Online marketing is featured with many divergent marketing channels that will be discussed in part 2 in short. The article is mainly focused on affiliate marketing as one of safer online marketing channels. Affiliate marketing and all its features from theoretical aspect are discussed in part 3, followed by part 4 where commissioner and marketing plans in practice, the requirements, and the environment are discussed.

2. ONLINE ADVERTISING

Online advertising is far more than placing a website on the Internet even if it is designed for promotional purposes . On the contrary, advertising is a mean to attract the potential customers to that website –i.e. traffic acquisition.

From the content providers' viewpoint, online advertising is an important source of income. According to the result-oriented compensation methods, affiliate marketing is a specific way of selling advertising space. To find out when it is beneficial, advantages and disadvantages of other options available for content providers must be examined.

are several definitions for the term *Internet advertising*. There has been an ongoing debate on whether to consider the Interne as another form of traditional advertising or its nature is more of direct marketing.

Companies, regardless of their industry, can be divided based on their role in business transactions. Zeff highlighted three parties that participate in online advertising:

- Sellers
- Buyers
- Infrastructure (Zeff, 1999; 18-21)

2.1 Sellers

Sellers are in charge of providing content that produce websites with advertising space to sell. Many of the modern sellers are media companies. Since inventory of the content providers is not easy to sell, they were quick in developing ad networks. The networks provide experienced sales force and the advertisers find them easier to deal with.

2.2 Buyers

Companies with products or services to promote are the buyers. Usually, they are represented by agencies that handle their advertising campaigns.

2.3 Infrastructure

Companies that run services of evaluating and auditing campaigns, targeting, and personalization companies, ad management software developers, and suppliers of technological infrastructure.

3. AFFILIATE MARKETING

Affiliate marketing is an economic way comparing with other online advertising forms as it cuts the administrative costs of buying advertising. When the program is well managed, it enables advertising to a great extent of website which is not acquirable otherwise.

The process of affiliate marketing is explained at first and then the parties engaged in the process and their role are discussed. A detailed analyses of the parties are provided along with an introduction to advertising methods and the commission payment models. Risks of affiliate marketing and the performing affiliate programs are also introduced.

The concept of affiliate marketing is designed based on reward in return for performance and here performance is a sale, registration for a web services, signing up for e-mail marketing, request for contact or any other measurable action. (Fig. 1)

Fig 1: The Basic Process of Affiliate Marketing

4. TERMINOLOGY

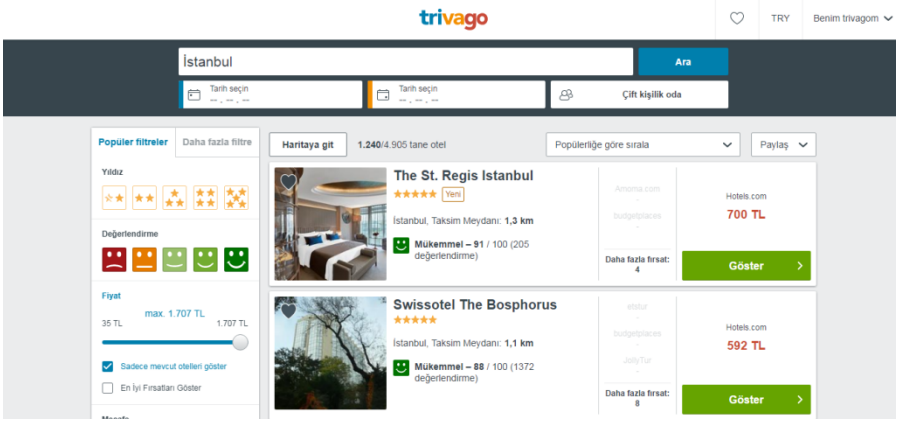
4.1 Merchant

The affiliate marketing merchant refers to a website looking for selling goods or services via online advertising. This is evident in online commerce from the web's largest sellers (e.g. Amazon) to mom-and-pop specialty sites.

4.2 Affiliate

The affiliate or publisher refers to a website that offers links to its visitors. For instance, along with posting a book to a blog or discussion forum, an affiliate may offer a link to Amazon to facilitate readers' purchase. In the same way, a travel guide site can provide link to a page at Trivago offering hotels in the area along with introducing a vacation destination. In the best scenario, the affiliate's links add to usefulness of the content along with paying the publisher. (Fig. 2)

Fig 2: The Sample of Affiliate Company



4.3 Network

The network provides connection between merchants and affiliates. Many merchants trust on networks for tracking, administration, and accounting purposes. This is needed to log which users clicked on what link and made what purchase, and consequently, develop a secure website for affiliates to obtain links and examine the results. The networks also offer effective consolidated payments to several affiliates on monthly bases.

However, many merchants have more tendency to the benefits of specialization. Networks introduce rules on permissible affiliate practices. By becoming a member of a network, the merchant can breach many of these rules or introduce other requirements of its own.

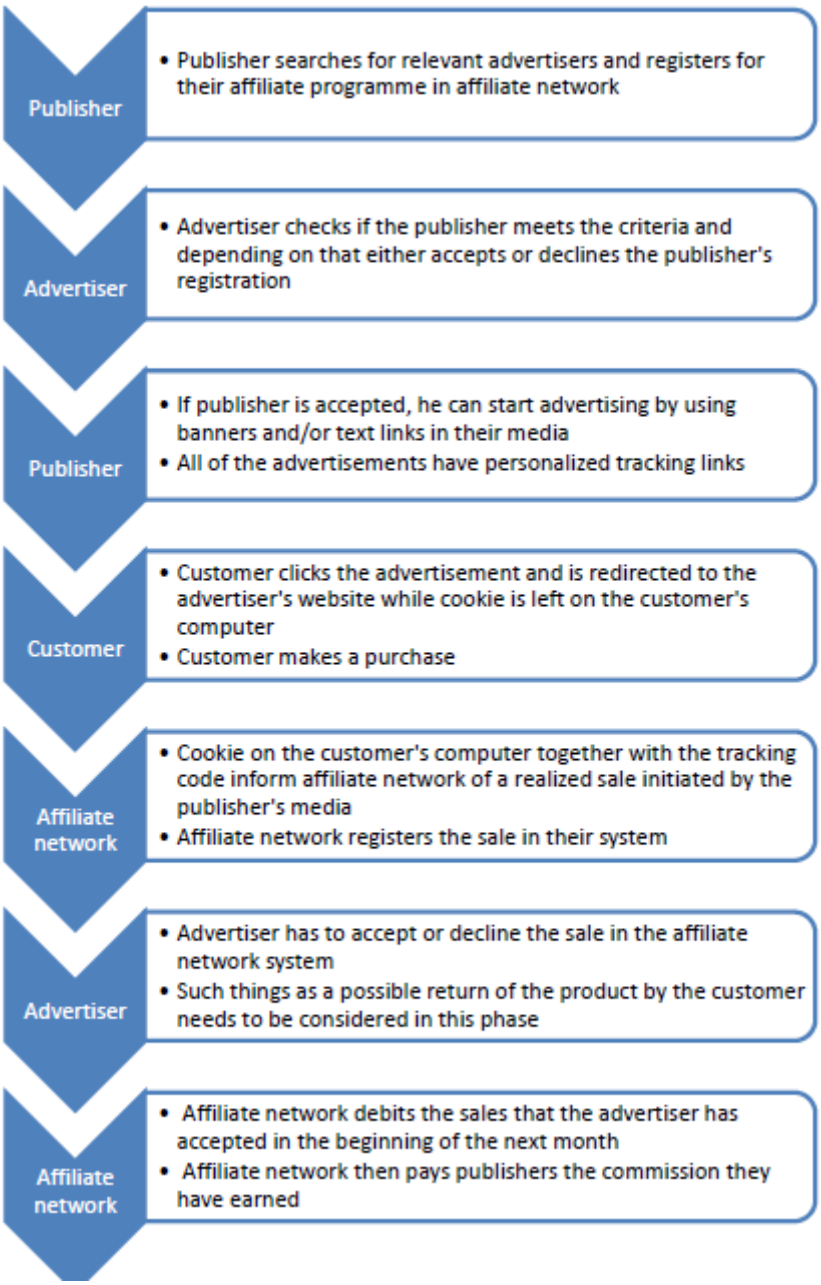
affiliate program: The rules of an affiliate program (e.g. how much affiliates will be paid, permitted behaviors, participation criteria for the affiliates) are set by the affiliate program managers.

Affiliate marketing as transaction oriented, which caused no liability for joint future success or exclusivity restrictions. Affiliate programs need little effort and investment, which makes it also easy to terminate.

Currently, the number of companies that have initiated their own affiliate programs to reach new customers is growing. Affiliate networks have emerged in the form of intermediaries between content providers and merchants and provide solutions to handle affiliate programs.

Although, it is possible for the companies to select a few content providers and provide an opportunity for participation in an affiliate program, in practice it is the content providers who make decision about the ad content suitable for their websites. Given the increasing number of affiliate programs found on the internet, content providers can choose from a large diversity.

Figure 3:Role and Responsibilities in Affiliate Marketing



Affiliate networks provide affiliate programs to websites so that instead of running an affiliate program, the seller can use the affiliate network. The network provides services such as tracking publisher's sales, pay out the commissions and provide consultancy service to advertisers and publishers should a problem or question is raised. It also charges advertisers based on the results. Moreover, affiliate network determines minimum transaction revenue for the advertiser.

By clicking on an affiliate link or banner, the customer allows cookies being saved in their computer. The information of the cookie is sent to the affiliate network along with details such as the timestamp and technical details of the customer's browser and computer. The cookies stay in the customer's computer for 30 days, in general. If a purchase is made during these 30 days, the publisher will receive a commission for the sale. Therefore, it is not necessary to make a purchaser immediately of the clink. After completion of the purchase the final "thanks for the purchase" page send the order number, time stamp, and publisher identification number so that the affiliate network can track the sale.

5. TYPES OF AFFILIATE MARKETING

By definition, affiliate marketing is defined as an agreement achieved between a merchant and content providers for promoting merchant's products/services at their websites. Based on the agreement, the content provider is paid when the visitors from their website perform a specific action.

Therefore, in an affiliate marketing arrangement, in addition to delivering the advertisement, the content provider needs to persuade users to become customers. Thus, affiliate marketing is also called performance marketing.

There are two categories of affiliate marketing; one-to-one affiliate marketing and one-to-many affiliate marketing. In the former, the merchant signs a contract with a selected affiliate, in which the terms are negotiable without influencing contracts with other affiliates. This form of affiliate arrangement is usually adopted by major players at the market who has the capacity to lure many potential customers to the merchant's website.

By using one-to-many affiliate programs, the merchant offers identical condition for all affiliates and they are free to join the program. This forms of arrangement is used when the merchant cooperates with several affiliates and negotiating contract terms with all providers is too costly. Also affiliate marketing is specifically beneficial to small websites as without such arrangements, they do not have the opportunity of selling to major advertisers. Thus, these websites can enjoy great advantages from affiliate networks.

Companies can either develop a private affiliate program at their website or join third-party networks and take advantage of their technological solutions . According Marketing Terms (2007), affiliate network is "a value-added intermediary providing services, including aggregation, for affiliate merchants and affiliates."

6. STANDARDIZATION IN AFFILIATE MARKETING

Affiliate marketing enables notable performance incentives with wider efficiencies of online advertising.

The rule of affiliate marketing payments are aimed at protecting advertisers against wasted expenses. In this regard, the payment structures and the consequential risks of other online ads systems are notable. For instance, a reasonable concern of the advertiser might be that few users would click on its buying display ads; this might be due to irrelevance of the ads to the users' interests or poor placement of the ads. The advertiser does not have control over all these factors: standard contracts enable the advertising networks determine the sites that will display a specific ad.

For example one of the top affiliate networks, LinkShare, promised the advertisers that they are required to make payment to an affiliate only when a sale or other specific action is performed by the user. LinkShare believes that this method is highly efficient. Following the same path, Commission Junction affiliate network states that only when a determined action is taken by the user, the advertiser will be charged.

It is true that practitioners believe that affiliate marketing is structurally free of fraud, in practice however, there are significant risks to be concerned about that will be explain in next section.

7. ADVANTAGE OF AFFILIATE MARKETING

Although, there are theoretical definition for the advantages of affiliate marketing for merchants, there is no proper description of the benefits for content providers. Content providers face the risk pertinent with using affiliate marketing. Still, there is no clear answer for the question why and when they decide to take the risk and not using other types of online advertising.

To make sure that affiliate programs achieve the expected results, companies and content providers need to establish a win-win relationship that may result in generating adequate sale at the websites of the merchants and sufficient earning for the content providers. The determining an attractive commission rate and commission model are essential to lure content providers to an affiliate program.

Advantages of affiliate marketing can be examined from two different perspectives; i) merchants who use affiliate marketing as a part of their marketing strategy, and ii) the content providers – affiliates, who display products/services of merchants on their websites Affiliate marketing provides opportunities of revenues from the website for the content providers; in addition, they can enjoy the opportunity to cooperate with a large number of merchants, which is not possible otherwise.

Participation in the affiliate program is most beneficial for the content providers when there is a strong connection between the website products/services promoted through the program. The connection is rooted in the product type, perception of brand, customer loyalty, and the like.

In addition to the information noted above, affiliate marketing provides a capacity to monitor activity of the customers after they click on the ads. This is done using a cookie that is stored on the potential customer's computer . Merchants are able to track the customers' activity for several days and weeks and if a sale is performed in this period, the content provider receives the commission.

8. DISADVANTAGE OF AFFILIATE MARKETING

There are some disadvantages cases by using affiliate marketing network that we discuss in this part of paper.

8.1 Fraud in Affiliate Marketing

Our representation of affiliate litigation (Edelman, 2012) highlights many disputes that are important enough to lead to a legal action. Since the methods under consideration are featured with the elements of common law fraud and they are considered as fraud in both civil and criminal litigation, they are known as affiliate fraud. Commission, in many affiliate marketing schedules, is payable only when a user performs a purchase. Thereby, to charge a merchant, the affiliate must make the merchant's record show that the affiliate has actually caused extra sale. In fact, an affiliate might penetrate server of the merchant to change the records directly. However, an attacker with privileged access to merchants' servers does not need to stop at affiliate fraud. In practice, affiliate fraud is mainly about focusing on the schemes to spot the users who had already made a purchase decision.

8.2 Fraud Categories

1. Adware: The user visits web page of a merchant using a computer that runs a specific advertising software; the software monitors the user's activity and takes the user to the affiliate's marketing link. If then the user performs the purchase, the affiliate will be paid as the assumed cause of the purchase.

2. Cookie-stuffing: the user visits a web page of which a section can claim to refer the user to a merchant. When the user makes a purchase from the merchant within the preset time period (usually seven to 30 days) the affiliate will receive a commission. In another scenario, the affiliate may design its cookie-stuffing as a component of other web pages.

3. Typosquatting: Affiliates are in charge of registering domain names that are misspelling of merchant's domain names. If a user misspells a merchant's domain name in a predictable way for the affiliate, the user is redirected to the affiliate's site. Then the user is taken to an affiliate link and then to the merchant. Afterward, the affiliate will received the commission when the user makes a payment.

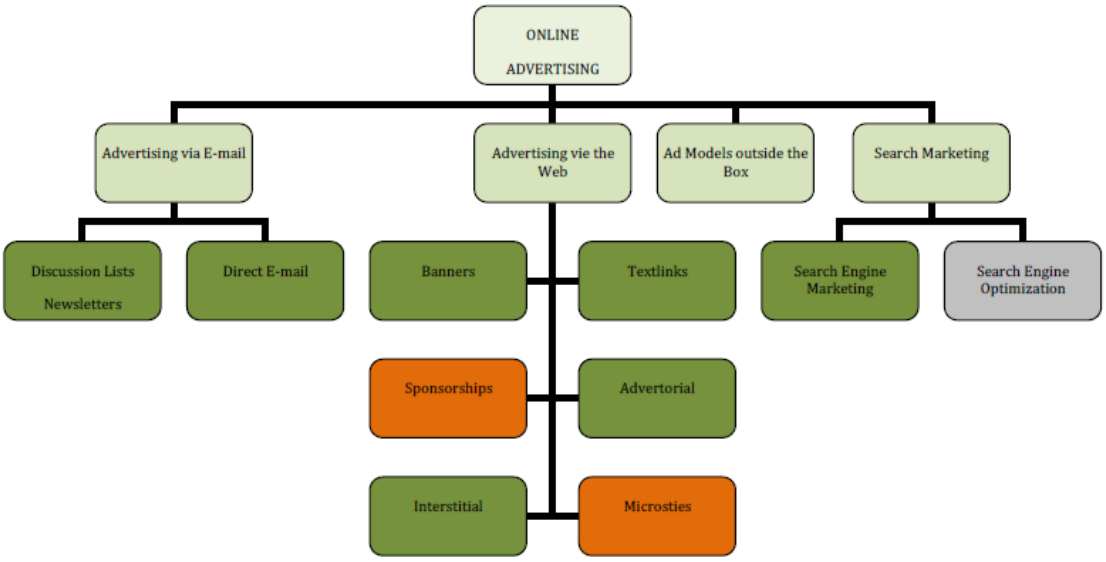
4. Loyalty software: by installing a loyalty software on a user's computer, affiliates ensure that the user is well informed about probable rebates, points, or other advantages of purchasing from specific merchants. The software leads the user via an affiliate's link when the user want to go directly to a merchant's website.

9. RELATIONSHIP BETWEEN ONLINE ADVERTISING AND AFFILATE MARKETING

The figure 4 illustrates the different forms of online advertising available to the merchants to acquire customers. The highlighted fields enjoy advantage of a network of content providers, while the dark green fields indicate the areas that

affiliate marketing is an option. As noted below, it is possible to use affiliate marketing in e-mail advertising through adding links with affiliate reference to the text of e-mail or newsletter. In the case of web-based advertising, affiliate links are connected to banners, textlinks, advertorials and interstitials. Sponsorship and microsites are not recommended as the visitors stay in the website of the content providers so that payment models cannot be used. Affiliate marketing can be used in search engine marketing; a topic that we will return to it later.

Figure 4: Overview of online advertising and its relationship to affiliate marketing



the some cases of using affiliate marketing are storefronts, so that content providers can market products/services of other companies at their own branded website using technological solutions offered to them by the advertiser. There is no need for the website visitors to leave the content providers' website to complete the purchase . Storefronts are usually called "white lables."

10. COMPENSATION MODELS IN AFFILATE MARKETING

As highlighted by Fiore and in affiliate marketing in particular, advertisers can choose one of the three compensation models explained above to determine commissions payable to the content providers.

- Pay per sale (PPS): commission is paid for each sale by the visitor from the affiliate website.
- Pay per lead (PPL): commission is paid to the affiliate when one of its visitors performs a specific action (e.g. subscribing a newsletter, filling out a questionnaire). Note: another term of PPL is pay per action (PPA)
- Pay per click (PPC): commission is paid to the content provider for each click at the link to the advertiser's website. (Fiore, 2001; 125)

10.1 CPS (Cost-Per-Sale)

The term CPS, also known as PPS (Pay Per Sale), refers to a low-risk, high-profit, revenue-sharing model utilized by marketers to attract new customers to their products or services. CPS pays a determined commission to the affiliate marketer who refers a lead that results in a purchase. CPS model is more preferred by the marketers as they are only required to pay a commission when they sell to a customer. In fact, it is a free marketing and advertising scheme where the affiliate produces the lead without receiving a upfront cost.

10.2 CPA (Cost-Per-Action)

Affiliate marketing is an easy way to address banner blindness, since content providers have the deepest insight into their customers and know what advertisings is more effective . Additionally, the content provider has all the reason to avoid banner blindness, since affiliate marketing campaigns are paid per action (PPA).

Under CPA setting, the affiliates receive a fee when a specific action is perform by the referral or lead. Among such actions are clicks, impressions, form submits, sign-ups, registrations, or opt-ins. Because CPS are not necessarily featured with a direct sale and more risk taking, the payout percentages are smaller than those in CPS setting.

Three basic plans introduced above might be mixed with other types. The most common combinations are two-tier and residual programs. The former functions based on multilevel marketing and the affiliates are paid in return of the actions taken by the newly recruited affiliates. In the latter, a commission is paid for all the purchases performed by previously referred visitors during a specific time period.

The commission paid by the merchants might be based on a flat fee (i.e. the affiliate receive a specific amount of money for every sale) or based on a percentage of sale value.

10.3 Business models of Content Providers

To have a better insight into the ways that content providers place themselves in the online value chain and earn money from their websites, one needs to learn about their sources of revenue. Evaluating the condition in which different forms of online advertising are efficient or not needs insight into the available business models.

There are seven sources of income for online content providers:

1. Online advertising
2. Subscription fee for online content
3. Online ordering of print publication
4. Syndication/revenue sharing with online services
5. Per-unit charges for online content
6. Online sale of non-content merchandise and services
7. Affiliate marketing

However, the mentioned article only focused on magazine and newspaper companies so that some of the mentioned sources are not available in general. Strauss et al. proposed a better model in which business models for cybermediaries are categorized as follows:

- Brokerage model
- Agent model
- Online retailing

Brokerage model

In this model, the cybermediary functions as a supplier of negotiations platform between buyers and sellers. The cybermediary might charge the buyer, the seller, or the both. These companies function on the basis of business to business (B2B), business to customer (B2C) or customer to customer (C2C) models. In general, they offer several value added services for both buyers and sellers. There are two common models in brokerage model:

- Online exchange
- Online auction

Agent model

As explained by Strauss, online agents act as representative and in interest of the buyer or seller. Following structures are notable in the agent model:

- Models representing sellers
    - Selling agent
    - Metamediaries
    - Virtual malls
- Models representing buyers
    - Purchasing agents
    - Reverse auction
    - Buyer cooperative

Selling Agents

Strauss illustrated selling agents as cypermediaries that act as representative of one or several companies and offer them services to promote their sale. In this arrangement, a commission model is used within an affiliate program. The agent receives a commission when it refers a user that will perform a purchase or other specification actions. While an agent in traditional marketing represents one producers, in online environment, an agent can represent an industry and recommend rival products.

This model is generally used in travel industry and the agents usually represent airlines, hotels, car rental companies and the like. From the customer's viewpoint, the agent provide valuable services as, through their services, they can find better deals and enjoy more convenient buying process.

Another notable model is catalog aggregators, in which information about products is collected in online catalogs and customers can choose among several products. The customer is also provided with real time prices and products information. This model is more general in B2B Solutions. At the most advanced level, catalog aggregators might provide direct link to venders' ERP (enterprise resource planning) systems.

Metamediaries

As explained by Strauss, the term "metamediary" refers to a website that creates connection between stakeholders and an important event like purchase of asset or marriage. Metamediaries shorten the customer's search for information, guarantee quality of suppliers, and supply pertinent information. They also act as facilitators of transactions. On the other hand, merchants enjoy the traffic to their website provided by metamediaries and co-branding. Metamediaries work based on commission models. Strauss et al. noted that what matters between the metamediary and customers is the trust and to preserve their unbiased image, some of them do not even utilize web ads.

Virtual malls

Virtual malls function like brick-and-mortar shopping malls. Like the metamediaries model, stores enjoy benefits of virtual mall traffic and co-branding. Customers, on the other hand, enjoy the benefits of having one shopping cart in all stores and do purchase while benefiting loyalty programs, gift registries, search facilities and shopping suggestions, recommendations, and product reviews.

Purchasing agents

A purchasing agent collects orders of individual customers who can remain anonymous, and perform large volume order for better price. Practically, the customer is in the position to determine the price of their preference.

Reserve auction

Reserve auction works alike purchasing agents in all aspects with the exception that the business keeps its separate identity. The customer can determine their preferred price and the vendors bid to reach the price or offer a lowest price to grab attention of the buyer.

Buyer cooperative

An alternation of purchasing agents is buyer cooperative, where, customers are grouped in order to achieve lower prices; however, in buyer cooperative, a relationship exists between the amount of buyers and the price. Buyers know the exact number of buyers they need to collect to reduce the price. In such a case, buyers are in charge of marketing and finding more buyers.

Online retailing

Online retailers means online stores, so that they buy goods from the suppliers and sell them to the customers.

11. CONCLUSION

It has been for decades or one as some say centuries that businesses have complained about effectiveness of advertisement activities. While the costs of advertisement needs to be paid beforehand and the expected benefits, if any, will come later, the advertisers are always at the risk of low performance or no-performance in their campaigns. In the face of this drawback, affiliate marketing brings in new changes; it introduces performance-based approach to online marketing and the advertisers are needed to pay only when a sale takes place. Thanks to the reliable online tacking that highlights the sales pertinent to affiliates, advertisers enjoy a notable decrease in their risk.

In this article affiliate marketing and all its features from theoretical aspect are discussed in this article, also commissioner and marketing plans in practice, the requirements, and the environment are discussed.

REFERENCES

Barbora benediktová , lukas nevosad (2008), Master Thesis , lulea teknik üniversitesi, ISSN: 1653-0187 , mart 2008 PP:30-55

Zeff, R. L., & Aronson. (1999). Advertising on the Internet. 2nd ed. New York: Wiley Computer Publishing , ISBN 0-471-34404PP 230-301

Chaffey, D. (2003). Internet Marketing: Strategy, Implementation and Practice. 2nd ed. New York: Financial Times.

Chatterjee, D.D., Pacini, C., and Sambamurthy, V.(2002) "The Shareholder-Wealth and Trading-Volume Effects of Information Technology Infrastructure Investments," Journal of Management Information Systems (19:2), Fall 2002, pp 7-42.

Edelman, Benjamin (2013),The design of online advertising markets. In Nir Vulkan, Alvin Roth, and Zvika Neeman, editors, The Handbook of Market Design. Oxford University Press, 2013.

Stefan Schwarz (2011) ,Affiliate Marketing : A case study of Ticket Express GesmBHoticket.com , Master Thises , Modul Vienna University

Ben Edelmen ,Weley Brandly (2014) ,Risk Information and Incentives in Online Affiliate Marketing ,Working paper ,Harward Business School , haziran 2014

United States of America v. Hogan (2010), U.S. District Court for the Northern District of California, CR 10-0495-JF. Pp:6-7

Baker, George P. and Thomas Hubbard (2003), "Make Versus Buy in Trucking: Asset Ownership, Job Design, and Information," American Economic Review , 93 (3), 551–72.

Blake, Thomas, Chris Nosko, and Steven Tadelis (2013), "Consumer Heterogeneity and Paid Search Effectiveness: A Large Scale Field Experiment," mimeo, eBay Research Labs.

Lafontaine, Francine and Margaret Slade (2007), "Vertical Integration and Firm Boundaries: The Evidence," Journal of Economic Literature , 45 (September), 629–685.

Libai, Barak, Eyal Biyalogorsky, and Eitan Gerstner (2003), "Setting Referral Fees in Affiliate Marketing," Journal of Service Research , Vol 5 ,No 4 (May), 303–315.

Moore, Tyler and Benjamin Edelman (2010), "Measuring the Perpetrators and Funders of Typosquatting," in Financial Cryptography Lecture Notes in Computer Science , Vol. 6052, RaduSion, ed. Berlin: Springer, 175–91.

Zhu, Yi and Kenneth Wilbur (2011), "Hybrid Advertising Auc-tions,"Marketing Science, 30 (2), PP .249–73.

Adamson, A.P. (2011). Brand Digital. Bratislava: Eastone Books. ISBN 978-80-8109- 192-6

Jurišová V. et al. (2012). Affiliate marketing. In Conference proceedings – International Scientific Conference New Trends in Marketing – Social Responsibility. Trnava: FMK UCM. ISBN 978-80-8105- 438-9

Matúšová, J. (2013). Budovanie a komunikácia značky: Značka v PR a reklame. Trnava: FMK UCM. ISBN 978-80-8105-440-2

Murár, P. (2010). Sociálne médiá ako prostredie marketingovej komunikácie firmy. In Sborník příspěvků: III. mezinárodní vědecké konference doktorandů a mladých vědeckých pracovníků. Karviná: Slezká univerzita v Opavě, Obchodně podnikatelská fakulta v Karviné. ISBN 978-80-7248-620-5

Pizano, V. – Kusá, A. (2012). Typológia slovenských internetových používateľov podľa vnímania nových foriem marketingovej komunikácie. Trnava: FMK UCM. ISBN 978-80-8105-425-9

D.M. Scott. (2010). Nové pravidlá marketingu a PR. Bratislava: Eastone Books. ISBN 978-80-8109-149-0

Dennis L. Duffy (2005), 'Affiliate Marketing and Its Impact on e-commerce', Journal of Consumer Marketing, vol. 22, no 3, pp.161 – 163.

Venu Gopal K., Saumendra das., Dr D.vishnu Murthy., (2012)., "EMAIL MARKETING: A PARADIGM SHIFT TO MARKETING" International Journal of Advanced Research in management and Social Sciences , Volume1, no1, July 2012, ISSN: 2278 -6236, pp.101-11

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

ENHANCING BREAST CANCER DETECTION USING DATA MINING CLASSIFICATION TECHNIQUES

DOI: 10.17261/Pressacademia.2017.605
PAP-WCTIE-V.5-2017(43)-p.310-316

Florije Ismaili[1], Luzana Shabani[2], Bujar Raufi[3], Jaumin Ajdari[4], Xhemal Zenuni[5]
[1]South East European University. f.ismaili@seeu.edu.mk
[2]State University of Tetovo. Luzana.shabani@unite.edu.mk
[3]South East European University. b.raufi@seeu.edu.mk
[4]South East European University. j.ajdari@seeu.edu.mk
[5]South East European University. xh.zenuni@seeu.edu.mk

ABSTRACT

Cancer is one of the crucial causes of death for both men and women. All over the world, breast cancer is one of the leading cause of cancer deaths in women. The most effective way to reduce cancer death is to detect it earlier but the detection of cancer in early stages is not an easy process. As result, many researches are focused on developing different systems for breast cancer detection. In this paper we have discussed various data mining approaches that have been utilized for breast cancer diagnosis and prognosis. We have proposed a breast cancer prediction framework consisting of four main modules: Data Collection, Data Preprocessing, Feature Selection, and Classification. Evaluation results are provided as well. The goal is to find the best combination for feature extraction algorithm and classification algorithm, which will improve the accuracy of mammograms classification process.

Keywords: Breast cancer, mammograms classification, data mining

1. INTRODUCTION

Breast Cancer is among the leading causes of cancer death in women. Although mammography is currently the most effective tool for early detection of breast cancer, it has some restrictions. Radiologists visually search mammograms for specific abnormalities, but detection of suspicious abnormalities is a repetitive and wearing task, thus an abnormality may be unnoticed. Consequently, research for breast cancer detection is focused in finding computer aided methods, developed to aid radiologists in detecting mammographic lesions that may indicate the presence of breast cancer [1-5].

In this paper we propose a breast cancer prediction framework based on data mining classification techniques. The proposed framework consists of four major steps of determining the breast cancer: collection of mammogram images, image preprocessing, classification and result evaluation. To evaluate the accuracy of the proposed model, experimental results are provided as well.

The rest of the paper is organized as follows: section two provides an overview of the related work done for breast cancer identification using data mining techniques. In section three the proposed framework is presented followed by classification methodology in section four. Section five elaborates the experimental results while section six concludes the paper.

2. LITERATURE REVIEW

A literature survey showed that there have been several studies on the breast cancer detection using data mining techniques.

Salama, Gouda I., M. B. Abdelhalim, and Magdy Abd-elghany Zeid [6] applied various classification algorithms on three different breast cancer databases: Wisconsin Breast Cancer (WBC), Wisconsin Diagnosis Breast Cancer (WDBC) and Wisconsin Prognosis Breast Cancer (WPBC). Experiments are performed using 10-fold cross validation method combined

with tree, Multi-LayerPerception, Naive Bayes, Sequential Minimal Optimization, and Instance-Based for K-Nearest neighbor. According to their results, classification using fusion of MLP and J48 shows better results than other classification approaches.

Mittal, Dishant, Dev Gaurav, and Sanjiban Sekhar Roy [7] proposed a hybrid method of breast cancer diagnosis which gave a significant accuracy over training set and testing set. The proposed hybrid method combines unsupervised self-organizing maps (SOM) with a supervised classifier called stochastic gradient descent (SGD). The experimental results are conducted by comparing their results with three supervised machine learning techniques: decision tree (DTs), random forests (RF) and support vector machine (SVM).

M. Vasantha et al., are concentrated in classifying the mammogram images into three categories (normal image, benign image and malignant image) decision[8]. Halawani et al. have applied different clustering algorithms in order to detect breast cancer. Experiments were conducted using Digital mammograms in the University of Erlangen-Nuremberg between 2003 and 2006 [9].

A survey done by [10] have analyzed a significant number of research done in the field of breast cancer detection. Their focus was on analyzing the different data mining techniques applied in breast cancer classification along with their advantages and disadvantages. Particularly, this survey discusses about use of the classification algorithms ID3 and C4.5 in breast cancer analysis.

A review on different machine learning applications in cancer prediction and prognosis is presented by Kourou et al. [11]. In the presented review of around 70 approaches, they came to conclusion that last years the research is focused on the development of predictive models using supervised ML methods and classification algorithms where the integration of multidimensional heterogeneous data is combined with the application of different techniques for feature selection.

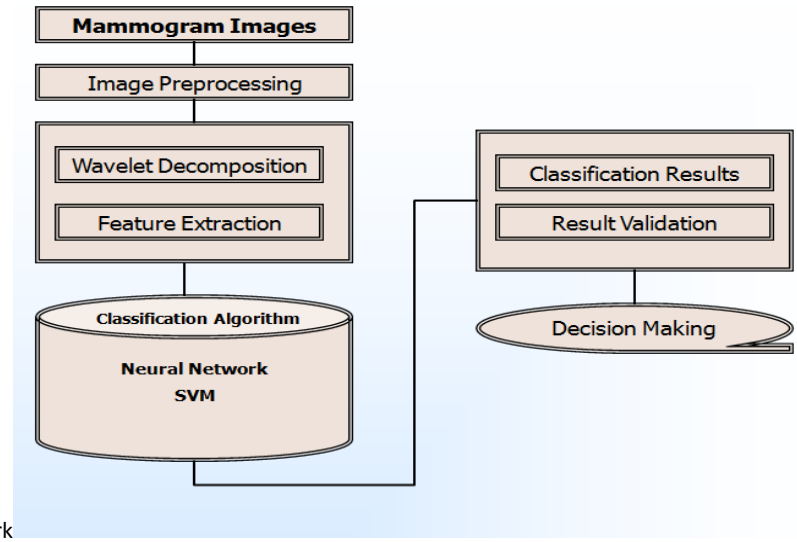
In the review done by Cruz et al. the performance of different machine learning that are being applied to cancer prediction and prognosis have been explained, compared and assessed [12]. They identified a number of trends with respect to the types of machine learning methods being used, the types of training data being integrated, the kinds of endpoint predictions being made, the types of cancers being studied and the overall performance of these methods in predicting cancer susceptibility or outcomes.

From related work it is obvious that even a huge work is done in the field of breast cancer detection, researches are still focused on enhancing the performance of cancer detection by using data mining techniques combined with the application of different techniques for feature selection, which proves the validity of the work presented in this paper.

3. PROPOSED FRAMEWORK

The aim of this research is to develop a tool for the prediction of breast cancer at its initial stage. Therefore we propose a “Breast Cancer Prediction Framework” which will contribute in decreasing mortality rate due to breast cancer. The overall architecture of the proposed framework is given in Figure 1.

Figure 1. Breast Cancer Prediction



Framework

The following are the key steps of the proposed framework:

- A corpus of data consisting of mammogram images is created.
- Mammograms are preprocessed in order to create feature vectors appropriate for classification. Feature vectors are created in two manners:
    - using global histogram equalization to obtain a uniform histogram for the output image where images were subjected to a decomposition process by wavelet transform and
    - feature extraction from image properties like radiologist's "truth"-markings on the locations of any abnormalities that may be present, character of background tissue, class of abnormality present, (x, y) image-coordinates of centre of abnormality and approximate radius (in pixels) of a circle enclosing the abnormality.
- Classifiers are trained using train data then classification algorithms are applied one by one to find out which one is producing better result in terms of accuracy for the given data set.
- Decision for breast cancer presence is made according to classification results evaluation.

4. DATA AND METHODOLOGY

This section, discusses the methodology which has been used for the proposed work. Breast Cancer Prediction Framework consists of four main modules namely: Data Collection, Data Pre-processing, Feature Selection, and Classification.

Database has been taken from Mammographic Image Analysis Society (MIAS), in order to perform experiments and evaluate the obtained results [13]. The details of all stages are discussed below.

4.1. Data Collection

First step is the collection of data. For experimental purposes we have used the MIAS digital mammography database, which consists of total 330 images. The mammograms are collected by United Kingdom National Breast Screening Programme and all mammograms follow the same criteria: only the medio-lateral oblique view is available, all films taken have been digitalized to 50 micron pixel edge with a Joyce-Loebl scanning microdensitometer.

The mammograms database includes 123 images with both benign and malignant forms of abnormalities and 207 normal mammograms. Additional information like radiologist's "truth"-markings on the locations of any abnormalities that may be present, character of background tissue (Fatty, Fatty-glandular, Denseglandular), class of abnormality present (Calcification, Welldefined/circumscribed masses, Spiculated masses, ill-defined mass, Architectural distortion, Asymmetry, Normal), (x, y) image-coordinates of centre of abnormality and approximate radius (in pixels) of a circle enclosing the abnormality are provided as well.

4.2. Data Preprocessing

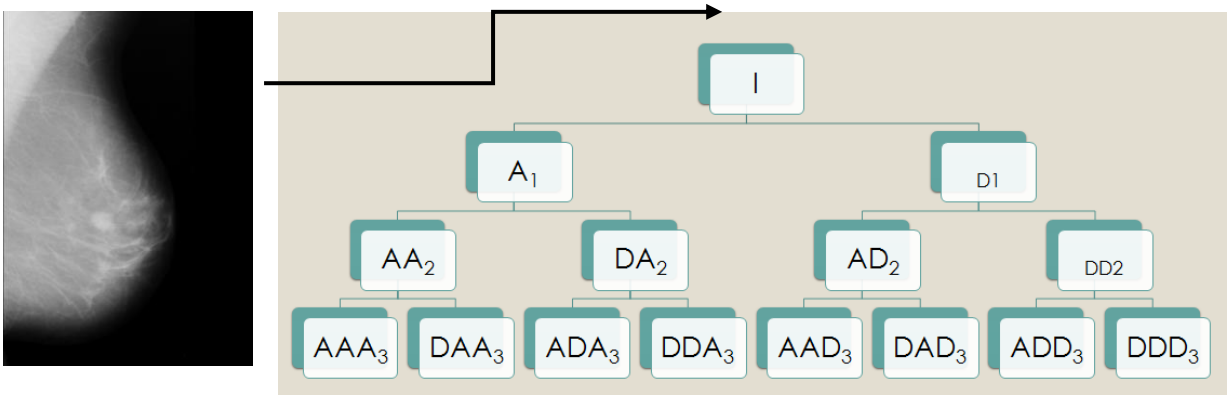
To improve the quality of the images, we use the global histogram equalization (GHE) in the pre-processing stage. The main objective of GHE is to obtain a uniform histogram for the output image. To perform histogram equalization, the running sum of the histogram values is discovered, normalized and then multiplied with maximum gray level value. These values are then mapped on the previous original values using one-to-one correspondence [14]. Basically, this method multiplies the scale factor from the normalized cumulative distribution of the brightness distribution of the original image with the original image to redistribute the intensity.

At this stage, images were subjected to a decomposition process by wavelet transform which involves decomposition of the signal. The decomposition of the multilevel wavelet transform can be expressed as follows [15]:

$$I = A_j + D_j + D_{j-1} + ... + D_2 + D_1 \quad (1)$$

where I represents the $i-th$ image, j represents the level of decomposition, A is approximation and D is detail coefficients. Figure2 demonstrates the decomposition process of the original image.

Figure 2: DWT Decomposition Process of the Original Image



The detail coefficients consist of noise, so for feature extraction, only approximation coefficients are used. Information loss can occur after level four since informative coefficients cannot be detected properly. Thus, to avoid misclassification, each image was decomposed up to four levels, i.e. As a result, by summing all the approximation coefficients at each level, a one-dimensional matrix is obtained as follows:

$$Mat = A_4 + A_3 + A_2 + A_1 \quad (2)$$

where A indicates the approximation coefficients at each level of decomposition and Mat represents the one-dimensional matrix obtained from summation of all the approximation coefficients. The resultant matrix is subjected to wavelet transform to generate the feature vectors. The general form of the wavelet transform can be written as:

$$W_y(a,b) = C(a,b) = \int_{-\infty}^{\infty} y(t)\frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right)dt \quad (3)$$

where $W_y(a,b) = C(a,b)$ is the wavelet coefficient (approximately directly proportional to the amplitude of a specific mode) with the scale a (inversely proportional to the wavelet center frequency) and position b and Ψ is the complex conjugated wavelet function [16]. To generate the feature vectors from one-dimensional matrix (M), the wavelet coefficient is converted into the mode of frequency (i.e., f_m) as follows:

$$f_v = \frac{f_{avg}(\Psi_{f.e})}{a(\Psi_{f.e})\Delta} \quad (4)$$

where $f_{avg}(\Psi_{f.e})$ is the average frequency of the wavelet function, $a(\Psi_{f.e})$ represents the approximation coefficients at all levels of decomposition, Δ indicates the image decomposition period, and f_v indicates the resultant feature vector for an image. To reduce the amount of data produced by the wavelet transform, the discrete wavelet transform (DWT) that uses a certain subset of scales- a , and positions - b is used[16].

4.3. Classification

For classifying mammogram images which are associated with benign, malign and normal classes, we used SVMs and Artificial Neural Networks (ANNs).

Neural Networks are statistical learning models used in machine learning. They are capable of a wide range of classification or pattern recognition problems since they are able to perform a range of statistical (linear, logistic and nonlinear regression) and logical operations or inferences (AND, OR, XOR, NOT, IF-THEN) as part of the classification process. They are trained to generate an output as a combination between the input variables. Multilayer Artificial Neural Network contains several intermediary called hidden layers between its input and output layers [17,18]. The basic structure of the Neural Network is given in Figure 3 below.

Figure 3: Basic structure of the Neural Network

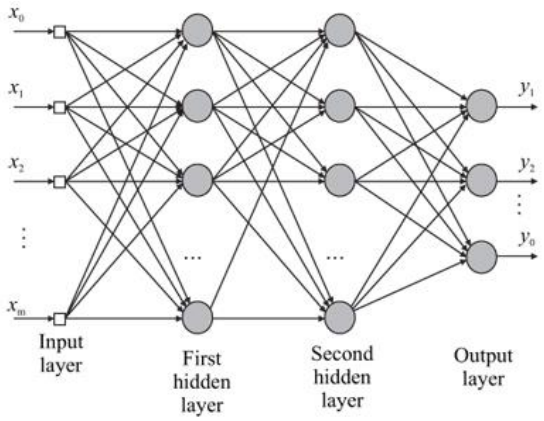
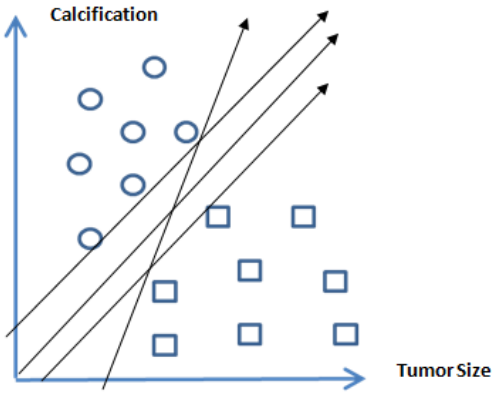


Figure 4. SVM data transformation into a higher dimensional space



Support Vector Machine (SVM) is another Data Mining technique used for data classification. Initially, SVM tries to identify the hyperplane that separates the data points into two classes by mapping the input vector into a feature space of higher dimensionality. Using non-linear kernel which is a mathematical function that transforms the data from a linear feature space to a non-linear feature space, SVM is able to perform nonlinear classification. Like ANNs, SVMs can be used in a wide range of pattern recognition and data classification [17,18]. Figure 4 illustrates how an SVM might classify benign and malignant tumors based on their size and calcification.

5. FINDINGS AND DISCUSSIONS

WEKA version 3.8.1 was utilized as a data mining tool to evaluate the performance and effectiveness of the breast cancer prediction models. Mammogram data are divided in test data and train data, then performance of various classifiers in combination with data pre-processing methods is evaluated using the breast cancer data set. The performance of a chosen classifier is validated based on precision, recall, F – measure and ROC area. Accuracy of the model is measured by the area under the ROC curve. An area of 1 represents a perfect test. Precision or sensitivity is defined by TP / (TP + FN), Recall or specificity is defined by TN / (TN + FP) while F – Measure is the harmonic mean of precision and recall.

True positive (TP) = number of examples predicted positive that are actually positive. False negative (FN) = number of examples predicted negative that are actually positive. False positive (FP) = number of examples predicted positive that are actually negative. True negative (TN) = number of examples predicted negative that are actually negative [17,18].

The analyses have been carried on using two algorithms namely, SMO and Multilayer Perceptron - MP for two types of generated feature vectors: a) feature extraction from properties associated with mammogram images and b) feature vectors created from image wavelet decomposition.

The following table demonstrates the detailed analysis of both classifications algorithms where the correctly and incorrectly classified instances show the percentage of test instances. Kappa statistics value should be high for a good model since Kappa is a chance-corrected measure of agreement between the classifications and the true classes. The Mean absolute error, Root means squared error, Relative absolute error, Root relative squared error are used to assess performance as they are frequently used measure of the differences between classes predicted by a model and the classes actually observed.

Table1: Summary of Mammogram Classification Models

	SVM		Neural Networks	
	Feature Extraction	**Wavelet Transformation**	**Feature Extraction**	**Wavelet Transformation**
Correctly Classified Instances	80.303%	86.9697 %	83.9394 %	95.1515 %
Incorrectly Classified Instances	16.697%	13.0303 %	16.0606 %	4.8485 %
Kappa statistic	0.6317	0.7562	0.7012	0.9093
Mean absolute error	0.266	0.2512	0.1162	0.0529

Root mean squared error	0.3433	0.321	0.3021	0.149
Relative absolute error	74.2453 %	70.1342 %	32.4274 %	14.7782 %
Root Relative squared error	81.2024 %	75.9352 %	71.4685 %	35.2557 %
Total Number of Instances	330	330	330	330

As can be seen in Table1, Neural Network and Support Vector Machine have comparable performances. From experimental results using two classification algorithms combined with two preprocessing techniques is obvious that both of classification algorithms shows better performance when used with image wavelet decomposition preprocessing technique in terms of higher classification accuracy and lower error rate .

Figure 4. Accuracy of SMO Predicting Breast Cancer using Feature Extraction

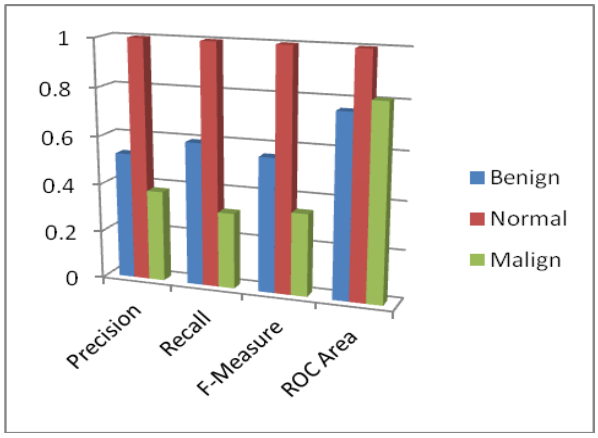


Figure 5. Accuracy of SMO Classifiers for Predicting Breast Cancer using Wavelet Decomposition

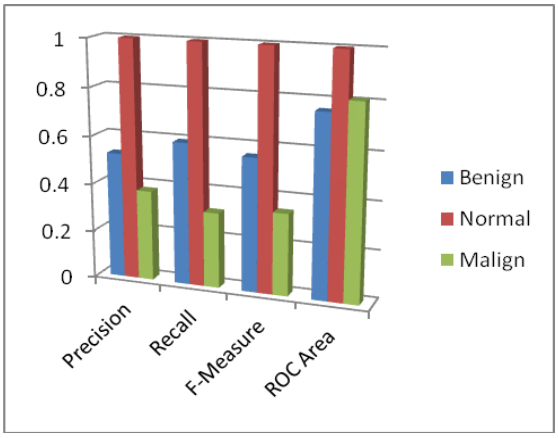


Figure 6. Accuracy of MP Predicting Breast Cancer using Feature Extraction

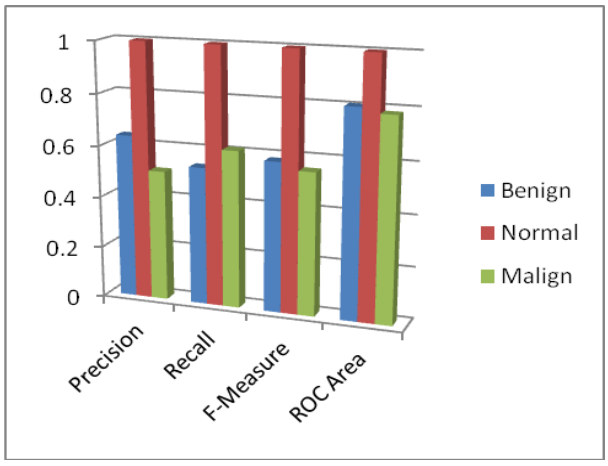


Figure 7. Accuracy of MP Classifiers for Predicting Breast Cancer using Wavelet Decomposition

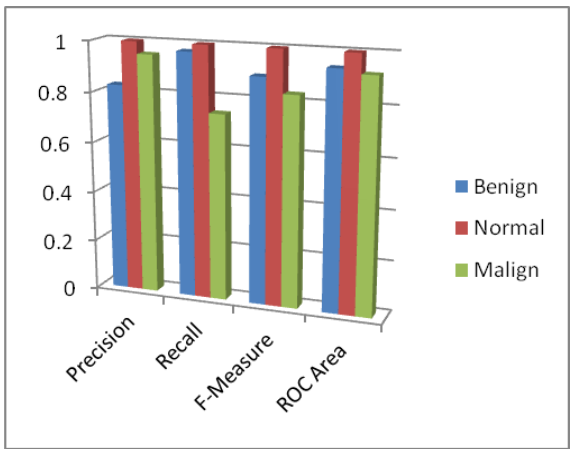


Figure 4, 5, 6, 7 shows the Precision, Recall, F – measure and ROC Area values for SMO and MP classification algorithms. It can be seen that normal mammograms are correctly classified in all cases while mammograms classified as benign and malign have higher precision (0.827, 0952), recall (0.971, 0741), F-measure (0.893, 0.833) and ROC Area (0.94, 0.923) values when Neural Network is used for classification and feature vectors are prepared using image wavelet decomposition technique.

Since cancer detection at very early stage is very sensitive, from experimental results we can conclude that the proposed framework is contributing in detecting predispositions of a patient for having breast cancer. A single user input data (mammogram) is fed into the system and gets preprocessed and classified according to the explained technique. For all

used technique combinations, non cancer data are correctly classified in all cases. If a mammogram is classified as benign or malign than the user should for surely proceed with further analyses. With each new entry getting appended to the model the process becomes intelligent and ensures accurate results.

6. CONCLUSION

This paper has outlined and discussed various data mining approaches and techniques for the problem of breast cancer detection. Although different classification techniques have been developed for cancer classification, there are still many drawbacks in their classification capability.

In order to enhance breast cancer classification, in this paper we proposed a new framework for breast cancer classification by combining mammogram wavelet transformation and neural network. According to results, classification based on locations of any abnormalities that may be present, character of background tissue, class of abnormality present, does not always shows the desired result. Finally, the evaluation and performance analysis of the proposed approach clearly shows that the preliminary results are promising in breast cancer discovery at early stage.

REFERENCES

[1] V. Gaike, R. Mhaske, S. Sonawane, N. Akhter, P. D. Deshmukh, "Clustering of breast cancer tumor using third order GLCM feature", , vol. 00, no. , pp. 318-322, 2015

[2] A. Sahar "Predicting the Serverity of Breast Masses with Data Mining Methods" International Journal of Computer Science Issues, Vol. 10, Issues 2, No 2, March 2013 ISSN (Print):1694-0814| ISSN (Online):1694-0784 www.IJCSI.org.

[3] S. Sondele and I. Saini, "Classification of Mammograms Using Bidimensional Empirical Mode Decomposition Based Features and Artificial Neural Network", International Journal of Bio-Science and Bio-Technology, Vol.5, No.6 (2013), pp.171-180.

[4] Z. K. Senturk and R. "Breast Cancer Diagnosis Via Data Mining: Performance Analysis of Seven Different Algorithms", Computer Science & Engineering: An International Journal (CSEIJ), Vol. 4, No. 1, February 2014.

[5] M. P. Sampat, M. K. Markey, and A. C. Bovik, "Computer-Aided Detection and Diagnosis in Mammography," in Handbook of Image and Video Processing, A. C. Bovik, Ed., 2nd ed, 2005.

[6] Salama Gouda I., M. B. Abdelhalim, and Magdy Abd-elghany Zeid. "Experimental comparison of classifiers for breast cancer diagnosis." Computer Engineering & Systems (ICCES), 2012 Seventh International Conference on. IEEE, 2012.

[7] Mittal Dishant, Dev Gaurav, and Sanjiban Sekhar Roy. "An effective hybridized classifier for breast cancer diagnosis." 2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM). IEEE, 2015.

[8] M. Vasantha, S. Bharathi V and R. Dhamodharan, "Medical image feature, extraction, selection and classification". Intl. J. Engg.Sci. & Technol. 2(6), 2071-2076, 2010.

[9] S M Halawani "A study of digital mammograms by using clustering algorithms" Journal of Scientific & Industrial Research Vol. 71, September 2012, pp. 594-600.

[10] B. Padmapriya and T. Velmurugan, "A survey on breast cancer analysis using data mining techniques," 2014 IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore, 2014, pp. 1-4.

[11] K. Kourou, T. P. Exarchos, K. P. Exarchos , M. V. Karamouzis, D. I. Fotiadis, "Machine learning applications in cancer prognosis and prediction", Computational and Structural Biotechnology Journal 13 (2015) 8–17.

[12] Cruz JA, Wishart DS. "Applications of Machine Learning in Cancer Prediction and Prognosis". Cancer Informatics. 2006;2:59-77.

[13] Mammographic Image Analysis, http://www.mammoimage.org/databases/. Last accessed: March, 2017.

[14] R. Sharmila and R. Uma, "A new approach to image contrast enhancement using weighted threshold histogram equalization with improved switching median filter," International Journal of Advanced Engineering Sciences and Technologies, Vol. 7, 2011, pp. 206-211.

[15] P. Kour, "Immage Processing using Discrete Wavelet Transformation", International Journal of Electronics & Communication (IIJEC), Volume 3, Issue 1, January 2015, ISSN 2321-5984.

[16] J. Turunen, "A wavelet-based method for estimating damping in power systems,"Ph.D. Thesis, Department of Electrical Engineering Power Transmission Systems, Aalto University, 2011.

[17] J. Han, M. Kamber, J. Pei, "Data Mining: Concepts and Techniques", Third Edition (The Morgan Kaufmann Series in Data Management), ISBN-13: 978-9380931913, 2011.

[18] P. N. Tan, M. Steinbach, V. Kumar," Introduction to Data Mining 1st Edition", ISBN-13: 978-0321321367, Pearson Education, 2014.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

ERGONOMIC RISK ASSESSMENT OF MANUAL MATERIAL HANDLING AT AN AUTOMOTIVE MANUFACTURING COMPANY

DOI: 10.17261/Pressacademia.2017.606
PAP-WCTIE-V.5-2017(44)-p.317-324

Baba Md Deros[1], Dian Darina Indah Daruis[2], Athirah Liyana Rosly[1], Ilhami Abd Aziz[1], Nor Syazwani Hishamuddin[1], Nur Hidayah Abd Hamid[1] & Siti Maisarah Roslin[1]
[1]Universiti Kebangsaan Malaysia, hjbaba@ukm.edu.my
[2]Universiti Pertahanan Nasional Malaysia

ABSTRACT

This study investigates ergonomic risks faced by manual material handling (MMH) workers performing repetitive loads carrying, holding, lifting, lowering, pushing and pulling activities. A total of 11 MMH workers was recruited to participate in the study. Methods used to collect data were Modified Nordic Questionnaire (MNQ), Rapid Entire Body Assessment (REBA) and NIOSH Lifting Equation. MNQ result showed the MMH workers' body parts that suffered pain are shoulders (63.6%), upper back (54.5%) and lower back (45.4%). The movements, which recorded high REBA score (i.e. 9) are twisting, lowering and lifting, which is considered as high risk. Based on NIOSH Lifting Equation, the value of RWL for MMH tasks is 15.82 and the lifting index (LI) is 0.73. The recommended control measures to be implemented in Company X include to limit the working time, administrative control, training, and reduce the weight of the load carried and lifted by the MMH workers.

Keywords: Ergonomics, manual handling, REBA, NIOSH, MNQ

1. INTRODUCTION

Normally, a production line in a manufacturing company mainly involves equipment, machinery, and human workers. In heavy industry, innovation in machinery and equipment have substituted human workers to increase safety and productivity. However, in the manufacturing industries, human workers are still required mainly to perform manual handling tasks and processes that are not done by machines. Manual Material Handling (MMH) tasks involve activities such as carrying, lifting, lowering, pushing or moving loads. Repetitive and incorrect MMH techniques could increase the ergonomic risk and have known to caused work-related musculoskeletal disorders (WMSDs). WMSDs are injuries in ligaments, tendons, muscles, nerves and joints (Nurmianto et al. 2015). WMSDs have been a major concern in the manufacturing industry as it affects on workers' productivity, absenteeism, turnover, compensation and work quality (Qutubuddin et al., 2013). However, lack of knowledge and initiative for ergonomics application in manufacturing industry could further increase the chance of exposing the human workers to ergonomic risk factors.

In this study, the three techniques used to estimate the level of workers' discomfort and to assess their ergonomic risks exposure are by using observational techniques and instruments, such as Modified Nordic Questionnaire (MNQ) (Kuorinka et al., 1987), Rapid Entire Body Assessment (REBA) (Hignet & McAtamney, 2000) and Revised NIOSH Lifting Equation (Waters et al. 1993). Observational techniques are used widely in the majority of companies because, they do not interfere with the work process, convenient and low cost. The natural working position is observed from the angular deviation of employee's body segments (Lasota, 2015). MMH tasks are closely related to musculoskeletal symptoms. According to Gallis (2006), the three factors which contribute towards developing musculoskeletal disorders are the work postures, load, and technique used during performing MMH activities. A study done by Syazwani et al. (2016) indicate long working hours with short rest period is also considered a major ergonomic risk factor.

Lin & Radwin (1998) study results showed repetition rate and force used during wrist flexion task using a power grip as risk factors for developing discomfort in hand-task. While historical changes, which demonstrates skeletal tissue, soft tissue and nerves may be sensitive to repetitive movement, and tissue disruption can lead to functional impairment (Revel et al. 1992). Latko et al. (1999) study also found that repetition can be associated with worker-reported discomfort at the wrist, hand and fingers and symptoms consistent with CTS. In other words, workers' performing high repetitive tasks have 2 to 3 times higher risk compared to low repetitive tasks. Thus, it can be concluded that repetitive movement can lead to WMSD if no further action was taken to solve this problem.

Work technique, work organization, and individual characteristics or combined risk factors may influence the musculoskeletal disorder (Gallis, 2006). Workers who exposed to ergonomic risk such as MMH have a high possibility to suffer pain on the entire body parts or muscle pain (Deros et al. 2016). Excessive physical effort can cause back injuries and shoulder injuries (Parejo et al. 2012). Yeung at al. (2002) have found that both lower back and shoulder are the most prevalent locations of musculoskeletal symptoms. This study was conducted to identify working postures practiced among MMH workers in Company X, to assess their associated ergonomic risks and to determine musculoskeletal discomfort among them.

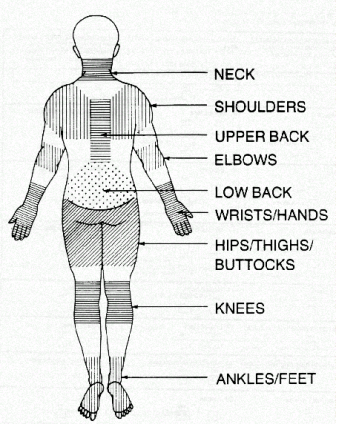
2. RESEARCH METHODOLOGY

This study was conducted at Company X located in Port Klang, Selangor, Malaysia. It was based on workers' complaints of suffering back pain symptoms and problems due to performing MMH activities. Company X is involved in metallurgy industry which produces sintered metal parts. Powder metal is the raw material for used for manufacturing the sintered metal parts. The powder metal will undergo a few processes to produce sintered automotive part. The processes involved are compacting, sintering, sizing, cutting, treatment, a magnetic crack detector (MCD), sorting, packing and quality check. This study focuses only one section which is a Magnetic Crack Detector (MCD) Process. There are 11 MMH workers working at MCD section. The manufacturing processes involved in manufacturing the sintered metal parts are magnetized, inspection under UV light and demagnetized. The majority of the processes involved manual handling of heavy poly-box, lifting, repetitive motion, twisting, lowering and reaching. The three methods used to identify Work-related Musculoskeletal Disorders (WMSDs) among the MMH workers were MNQ, observation on their working posture using the Rapid Entire Body Assessment (REBA) and NIOSH Lifting equation.

2.1. Modified Nordic Questionnaires (MNQ)

MNQ can be used to assessed body parts with WMSDs problem were distributed to the 11 MMH workers in the MCD section. This was done in order to obtain information on their perceptions and level of discomfort they had experienced. To identify the musculoskeletal symptom among MMH workers at MCD section, MNQ was used (Kuorinka et al. 1987). The main purpose of this MNQ is to identify the prevalence of musculoskeletal symptoms among MCD section MMH workers. MMH workers were guided on how to answer all the MNQ questions correctly. This MNQ is comprised of a diagram that showed nine body parts: neck, shoulder, upper back, lower back, elbow, hand, thigh, knee and leg as shown in Figure 1. The main purpose is to assist the 11 MMH workers identifying the correct body parts when answering the questions. The numeric scale used ranging from 1 (not painful) until 10 (very painful) can be used to determine the level of WMSDs experienced by the workers.

Figure 1: Nine Body Parts According to MNQ



2.2. Rapid Entire Body Assessment

Eleven MMH workers took part in the observation working posture while performing MMH activities using the Rapid Entire Body Assessment (REBA). The main purpose of this working posture assessment is to associate risk of musculoskeletal injury from the working posture practiced. REBA can be used to identify, assess and evaluate the entire body segments for Work-related Musculoskeletal Disorders (WMSDs) to entire body segments that include the following body parts: neck, shoulder, upper back, lower back, elbow, hand, thigh, knee, legs, trunk, arms, and wrist. Having done that, overall scores for all body parts were calculated using REBA worksheet and every posture has its scores that need to be assigned. Additional scores for muscles and force also were also calculated to get the final REBA score. The final REBA score will indicate the level musculoskeletal injury risk faced by the MMH workers.

2.3. Observation

During the observation MMH workers were asked to perform their usual activities, it was captured and recorded by using a video camera. This was done for the purpose to properly check and evaluate their daily working postures and to avoid any assessment mistakes that give a false or wrong result. All photos and videos recorded during the observation were utilized to identify ergonomic working postures of the MMH workers while performing their daily MMH tasks.

2.4. NIOSH Lifting Equation

The NIOSH Lifting Equation (Water et al. 1993) is a method to assess risk of low-back disorder with repeated lifting. The value of Recommended Weight Limit (RWL) and the Lifting Index (LI) need to be identified according to the following formula.

$$\mathbf{RWL = LC \times HM \times VM \times DM \times AM \times FM}$$

where:

LC : Load Constant (23 kg)

HM : Horizontal Multiplier Factor

VM : Vertical Multiplier Factor

DM : Distance Multiplier Factor

AM : Asymetric Multiplier factor

FM : Frequency multiplier Factor

CM : Coupling Multiplier Factor

$$\mathbf{LI = L/RWL}$$

where:

LI : Lifting Index

L : Actual load weight

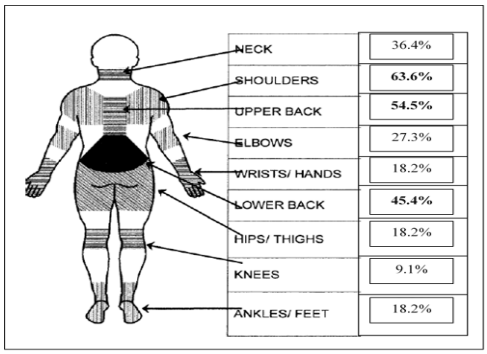
RWL: Recommended Weight Limit

3. RESULTS AND DISCUSSIONS

3.1. Modified Nordic Questionnaire (MNQ) Results

Based on the MNQ given to MMH workers, it was found many of them had experienced serious pain; 63.6% at the shoulders, 54.5% at the upper back, and 45.4% at the lower back areas. The least pain they suffered is at the knee with only 9.1% of MMH workers. Four out of 11 MMH workers have taken medical leave due to the shoulder and back pain in the last of 12 months. This is due to their routine work activity performing MMH of the heavy poly box. The activity involves lifting, lowering, reaching, twisting and repetitive. The percentage of MMH workers that experience musculoskeletal discomfort at the nine body parts according to the MNQ is shown in Figure 2.

Figure 2: Musculoskeletal discomfort at 9 body parts according to MNQ



3.2. REBA Score

REBA score as tabulated in Table 1 and Figure 3 are expressed on a scale of 1 to 15. The five categories of scoring used in REBA are from negligible risk until very high risk. Score 1 is categorized as a negligible risk. Score 2 and 3 is categorized as low risk, which may require changes. Score 4 to 7 falls under medium risk, where further investigation and necessary changes are required. While, score 8 to 10 falls into a high risk category, which required further investigation and implementation changes to reduce risk. Score 11 to 15 is considered very high risk and needs to carry out further investigation and implement changes immediately. The three types of postural movements that have REBA score of 9 are twisting, lowering and lifting are considered as high risk. Meanwhile, both repetitive and reaching postural movements have REBA score of 6 and 5 respectively are considered as a medium risk. Figure 4 show the five most common tasks and postures practiced by the MMH workers in MCD section at Company X.

Table 1: Risk Levels of Ergonomic Injury Corresponding to REBA Scores and Indication

Posture	REBA Score	Risk Level (REBA)	Action
Repetitive	6	Medium	Further investigate. Change soon.
Lowering	9	High	Investigate and implement change.
Twisting	9	High	Investigate and implement change.
Reaching	5	Medium	Further investigate. Change soon.
Lifting	9	High	Investigate and implement change.

Figure 3: REBA Score

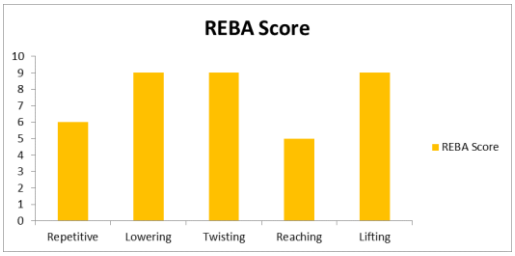


Figure 4: Workers' MMH tasks

Based on the REBA assessments carried out in Company X, repetitive hand movement falls under medium risk score, which need to take further action and changes in the work posture. Repetitive movements involving the same joints and muscle over and over again when doing the same motion can become hazardous thus increasing the risk of WMSD. Muscles fatigue and joints tiredness due to repetitive works indicate the need to change the work pattern. If no further action was taken, it will cause more pain and inflammation to the muscles and joint such as tendinitis, tenosynovitis and Carpal Tunnel Syndrome (CTS).

REBA result for reaching was classified as medium risk, which means need further investigation and change soon. Compared to the five MMH activities, reaching is the lowest risk to musculoskeletal disorder. This is due to the work activity is very mild. The worker need to stand for certain period of time and to reach and collect the part, which is located just in front of the MMH worker. There is no twisting or take far object. The worker need to move his hand forward to collect and place the part. The weight of the part is less than 2 kg. Introducing frequent rest time a few minutes every hour can reduce the risk of musculoskeletal disorder.

One of the step to prevent this problem is to limit the time duration by the worker doing the same motion and tasks over and over again. When analyzing specific job for solving ergonomic problems, a few factors such as the weight of object being handle, body position, repetition of certain movement and how worker grip the objects must be taken into considerations. Engineering controls like workstation design, working method, tool and handle design can also be used control this problem. At the same time, the worker should always practice proper working technique, employee conditioning, inspection, feedback and equipment maintenance. In term of administrative control, should reduce the duration, frequency, and severity of exposures to the ergonomic hazard. For example, reducing the number of repetition per hour or provide short rest period to relieve fatigue.

For lowering task movement, after conducting the observation and scoring using REBA, it was classified as high risk. This type of posture or task movement needs to be investigated and implement the required change. Due to forward movement, shoulders, upper and lower back positions are the highest frequency of movements performed by the MMH workers. Workspace design and task requirement are the contributing factors that can lead to good working posture (Zein et al. 2015). At present, the majority of the industrial workplace is poorly designed, and workers needs to adapt with the workspace design to complete the task given. In this study, the lowering posture need the MMH employee to handle a load

of more than 15 kg. Thus, one of the recommendation is to reduce the load that MMH employee needs to lower. Company X, management need to provide training and education to MMH workers on ergonomics posture, when performing lifting and lowering tasks. This would create awareness towards the importance of ergonomics risk factors at the workplace and its association with MMH employee safety and health.

Based on the REBA result, lifting activity score is 9. It can be classified as high risk and Company X, management need to conduct a proper investigation and implement the recommended changes to improve their workplace design and to reduce the ergonomic risk factor.

3.3. NIOSH Lifting Equation

The highest frequency and severity is lifting and lowering. Lifting and lowering activities involving a load of 11 to 15 kg per polybox. Based on the NIOSH Lifting Equation, the calculated value of Recommended Weight Limit (RWL) is 15.82 kg, which indicates that the load lifted is acceptable. According to NIOSH Lifiting Equation, if Lifting Index (LI) value is more than 1.0, then the workstation and task should be redesigned. In this case, the calculated Lifting Index (LI) value is 0.73 (i.e. less than 1.0), therefore; there is no need to redesign the workstation, lifting and lowering tasks. However, MMH workers may still suffer from backpain and shoulder pain because they have to manually carry on average 40 - 60 polybox per day and using improper lifting posture.

4. RECOMMENDATIONS

4.1. Administrative Improvements

Observe how different MMH workers perform the same tasks to get ideas for improving work practices or organizing the work. Administrative improvements that can be carried out, such as: job rotation, can help reduce MMH workers' exposures to risk factors by limiting the amount of time workers spend on "problem jobs". Atishey et al. (2013) have recommended to implement the following adminstrative improvement activities:

a. Alternate heavy tasks with light tasks.

b. Provide variety in jobs to eliminate or reduce repetition (i.e. overuse of the same muscle groups).

c. Adjust work schedules, work pace, or work practices.

d. Provide recovery time (e.g. short rest breaks).

e. Rotate workers through jobs that use different muscles, body parts, or postures.

4.2. Training

Training and education are an effective way of increasing awareness on ergonomics issues (Tayyari & Smith, 1997). In that case, organization need to provide training to workers especially about safety and health matters to ensure that particularly sensitive groups of workers are protected against any dangers which specifically affect them including in relation to the manual handling of loads. The manual handling training needs to be specific to the tasks involved. It should aim to ensure that the employee understands the reasons for doing the job with least risk, can recognize the risks and decide the best way to control the risk and can perform the task in that way. The instructor should ensure that all participants understand and can apply good handling techniques. Manual handling skills and the review of the findings of the risk assessment process should be an integral part of the training programme.

In this case, employers need to engage external competent trainers. Suitable trainers must have knowledge, skill and experience in the subject, have the ability to communicate effectively, are able to give practical demonstrations and have appropriate teaching materials. Besides that, refresher training is also important and will be conducted at intervals not more than every three years and when there is any major change in the work involved or equipment used or when an employee is transferred to another activity requiring different loads to be handled.

4.3. Reduce Weight of Load

Reduce the load weight especially if the same operation is repeated many times. As a rough guide, reduce the weights by 30% if the operation is repeated once or twice a minute, by 50% if it is repeated 5 - 8 times a minute, and by 80% when it is repeated more than 12 times a minute. Also, practice team lifting as a temporary measure until a more permanent improvement can be found. If possible, try to find a co-worker of similar height to help with the lift. Team lifting actually can reduce the load in half.

REBA result for twisting posture is 9 which is classified as high risk. Management need conduct an investigation and implement new changes in the workplace design or procedures in order to reduce the ergonomic risk factor. Twisting could cause back injury. Based on the video recording, MMH workers did a repetitive range action of twisting while moving the

load to the side. Although the weight load is less than 5kg, the repetitive motion could cause serious damage to the back tissue. One of the recommendation is to change the working techniques. The Company X management need to provide training on the proper MMH techniques of working to avoid twisting by turning the whole body while transferring the load. In addition, workspace design need to be taken into account by ensuring there is enough spaces for workers to turn around when moving the load to avoid twisting.

4.4. Stretching and Exercise Programs

In addition, it is recommended that the MMH workers perform daily body stretching exercises to prevent from them from stiffness and can relax their muscle. This is in-line with Choi & Woletz (2010) findings that musculoskeletal disorders or injury can be prevented by stretching and exercise programs in different occupations.

5. CONCLUSION

Results obtained from the Modified Nordic Questionnaire (MNQ) showed 63.6% of the MMH workers had experienced serious pain at the shoulders, 54.5% at the upper back, and 45.4% at the lower back areas. In the last 12 months, 36% of MMH workers have taken medical leave due to the shoulder and back pain problem. Twisting, lowering and lifting tasks postural movements that have REBA score of 9 are considered as high risk. Meanwhile, both repetitive and reaching postural movements have REBA score of 6 and 5 respectively, which are considered as medium risk. For administrative control, Company X management should reduce the MMH workers exposures to ergonomic hazard, with respect to duration, frequency and severity. For example, reducing the number of repetition per hour or provide short rest period to relieve fatigue. The pain that the MMH workers felt on their shoulders, upper back and lower back are due to repetitive motions such as lifting, lowering heavy polybox, and twisting. In short, the ergonomic risk factors in Magnetic Crack Detection section in Company X is high especially that involves lifting, lowering and twisting postures. Company X management needs to carry out investigation, implement changes and improvement in their working procedures, techniques and workspace design to reduce WMSDs risks.

REFERENCES

Choi, S.D., Woletz, T. 2010. “Do stretching programs prevent work-related musculoskeletal disorders”. Journal of Safety, Health and Environmental Research, vol. 6, no. 3, pp. 1-19.

Deros, B. M., Daruis, D.D.I., Ghani, J. A., Salleh, C., Wazir, R. M., Yasim, M. K., Hassan, N.H., Saidi, M. S. 2016. “A case study on manual handling at an electronic component manufacturing company”. Malaysian Journal of Public Health Medicine, vol. 1, (Special Issue 1), pp. 1-6.

Gallis, C. 2006. “Work-Related Prevalence of Musculoskeletal Symptoms among Greek Forest Workers”. International Journal of Industrial Ergonomics, vol. 36, no. 8, pp. 731–736.

Hignett, S. & McAtamney, L. 2000. “Rapid entire body assessment (REBA)”, Applied ergonomics, vol. 31, no. 2, pp. 201-205.

Kuorinka, I., Jonsson, B., Kilbom, A., Vinterberg, H., Biering-Sørensen, F., Andersson, G., Jørgensen, K. 1987. “Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms”, Applied ergonomics, vol. 18, no. 3, pp. 233-237.
Latko, W.A., Armstrong T.J., Franzblau, A., Ulin, S.S., Werner, R.A., Albers, J.W. 1999. “Cross-Sectional Studies of the relationship between repetitive work and prevalence of upper limb musculoskeletal disorders”, American Journal of Industrial Medicine, vol. 36, pp. 248-259.

Lasota, A. M. 2015. “Ergonomic evaluation of physical risk for packing line operators”, Logistics and Transport, vol. 26, no. 2, pp. 11-20.

Lin, M.I., Radwin, R.G. 1998. “Agreement between a frequency-weighted filter for continuous biomechanical measurements of repetitive wrist flexion against a load and published pshcophysical data”, Ergonomics, vol. 4, pp. 459-475.

Mittal, A., Sharma, H. K., Mittal, K. 2013. “Ergonomic Risk Controls in Construction Industry - A Literature Review”, International Journal of Emerging Research Management & Technology, vol. 2, pp. 28-33.

Nurmianto, E., Ciptomulyono, U., & Kromodihardjo, S. 2015. “Manual Handling Problem Identification in Mining Industry: An ergonomic perspective”, Procedia Manufacturing, vol. 4, pp. 89-97.

Parejo-Moscoso, J. M., Rubio-Romero, J. C., Pérez-Canto, S. 2012. “Occupational Accident Rate in Olive Oil Mills”, Safety Science, vol. 50, no. 2, pp. 285–293.

Qutubuddin, S. M., Hebbal, S. S., Kumar, A. C. S. 2013. “An ergonomic study of work related musculoskeletal disorder risks in Indian Saw Mills”, Journal of Mechanical and Civil Engineering, vol. 7, no. 5, pp. 7-13.

Revel, M., Andre-deshays, C., Roudier, R., Roudier, B., Hamard, G., Amor, B. 1992. “Effects of repetitive strains on vertebral end plates in young rats”, Clinical orthopaedics and related research, vol. 279, pp. 303-309.

Syazwani, N., Deros, B. M., Ab Rahman, M. N., Ezrin Hani, S., Norani, N., Shamsul Bahri, M. T., Shahriman, A.B., Rambely, A. S. 2016.

"Ergonomic risk assessment of manual handling tools by oil palm collectors and loaders", Malaysian Journal of Public Health Medicine, vol. 16, pp. 56-60.

Tayyari, F., Smith J.L. 1997. "Occupational Ergonomics Principle And Applications", Chapman & Hall (London)Te H.P, & Kleiner B.H. 2001. New Development Concerning The Occupational Safety And Health Act, vol. 43, Number ½ 2001.

Waters, T. R., Putz-Anderson, V., Garg, A., Fine, L. J. 1993. "Revised NIOSH equation for the design and evaluation of manual lifting tasks". Ergonomics, vol. 36, no. 7, pp. 749-776.

Yeung, S. S., Genaidy, A., Deddens, J., Alhemood, A., Leung, P. C. 2002. "Prevalence of musculoskeletal symptoms in single and multiple body regions and effects of perceived risk of injury among manual handling workers", Spine, vol. 27, no. 19, pp. 2166-2172.

Zein, R. M., Halim, I., Azis, N. A., Saptari, A., Kamat, S. R. 2015. "A Survey on Working Postures among Malaysian Industrial Workers", Procedia Manufacturing, vol. 2, pp. 450-459.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

PREDICTING PRESSURE LEVEL FELT UNDER THE SEAT PAN DUE TO CHANGES IN DRIVING POSITION

DOI: 10.17261/Pressacademia.2017.607
PAP-WCTIE-V.5-2017(45)-p.325-334

Khamis, N.K.[1], Deros, B.M.[1], Schramm, D.[2], Maas, N.[2], Nuawi, M.Z.[1], and Koppers, M.[2]
[1]Universiti Kebangsaan Malaysia, Malaysia, kamaliana@ukm.edu.my
[2]University of Duisburg-Essen, Germany

ABSTRACT

In ergonomics perspective, car seat is defined as one of the main workstation when the driver is performing driving task. The car seat itself provides support and space for the driver over the head, at the upper and lower back, at the buttock and also at the thigh. There has been abundant research regarding car seat in the past decades. However, in spite of vast studies on this issue, limited studies were found on sitting effect to the pressure when adopting different driving position. Hence, this study integrates objective and subjective evaluation to predict pressure felt level. Mapping of the pressures distribution was taken to the seat pan. This study is a static field experiment where data acquired are in the quantitative form. Findings show that the pressure of the heavier subject is more scattered at the buttock area, while the lighter subject has mild stress concentrated under ischium tuberosity.

Keywords: Pressure, car,seat pan, driver, discomfort

1. INTRODUCTION

Prolonged sitting and the near static seated posture while driving imposes the driver to be more restricted and has been associated with an increased risk of musculoskeletal disorders (MSDs) such as lower back pain (LBP), neck pain and shoulder pain (Andersson 1980; Balasubramanian & Prasad 2007). It was caused by the high pressure due to driver's body weight, which leads to an elevated force on the muscles that are functioning in an anaerobic setting. Due to this compressive force at the driver and seat interface, the blood flow will be obstructed through the large vessels to the lower body part. Consequently, it leads to oxygen deficiency, which can be expressed as discomfort and fatigue. In the long run, it will turn to pain and injury (Yamazaki 1992; Graf, Guggenbiihl, & Krueger 1993; Graf et al. 1995; Wilke et al. 1999; Lueder 2004; Ng, Cassar, & Gross 1995; Gross et al. 1994)

Many studies in the literature focused on seat improvement. For instance, Grandjean (1980), Hubbard & Reynolds (1984), Coelho & Dahlman (1999) and Wu, Rakheja& Boileau (1999) focused on seat design parameters and features by considering postural angle, design, and materials of the seat. In addition, several studies had combined subjective assessment tools such as comfort rating with the pressure distribution data. Past studies showed that there is a good correlation between pressure distribution data and seat comfort rating. Ng et al. (1995) had conducted another study with the same approach by developing an intelligent seat system based on the pressure data adjustment on the seat. Subjective comfort ratings (from 1=very poor to 10=very good) and anthropometric measurements were also carried out in this study when 20 respondents were required to simulate driving position in a seating buck. Thakurta et al. (1995) have compared subjective assessment of short and long driving on the 80-mile highway. Thirty-six respondents have evaluated five small cars by using comfort assessment questionnaires and the pressure distribution was mapped before and after driving. Overall, up to this date, limited studies examine the pressure distribution pattern when sitting in a different driving position with the complete set of car setup, and its effect on the driver. There were three studies which are closely associated with this matter. Andreoni et al. (2002) have recorded the posture and pressure distribution when the subject put his foot on the car pedal. However, this study did not have fixed position. In addition, this study did not require the test subjects to drive the

car. Therefore, there was no result on the pre and post activity. Porter, Gyi & Tait (2003) conducted a study by asking the subject to drive the car. However, there was no specific posture for this study. Na et al. (2005) had carried out the study by setting the seat back angle at 115^0 based on Korean population. However, there was no comparison between different posture.

This paper aims to combine this matter by examining the impact of adopting different fixed driving position based on Malaysian population towards driver when sitting on the actual car's seat pan. Subjective and objective measures assessments were used in this study. Basically, this paper contributes to the literature on the impact of pressure distribution in three ways. First, this paper document an important relationship between characteristics of the sitter, such as weight with pressure distribution. Second, our results support the hypothesis, that there is a good correlation between objective and subjective measure by developing the linear model. Third, our results depict the association between pressure interface pattern as well as different driving positions and its impact on the driver when sitting on the car seat. The paper is organized as follows. Section 2 exhibits methodology of this research work. Section 3 reports and explains the empirical findings and analysis of this research. And the final section provides conclusions for this study.

2. RESEARCH METHODOLOGY

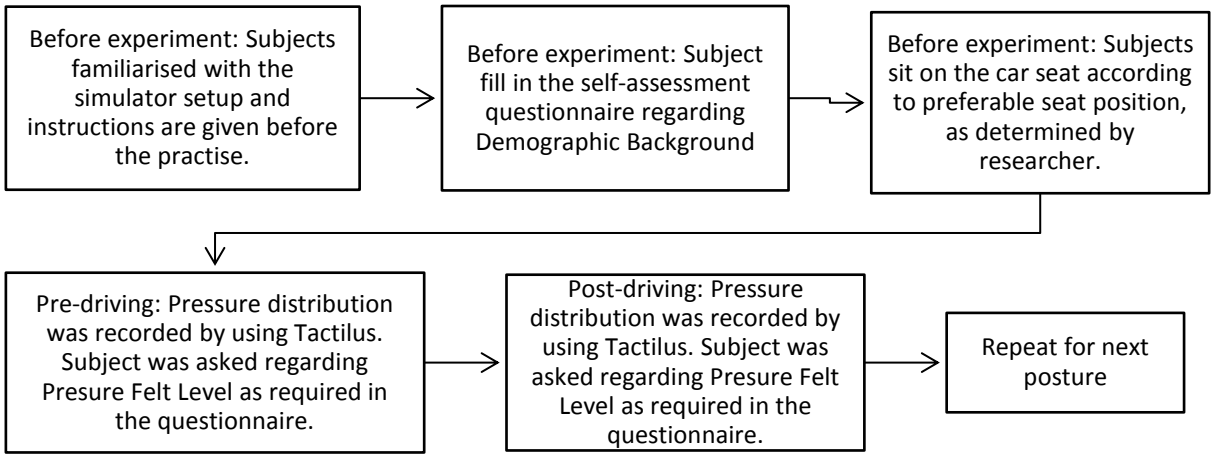
2.1. Test Subjects

Twelve female subjects (mean age of 29 years, mean height = 156 cm, mean weight = 54 kg) were assigned from the population of staff and students at Universiti Kebangsaan Malaysia (UKM) to take part in this simulator study. Gender criteria was focused on female only to reduce variations in the findings, due to male and female have major different in body charateristics and driving behavior. Each subject was required to attend one session, either in the morning session (from 9 am to 12 pm) or in the evening session (from 2 pm to 5 pm). The inclusion criteria were all respondents held a full Malaysia driving license, had at least 3 years of driving experiences and aged between 22 to 35 years. The constriction of the age range was proposed to reduce variations in the results due to age since even in normal aging, people present slight perceptive variations that have a direct attitude towards driving (Antonson et al. 2014). However, for this study there were three cases of data conflict and error during data collection. Therefore, only data from nine subjects had been used for further analysis.

All subjects were allowed to adapt with the car simulator setup and car seat adjustment before starting the experiment. The experiment was started after five minutes the subject had been in the driving position to allow them to adapt with the seat environment and fabrics. All subjects understood and complied with the oral and written instructions provided by researcher for this experiment. Information about the experiment procedure and questionnaires used was included. After receiving the complete information on the study, each subject signed an informed consent. However, before starting the experiment, the subjects were required to test the simulator in order to ensure they were familiar with the car's component; gears, steering, and acceleration as well as the simulator road condition and landscape. All subjects were instructed to drive and obey road rules for 15 minutes for each driving positions. Figure 1 demonstrates the process flow of the instruction for the subject. Section 2.2 describes the design of the simulator setup.

All subjects were allowed to adapt with the car simulator setup and car seat adjustment before starting the experiment. The experiment was started after five minutes the subject had been in the driving position to allow them to adapt with the seat environment and fabrics. All subjects understood and complied with the oral and written instructions provided by researcher for this experiment. Information about the experiment procedure and questionnaires used was included. After receiving the complete information on the study, each subject signed an informed consent. However, before starting the experiment, the subjects were required to test the simulator in order to ensure they were familiar with the car's component; gears, steering, and acceleration as well as the simulator road condition and landscape. All subjects were instructed to drive and obey road rules for 15 minutes for each driving positions. Figure 1 demonstrates the process flow of the instruction for the subject. Section 2.2 describes the design of the simulator setup.

Figure 1: Experiment Flow for Each Subject



2.2. Simulator Setup

A simulator was used in this study as displayed in Figure 2. This simulator was located in the Ergonomics Laboratory, Department of Mechanical and Materials Engineering, Faculty of Engineering and Built Environment, UKM. This simulator design and car seat parameters were quite similar to one of the national compact cars in Malaysia. The simulator consists of an adjustable driver's seat (inclination of the backrest, lower or elevate head rest, forward or backward seat), steering wheel, clutch, accelerator and brake pedals, handbrake, and manual gear shift. The screen was arranged in front of the driver and has the virtual dashboard on it when using the simulator.

Figure 2: Simulator Setup

The driving task was reduced to a lane keeping task to induce task monotony: no traffic, driving consisted in following a lane (no itinerary involved) with speed in between 50 to 70 kilometres per hour, without having to stop the car (no red traffic lights, stops) or having the need for frequent breaking intervals (no T junctions inter-sections or perpendicular turns), or having the necessities for gear and lane changings during the driving task (only change gear to gear 5 at the beginning), as well as turn signals activation. In addition, driving at the suburban scene was selected for this experiment. The simulated driving task was designed with the following requirements: the route was simple so that the drivers could complete the task easily, there were few scenery changes, there was no inclination on driving route to reduce outside stimuli, and light curvature was chosen so that drivers had to pay attention. However, there are different road characteristics along the driving journey. Some road surfaces is quite bumpy and the driver can see this change in the scene when driving.

2.3. Experiment Design and Procedure

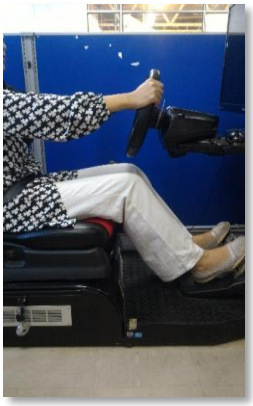
Two different driving positions with fixed backrest position at 100^0 were carried out in this study, as depicted in Figure 3(a) and (b). This fixed position at the back rest was based on the past study conducted by Daruis (2010), Mohamad et al. (2010) and Mohamad et al. (2016). Two different driving position were: i) Posture A: closest distant to car controls and ii) Posture B: far distant from car controls, as long as the test subjects could operate the car controls and sat comfortably leaning against backrest of the car seat. During recording the pressure distribution, the test subject have to place their hand at 10

and 2 o'clock as shown in Figure 3. In addition, the subject needed to ensure the right leg at the car pedal, while the left leg at the car simulator floor, near to clutch pedal.

Figure 3: Driving posture

(a) Posture A

(b) Posture B



As mentioned in Section 2.1, each subject was required to drive the simulator for 15 minutes. As mentioned by Porter, Gyi & Tait (2003), some seats are considered uncomfortable after approximately 15 minutes. İn addition, Shackel et al. (1969), also took 15 minutes as the ultimate time to evaluate the sitting comfort on ten chairs. Therefore, 15 minutes of driving task for each driving position should be enough to investigate the driver's state pattern. This pattern can be seen from pressure distribution measurement and pressure felt assessment on pre (before) and post (after) driving activity. Section 2.4 and 2.5 explained on the pressure distribution measurement and self-assessment questionnaire design.

2.4. Pressure Distribution Measurement

Pressure measure is sensitive to postural changes of varied angulation and has good correlation with subjective comforts, by determining the maximum pressure, average pressure ratio and maximum pressure gradient (de Looze, Kuijt-Evers, & van Dieën, 2003; Shen & Galer, 2015). In this study, the Tactilus® pressure mapping from Sensor Products Incorporations (SPI) was used as shown in Figure 4. The system includes a 22 x 22 sensor pad calibrated 0-5psi with 32 x 32 sensor matrix. The interface pressure uses thin and flexible sensor arrays. By scanning the grid and measuring the electrical resistance at each grid point, the pressure distribution on the sensor's surface can be determined. The scanning electronics are packaged in a handle assembly that clips onto the sensor's interface tab and provides the electrical connection to each sensing cell.

Figure 4: Tactilus® Pressure Mat

All subjects were requested to wear suitable clothes for driving but without heavy seams, buttons or pockets in order that there was minimal effect on the pressure readings. This requirement is necessary to avoid false seat or backrest interface pressure readings. These mats were securely attached to the seat using strips of masking tape. Care was exercised to ensure that the mats were placed in a consistent location from subject and seat or backrest. Subjects adopted the driving positions for this measurement (hands on the steering wheel and looking ahead), held for 30 seconds. Then, the pressure distribution measurement of pre and post driving for the car seat was taken about one minute for each position. The mats were removed and the occupant was asked to re-enter the seat in order to complete the survey without interference from the mats. The reason behind this instruction was, subjects faced difficulty in rating the appearance of the seat if they were sitting on it.

2.5. Design of Self-Assessment Questionnaire

The first section required the subject to provide the information regarding age, gender, height, weight and driving experience. Next section required the test subject to identify the perception of pressure felt level based on driving posture. In this section, seat part assessment was categorized into two segments, buttock and thigh. This section used the Visual Analogue Scale (VAS) for perception of pressure felt. Basically, VAS is referred to 10 cm continuous horizontal line with each point have a different definition. In this case, 0 was referred as no pressure felt, and 10 referred to the extreme pressure felt.

2.6. Statistical Analysis

Data from the experiment was used to predict and estimate the dependent variable (DV) based on independence variable (IV) by using the Regression Method. It was used to predict the value of a variable based on the value of another variable. The variable to predict is called the DV, while the variable to predict the other variable's value is called the IV or known as the predictor variable. In this study, the DV is based on the subjective assessment, while the IV is referred to the objective assessment (pressure distribution map measurement). The highest mean score for each IV had been used to develop the model in this section. Only data from pre-driving activity was used because it provides more accuracy, where the subject have enough break time in between driving position before conducting the experiment for each task.

Based on this method, the regression coefficient (K), regression constant (c), multiple correlation coefficients (R), coefficients of determination (R^2) and significance level (p) can be determined. The acceptance or rejection of the null hypothesis (Ho) for this study can also be determined based on the newly developed model (H_o: $\beta1=0$ and $\beta2=0$: $\beta1\neq0$ and $\beta2\neq0$). The H_o can be rejected if the p-value is low ($p<0.05$). As a result, it can estimate the drivers' condition based on these integrated assessment methods. In other words, a predictor that has a low p-value is likely to be a significant predictor to the model because changes in the predictor's value are connected to changes in the response variable. The linear equation to predict the drivers' state was demonstrated in Equation (1).

$$Y= K_1X_1 + c \qquad (1)$$

Where;

Y = Drivers' state based on subjective assessment (DV)

K = Regression coefficient as the contributor factor towards pressure pattern

X = Factor that contributes towards pressure pattern (IV)

c = Regression constant

3. FINDINGS AND DISCUSSIONS

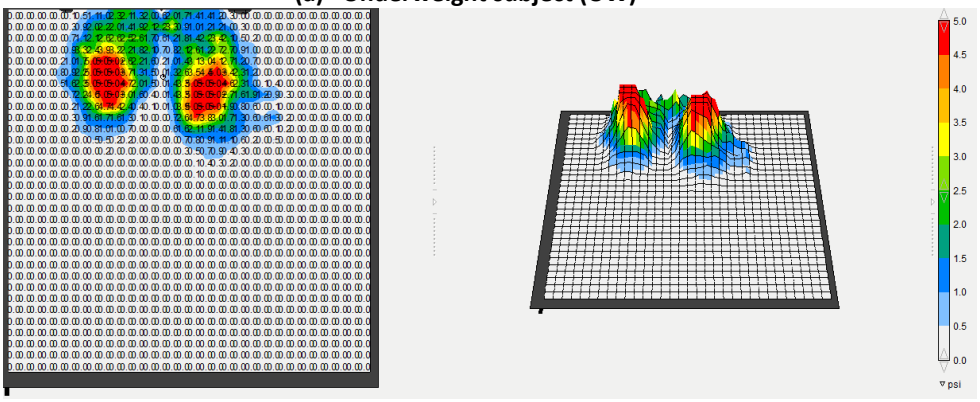
This section describes the findings for the pre-driving task and posts driving task from the seat pan. This section explains the seat pan pressure distribution's findings by focusing on the Body Mass Index (BMI) and percentile group of all subjects.

3.1 Seat Pan Interface Pressure Between Subjects According to the BMI and Percentile

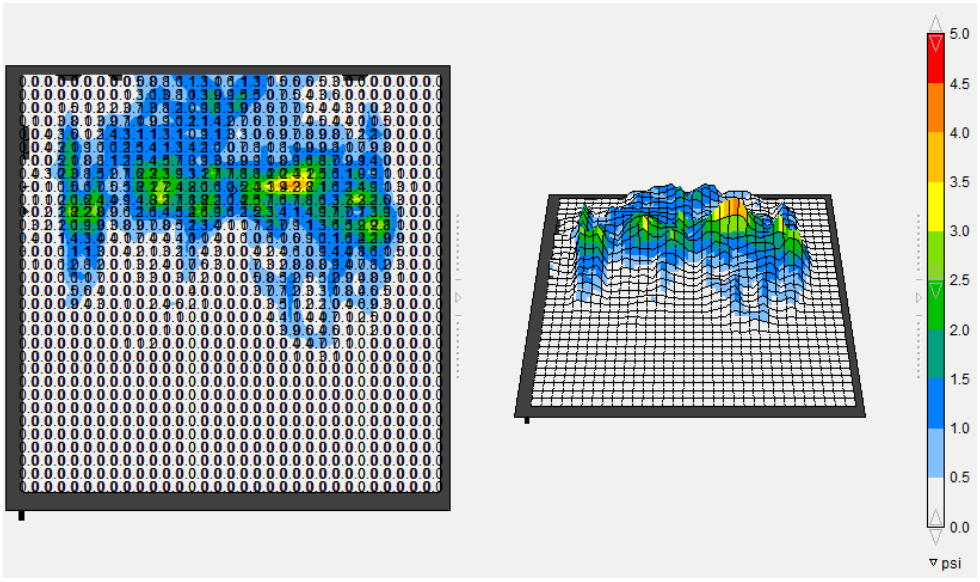
Figure 5(a), (b) and (c) illustrate three examples of the pressure distribution for the seat pan from three representatives of the underweight subject (BMI=17.1), normal weight subject (BMI=20.8) and overweight subject (BMI=27.3) from the Tactilus software and conversion Excel 32 x 32. With regards to the Figure 5 (a) to (c), the pressure of the heavier subject is more scattered at the buttock area, while the lighter subject has mild stress concentrated under ischium tuberosity.

Figure 5: Pressure distribution pattern based on the BMI for the seat pan

(a) Underweight subject (UW)



(b) Normal weight subject (NW)



(c) Overweight subject (OW)

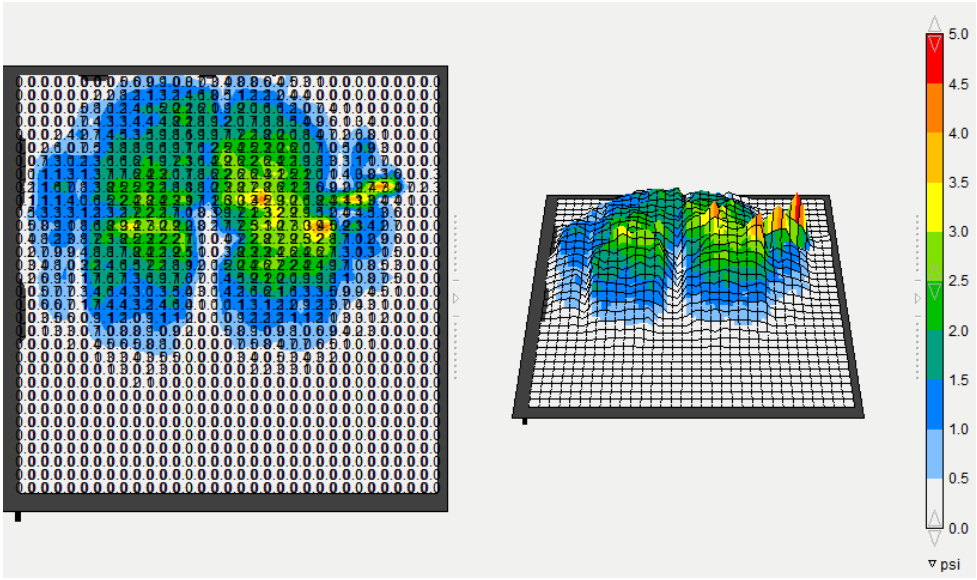


Table 1 demonstrates the mean pressure of each driving position for all subjects according to the percentile group. The percentile value used in this study is primarily referred to past study conducted by Daruis (2010) based on Malaysian population. Mean percentile or denoted as the 50th percentile for this case is 1.567 cm (height) and 54 kg (weight). Based on Table 1 without considering the percentile group, the buttock part is significantly higher than the thigh part, with the mean pressure is in between 1 to 2 psi. Moreover, generally, the group with below 50th percentile produces the highest mean pressure for the buttock part. It is because the highest mean pressure was concentrated under ischium tuberosity. As a result, the pressure value at the buttock part of this group was significantly higher. Meanwhile, the group with above 50th percentile generates the highest mean pressure at the thigh due to a scattered pattern of the pressure distribution. Next subsection clarifies the actual pattern based on these findings with the aid of the graph illustration and statistical analysis.

Table 1: Seat Pan Pressure Distribution According to the Body Size for Posture A and B

	Percentile < 50th		Percentile > 50th	
Percentile	**Mean**	**SD**	**Mean**	**SD**
A pre-buttock	2.71	0.74	2.64	0.31
B pre-buttock	2.20	0.41	1.96	0.11
A post buttock	3.59	1.06	2.79	0.73
B post buttock	3.20	0.80	2.47	0.50
A pre-thigh	0.56	0.54	1.05	0.25
B pre-thigh	0.94	0.43	1.14	0.40
A post thigh	0.74	0.52	0.98	0.45
B post thigh	1.40	0.51	1.40	0.51

3.2 Relationship Between Pressure Distribution Map And Pressure Felt Level

As described in Section 2.6, from the subjective and objective assessment for pre-driving activity, the highest pressure value was indicated at the buttock part at posture A with the mean pressure score were 6.24 and 2.67 psi from both assessments respectively. Therefore, for this part, the DV was the data from the subjective assessment (posture A-buttock pre-activity), meanwhile, the IV was the data from the pressure distribution measurement at the buttock part at posture A (pre-activity). Before performing the Regression Analysis, the unit of the pressure distribution had been converted to SI unit (from psi to kg/cm-2). This action was applied to ensure the standardization of the unit for all variables in the model.

Table 2 shows the findings from the Regression Method. This table provides the R, R^2, Adjusted R^2, and the Standard Error of the Estimate (SEE), which can be used to determine how well a regression model fits the data. R can be considered to be one measure of the quality of the prediction of the DV; in this case, the pressure felt by the subject. A value of 0.963 indicates a good level of prediction. In general, if R value is more than 0.71, it shows that there was a strong correlation between variable (Piaw 2006). The R Square column represents the R^2 value, which is the proportion of variance in the dependent variable that can be explained by the IV. Based on Table 5.1, 92.7% explained the variability of the DV. The Adjusted R Square is intended to control for overestimates of the population R square resulting from small samples. In this case, the Adjusted R Square was 0.916, smaller than R Square. Meanwhile, the SEE is the standard deviation of the residuals. As the R Square increases, the SEE will decrease. In this case, the SEE is 0.281.

Table 2: Summary of the Model for Pressure Distribution and Pressure Felt Level

Model	R	R Square	Adjusted R Square	SEE
1	0.963	0.927	0.916	0.281

Table 3 shows the ANOVA table for the pressure model at the buttock. The F-ratio in Table 5.2 indicates whether the overall regression model is a good fit for the data. The table shows that the IV statistically significantly predict the DV, F (1, 7) = 88.254, p < .0005.

Table 3: ANOVA Table of the Pressure Distribution and Pressure Felt Perception

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.951	1	6.951	88.254	0.000
	Residual	0.551	7	0.079		
	Total	7.502	8			

Table 4 depicts the coefficient table for this model. This table provides the general form of the equation to predict the pressure felt by the subject. Unstandardized coefficients indicate how much the DV varies with an IV when all other IV (if any) are held constant. According to Table 5.3, the significant level for constant and the IV (pressure distribution at the buttock) were less than 0.05. It indicates the possibility to obtain t value for the constant was 2.393, the slope for pressure distribution at the buttock was 9.394. Therefore, the H0 would be rejected (H0: β1=0 and β2=0).

Table 4: Coefficient Table for the Pressure Distribution and Pressure Felt Perception

Model		Unstandardized Coefficient		Standardized Coefficient	t	Sig.
		B	Std. Error	Beta		
1	Constant	1.284	0.536		2.393	0.048
	Pressure distribution map at the buttock	1.860	0.198	0.963	9.394	0.000

With regards to Table 4, the equation model to predict the drivers' state based on pressure felt at the buttock can be used. It was indicated in Equation (2):

$$Y = 1.86X_1 + 1.284 \quad (2)$$

Based on the findings, there was a strong evidence that the pressure distribution data at the buttock was a significant predictor to estimate the perception on pressure felt at the buttock. It was shown by the standardized coefficient magnitude, Beta (β) = 0.963. In addition, as mentioned earlier, pressure distribution data at the buttock explained 92.7% from the variance occurred. It was explained at the significance level, α = 0.05 (5%) or at confidence level 95%.

5. CONCLUSION

All in all, this study shows that the distribution of pressure over seat pan was slightly influenced by the characteristics of the sitter's body part, in term of the weight and also buttock-knee length as mentioned in the thorough analysis of the seat pan. Based on the findings from the seat pan, the pressure of the heavier subject is more scattered at the buttock area, while the lighter subject has mild stress concentrated under ischium tuberosity. With regards to driving position, posture A depicted the highest mean pressure at the buttock compared to posture B. In term of driving position, there is a difference in all postures at the buttock. Based on the model, a linear regression was run to predict the pressure felt from pressure distribution map. This variable statistical significantly predicted the pressure felt, F(1, 7) = 88.254, p < 0.0005, R2 = 0.927. In

addition, due to the low value of p ($p<0.05$), the Ho (Ho: $\beta1=0$) was rejected. Hence, the linear model to predict the pressure felt at the buttock was: $Y = 1.86X_1 + 1.284$.

REFERENCES

Andersson, G.B.J. 1980. *The load on the lumbar spine in sitting postures*. D.J. Oborne and J.A. Levis (eds.) Human Factors in Transport Research. New York: Academic Press, 231-239.

Andreoni, G., Santambrogio, G.C., Rabuffetti, M., & Pedotti, A. 2002. Method for the analysis of posture and interface pressure of car drivers. Applied Ergonomics, 33(6): 511-522.

Antonson, H., Ahlström, C., Mårdh, S., Blomqvist, G., & Wiklund, M. 2014. Landscape heritage objects' effect on driving: a combined driving simulator and questionnaire study. Accident Analysis and Prevention, 62: 168-177.

Balasubramanian, V., & Adalarasu, K. 2007. EMG-based analysis of change in muscle activity during simulated driving. Journal of Bodywork and Movement Therapies, 11(2): 151-158.

Coelho, D.A., & Dahlman, S. 1999. A pilot evaluation of car seat side support : Leading to a redefinition of the problem. International Journal of Industrial Ergonomics, 24: 201-210.

Daruis, D.D.I. 2010. Model bersepadu untuk pengukuran dan penilaian tahap ketidakselesaan tempat duduk kereta. Tesis Doktor Falsafah, Universiti Kebangsaan Malaysia.

De Looze, M.P., Kuijt-Evers, L.F.M., & van Dieën, J. 2003. Sitting comfort and discomfort and the relationships with objective measures. Ergonomics, 46(10): 985-997.

Graf, M., Guggenbiihl, U., & Krueger, H. 1993. Investigations on the effects of seat shape and slope on posture, comfort and back muscle activity. International Journal of Industrial Ergonomics, 12: 91-103.

Graf, M., Guggenbuhl, U., & Krueger, H. An assessment of seated activity and postures at five workplaces. International Journal of Industrial Ergonomics, 15(2): 81-90.

Grandjean, E., 1980. Sitting posture of car drivers forms the point of view of ergonomics. Hum. Fact. Trans. Res., 2: 205-213.

Gross, C.M., Goonetilleke, R.S. and Menon, K.K., 1994. New developments in the biomechanical assessment and prediction of seat comfort. Hard facts about soft machines: The ergonomics of seating.

Hubbard, R.P., & Reynolds, H. M. 1984. Anatomical geometry and seating. SAE Technical Paper 840506. Warrendale, PA: Society of Automotive Engineers, Inc

Lueder, R. 2004. Ergonomics of seated movement: A review of the scientific literature. Humanics Ergosystem Inc.

Mohamad, D., Deros, B.M., Daruis, D.D.I., Ramli, N.F., & Sukadarin, E.H. 2016. Comfortable driver's car seat dimensions based on anthropometrics data. Iranian Journal of Public Health, 45(1): 106-113.

Mohamad, D., Deros, B.M., Wahab, D.A., Daruis, D.D.I., & Ismail, A.R. 2010. Integration of comfort into a driver's car seat design using image analysis. American Journal of Applied Sciences, 7(7): 937-942.

Na, S., Lim, S., Choi, H.-S., & Chung, M.K. 2005. Evaluation of driver's discomfort and postural change using dynamic body pressure distribution. International Journal of Industrial Ergonomics, 35(12): 1085-1096.

Ng, D., Cassar, T., & Gross, C.M. 1995. Evaluation of an intelligent seat system. Applied Ergonomics, 26(2): 109-116.

Piaw, C.Y. 2006. Asas statistik penyelidikan. Cetakan Pertama. Shah Alam, Selangor: McGrraw-Hill (Malaysia).

Porter, J.M., Gyi, D.E. & Tait, H.A., 2003. Interface pressure data and the prediction of driver discomfort in road trials. Applied Ergonomics, 34(3): 207-214.

Porter, J.M., Gyi, D.E. & Tait, H.A., 2003. Interface pressure data and the prediction of driver discomfort in road trials. Applied Ergonomics, 34(3): 207-214.

Shackel, B., Chidsey, K.D., Shipley, P., 1969. The assessment of chair comfort, Ergonomics, 12(2): 269-306

Shen, W., & Galer, I.A.R. 2015. Development of a pressure related assessment model of seating discomfort. Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Thakurta, K., Koester, D., Bush, N., Bachle, S. 1995. Evaluating short and long term seating comfort. Society of Automotive Engineers, 950144.

Wilke, H.J., Neef, P., Caimi, M., Hoogland, T. & Claes, L.E. 1999. New in vivo measurements of pressures in the intervertebral disc in daily life. Spine, 24(8): 755-762

Wu, X., Rakheja, S., & Boileau, P.-É. 1999. Distribution of human–seat interface pressure on a soft automotive seat under vertical vibration. International Journal of Industrial Ergonomics, 24(5): 545-557.

Yamazaki, N. 1992. Analysis of sitting comfortability of driver's seat by contact shape. Ergonomics. 35(5-6): 677-692.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

ON THE PREDICTION OF STRUCTURAL REACTIONS TO BIG EARTHQUAKES IN TURKEY

DOI: 10.17261/Pressacademia.2017.608
PAP-WCTIE-V.5-2017(46)-p.335-340

Guvenc Arslan[1], Diaa Fawzy[2], Coskun Atay[2]
[1] Kirikkale University, Ankara, Turkey
[2] Izmir University of Eonomics, Izmir, Turkey

ABSTRACT

The prediction of structural reactions to big earthquakes is vital in giving warnings for potential damages early enough to minimize losses of life and properties. In the current study we describe buildings by fixed construction and environmental related parameters. Our models are based on real data of damaged buildings collected after the occurrence of three big earthquakes in Turkey. We extend our previous work to include the soil type for damaged buildings. We employ different techniques, namely neural networks (NN) and support vector machines (SVM) to improve the prediction accuracy. The results show that support vector machines, and in particular support vector regression gives better results compared to neural networks. Although we only used averages of soil type for each region, we observed that adding soil type has improved accuracy of predictions for building damages. It is to be noted that these types of predictions are important to ensure the serviceability and safety of existing structures. Our models are vital for the authorities to make fast and reliable decisions and can be also used to improve the development of new constructions codes.

1. INTRODUCTION

Since the formation of the earth, earthquakes have occurred in succession in seismically active areas, resulting in the disappearance of millions of people and buildings. Turkey lies on one of the world's most active regions in terms of earthquakes. Approximately 60,000 people have lost their lives, 125,000 people have been injured and 415,000 buildings have been destroyed in earthquakes in our country during the last 60 years, 96% of our country's lands show the distribution of the first four earthquake zones and 98% of the total population live in these regions. Because of these reasons, we should always be ready to earthquakes and earthquake as an important factor should always be taken into account in settlement planning.

The earthquakes in Turkey located in the Alpine-Himalayan earthquake belt are related to the movement of the African-Arabian plates to the north-northeast direction due to the spreading of the Atlantic Ocean's back toward the two sides. Moreover, the Arabian plate is being pushed to the north and forced to dive under the plate of Eurasia, due to the spread of the sea floor which continues today along the long axis of the Red Sea. Due to this strain, there is an intense jamming effect in the eastern Anatolian region between the Arabian plate and the Eurasian continent. This compression, which has caused major fractures such as the North Anatolian Fault and the Eastern Anatolian Fault, has been going on for millions of years and is the main cause of the earthquakes that we are experiencing today.

The North Anatolian Fault is 1400-1500 km in length. The Anatolian plate between the North Anatolian Fault and the East Anatolian Fault moves westward at a speed of 13-27 mm per year and in the westernmost direction it is curved to the left to the Crete seduction zone.

In November 1999, a devastating earthquake of 7.2 magnitude occurred at the epicenter of Düzce. The earthquake occurred in the direction of East-West between Akyazı-Bolu Tunnel in the direction of Düzce Fault extending parallel to the North Anatolian Fault. In this earthquake, it is understood that the eastern part of the Düzce Fault, which is broken down during the August 17, 1999 earthquake, is broken up, mainly in the 30 km segment in the westernmost part.

In February 2002, a devastating earthquake (USGS Ms: 6.5) struck Afyon in the south east. The earthquake was strongly felt in Konya-Ankara-Eskisehir-Kütahya-Isparta provinces where Sultandağı-Çay-Bolvadin districts of Afyon province were the center. 46 people died in the earthquake, 318 people were injured and 622 buildings were seriously damaged. In large settlements, the most damage occurred in the Çay district, and the Sultandağı and Bolvadin districts also suffered damage. The damage to the rural settlements is concentrated in the villages which are heavily dependent on the provinces of Çay and Sultandağı. The surface fracture developed is faded in the alluvial rust, which consists of marsh at the westernmost end. At the eastern end, the fracture lengths of the bifurcated surface consist of intermittent particles ranging from a few tens of meters to several hundreds of meters and distributed in two main directions. The surface faulting in the Çay district and the parts observed to the east of it were observed in metamorphic foundation rocks. In the Maltepe district, the fracture developed completely in the Pleistocene and Holocene deposits. In the Sultandağı earthquake, although there were large alluvial plains in the fault ceiling block, a few field liquefaction events could be observed. The most prominent example of liquefaction was found in the vicinity of the Çay district of the current bed near the industrial site. Numerous and varied sizes of this area, similar to the liquefaction cone, have observed cone-shaped morphologic sand and gravel pile. The majority of these territories are composed of pebbles and the few are of silt and sandy material. Some of these soil deposits, especially salty material, are notable for their mudflows and offer similarity to liquefaction. It has been reported that the largest structural damage occurred on sandy and marshy grounds where the groundwater level is high.

The vibration-based structural damage detection methods are based on correlating model parameters with the physical properties of the structure. The time history of the measured natural frequency changes can be analyzed and modeled (Friswell and Penny; 1997). Since the natural frequency is a global parameter of the whole structure, these methods can not offer detailed information about local damages. Although the success is shown by the vibration-based methods, they are found to be very sensitive to the measurement noise and to the environmental parameter more than the structural parameters themselves. Housner (1997) has shown that once the data are available for damaged structures, statistical based methods and supervised learning approaches are very promising in detecting damages.

These considerations imply the urgent need for early warning systems. They also show that the ability of predicting future disasters is of great interest to the society. However, predicting future earthquakes is a very complicated task. New methods based on combining shifts in the geomagnetic fields with changes on the ionosphere and shifts of the Earth's crust are found to be very promising (see, e.g., Ruzhin et al. 2014; Duma, G. and Ruzhin, Y. 2003; Davis et al. 1980; Rikitake et al. 1980)

The current work is a continuation of our previous study Fawzy & Arslan (2015), hereafter called paper I. In our previous work we tried to predict the reaction of buildings to big earthquakes, we trained a neural network model with real data of damaged buildings collected after the occurrence of three big earthquakes Afyon (2002; Mw = 6.0), Bingöl (2003; Mw = 6.4) and Düzce (1999; Mw = 7.2). In paper I, we took into account the five basic building parameters, namely, age of the building, number of stories, average floor area, and column area. As an extension of our previous work, we add the average soil type underneath buildings as an extra parameter. The aim of the current study is three folds; (1) extending our previous data set to obtain more statistically reliable results, (2) generating more realistic models by adding the soil type as a parameter, and (3) comparing the results from different mathematical approaches to increase the prediction accuracy. It is also important to mention that the models are based on real data and not on probabilistic models.

2. METHODOLOGY AND THE PREDICTION MODELS

Different approaches for developing models to predict the degree of damages of buildings may be considered. In this study we concentrate on neural networks and support vector machines. One of our main goals is to extend our previous work in Fawzy & Arslan (2015). In this respect we have added more cases to the data set resulting in a total of 58 instances. For a more detailed description of the data set we refer the reader to Fawzy & Arslan (2015). In this study we also added soil type as a parameter, thus increasing the number of parameters.

2.1. Input Parameters

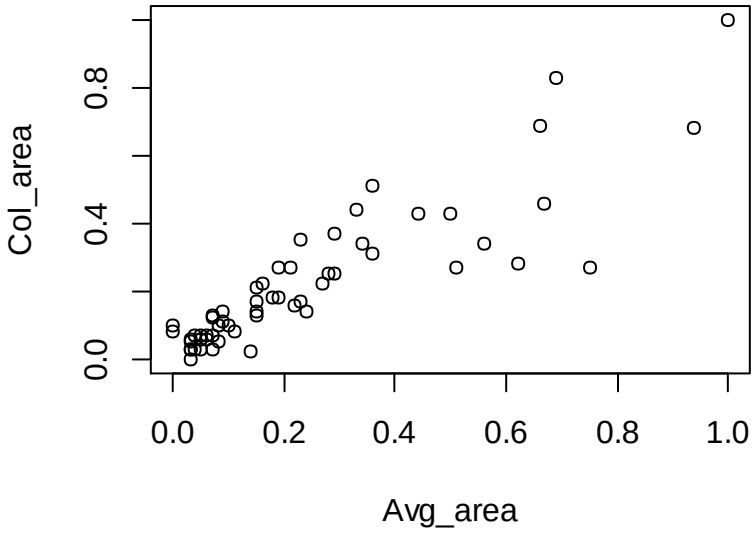
Although there are many factors which affect the degree of damage for buildings we have considered the following basic input parameters in this study:

3. intensity of earthquake,
4. Building type (only reinforced concrete)
5. number of floors,
6. building age,
7. average area,
8. total column area, and
9. soil type.

By using basic data analysis it can be seen that there is a high correlation between the two input parameters “average area” and “column area” (see Fig. 1). Therefore, we decided to exclude “column area” and to include the ratio of average area / column area as a new parameter, which is denoted as “column index”. In this study we also added soil type to our analysis. It is known that soil conditions affect amplitude, frequency and duration of the ground movements (Bol 2012). On the other hand, it is important to mention here that the exact determination of soil type is a difficult task. It may vary drastically even between sub regions of a particular study area. Therefore we decided to use an average soil type indicator for each of the three earthquake regions considered in the current study. The following ordinal variable was used to indicate average soil conditions in the considered regions.

1. Very hard
2. Solid
3. Moderately Solid
4. Loose Ground

Figure 1: Scatter Plot for COL_Area and Avg_Area



2.2 Damage Classification

The calssification of damages we employ for the current study is based on study by Grünthal (1998). Damge levels are divided into the following 5 different categories:

Grade 1: Negligible to slight damage (no structural damage, slight non-structural damage)

Grade 2: Moderate damage (slight structural damage, moderate non-structural damage)

Grade 3: Substantial to heavy damage (moderate structural damage, heavy non-structural damage)

Grade 4: Very heavy damage (heavy structural damage, very heavy non-structural damage)

Grade 5: Destruction (very heavy structural damage)

2.3 Neural Networks (NN)

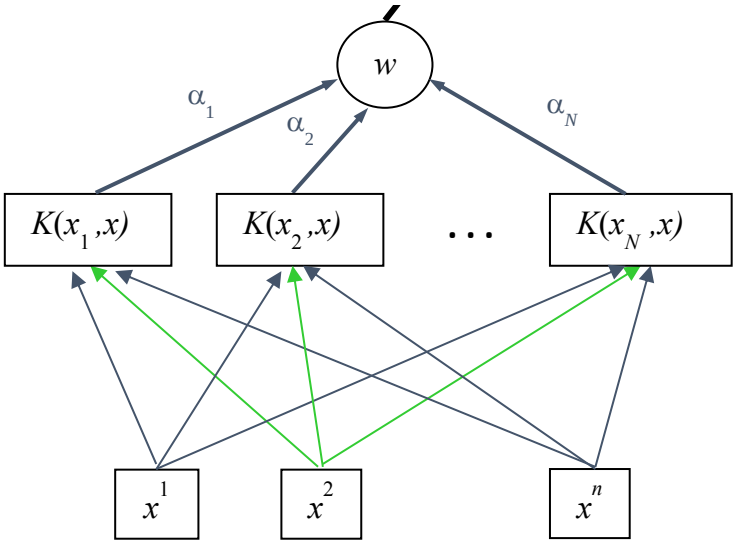
Neural networks learn by training approach similar to their biological counterparts. Although there are different types of learning methods, we concentrate here on the supervised feedforward learning techniques. It should be noted that with enough number of data points used for training of the network and several training cycles an optimum set of weights can be found. The performance function is then computed based on the error between predicted values and the actual cases. For more detailed explanation and information about neural networks the reader is referred to Hornik (1991) and Haykin (1999), among many other references.

2.4 Support Vector Machine (SVM)

Support vector machines (SVM) are one of the state of the art learning algorithms preferred by many researchers in recent years. Classically developed for two-class classification problems and based on the idea of maximum margin classification, they have been extended to more than two-class problems and generalized in various ways. SVMs exploit the idea of mapping the input space into a higher dimensional reproducing kernel Hilbert space in order to obtain a separable space in the new space. Therefore, they are also referred to as kernel machines. SVMs were developed in the framework of statistical learning theory, which was mainly developed by Vapnik and Chervonenkis (1964, 1974, 1998).

The basic idea of SVMs for classification can also be used for regression. The first step, in general, is to map the input variables into a feature space by a particular map ϕ. From a given training set of examples N are selected as support vectors. For any given examples x and x' in the input space the map ϕ produces an inner product, which is denoted by $K(x, x') = \langle\phi(x), \phi(x')\rangle$. The output of the SVM is then obtained as $\Sigma\alpha_i K(x, x_i) + w_0$. For more information on support vector regression the reader is referred to Smola and Schölkopf (2004).

Figure 2: SVM with *n* Inputs and *N* Support Vectors



3. Results and Discussion

In this study we used a data set consisting of 58 training examples representing data of damaged buildings for three major earthquakes in Afyon, Bingöl and Düzce.

Figure 3: Neural Network Model with 2 Hidden Layers

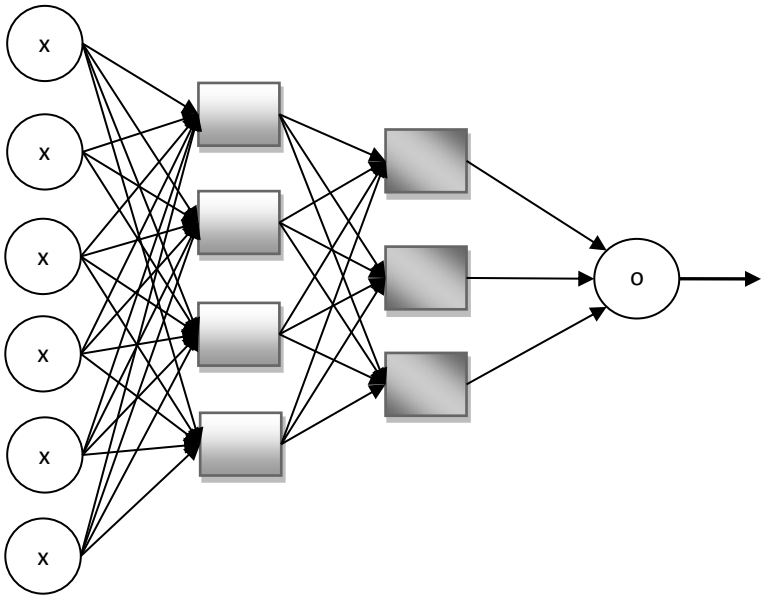
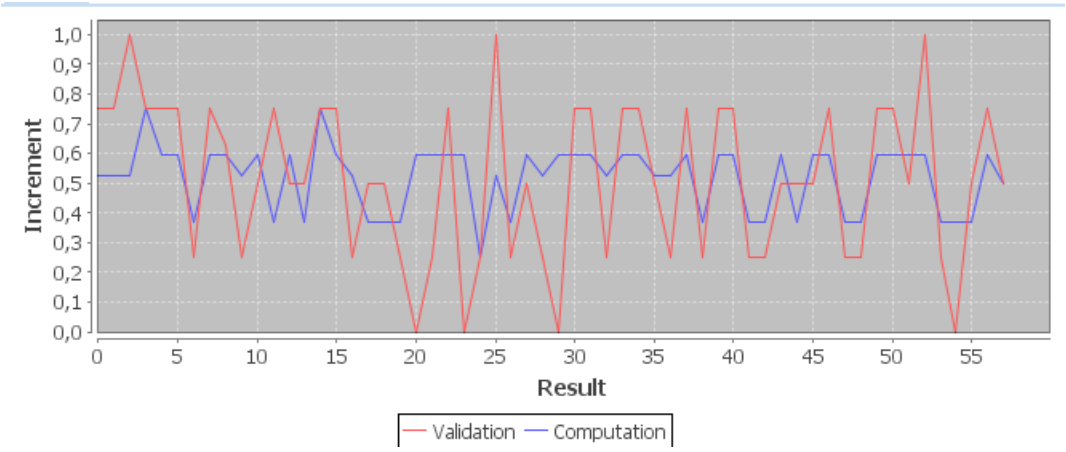


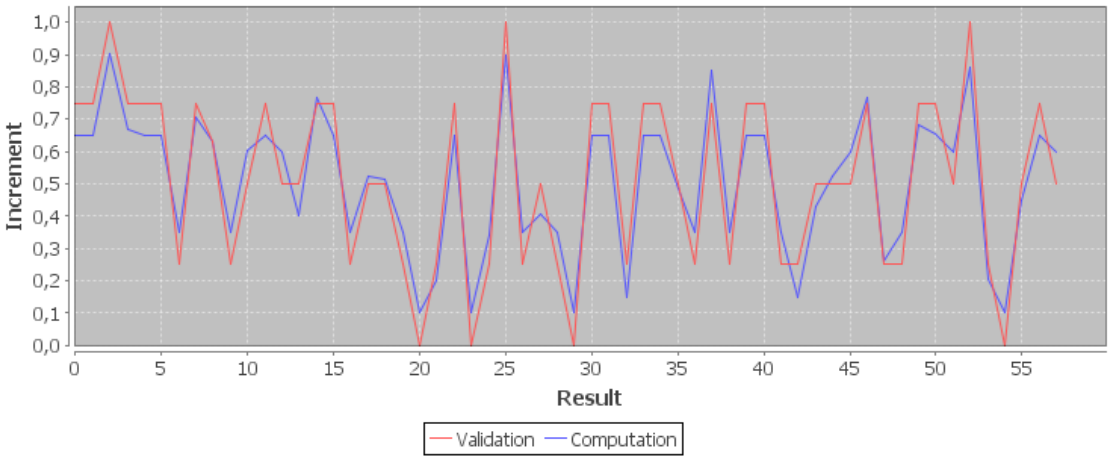
Figure 3 shows the feedforward neural network that we obtained for the earthquake data using Encog software (Heaton, 2015). After trying several two layer models such as five neurons in the first layer and three neurons in the second layer we used the best two layer model structure that we obtained, which is shown in Figure 3. This model has two hidden layers with four and three neurons in the first and second hidden layer, respectively. We note that the training of the neural network takes more time and also gives a training error rate of about 5.75%. The validation chart for this model is given in Figure 4.

Figure 4: Validation Graph for Neural Network Model



The results show that support vector machines, and in particular support vector regression gives better results compared to neural networks. Although we only used averages of soil type for each region, we observed that adding soil type has improved accuracy of predictions for building damages. Figure 5 shows the results of a support vector regression model implemented in the Encog software (Heaton, 2015). The training error rate for this model is less than 1%. We note also that we used the following input parameters in both models: intensity, number of floors, building age, average area, column index and soil type. It can be seen from Figures 4 and 5 that support vector regression clearly gives better results in terms of accuracy. In addition, it should be noted that the support vector regression model should be preferred because there is no need to determine an appropriate network structure. Only an appropriate kernel has to be determined. In this study we used radial basis functions for the kernel type.

Figure 5: Validation Graph for Support Vector Regression



REFERENCES

1. Bol, E., "Determination of the relationship between soil properties and earthquake damage with the aid of neural networks: a case study in Adapazarı", Turkey, Nat. Hazards Earth Syst. Sci 12: 2965–2975, 2012.
2. Davis, Paul M.; Jakson, David D.; Johnston Malcom J.S.: "Further evidence of localized geomagnetic field changes before the 1974 Thanksgiving Day Earthquake, Hollister, California", Geophysical Reasearch Letter, Vol. 7, Issue 7 (513-516), 1980.
3. Duma, G. and Ruzhin, Y.: " Diurnal changes of earthquake activity and geomagnetic Sq-variation", Natural Hazards and Earth Systems Sciences, 3 (171-177), 2003.
4. Fawzy, D.E. and Arslan, G.; "Development of building damage functions for big earthquakes in Turkey", Procedia – Social and Behavioral Sciences, 195 (2290-2297), 2015.
5. Friswell, M.I., Penny, Jet, and Gravey, SD, " A combined gentic and eigensensitivity algorithm for the location of damage in structure", Computers and Structures, 69: 547-556, 1998.
6. Grünthal, G., (ed.). *"European Macroseismic Scale 1998",* Cahiers du Centre Européen de Géodynamique et de Séismologie Volume 15, Luxembourg, 1998
7. Housner, G.W., Bergman, L.A., et al. "Structural control: past, present and future", Journal of Engineering Mechanics, 123 (9): 897-971, 1997.
8. Heaton, J., Encog: "Library of Interchangeable Machine Learning Models for Java and C#", Journal of Machine Learning Research 16: 1243-1247, 2015.
9. Haykin, S.; "Neural Networks: A Comprehensive Foundation", Prentice Hall, 1999.
10. Hornik, K.: "Approximation Capabilities of Multilayer Feedforward Networks, Neural Networks", 4(2), 251–257, 1991
11. MTA (Mineral Research & Exploration General Directorate) mta.gov.tr.
12. Rikitake T., et al.: "Changes in the Geomagnetic Field Associated with Earthquakes in the Izu Peninsula, Japan", Journal of geomagnetism and geoelectricity, Vol. 32, 12 (721-739), 1980.
13. Ruzhin, Yuri; Kamogawa Masashi; Novikov, Victor: "Interrelation of geomagnetic storms and earthquakes: Insight from lab experiments and field observations", 40th COSPAR Scientific Assembly. Held 2-10 August 2014, in Moscow, Russia, Abstract A3.1-68-14
14. Smola, A.J., Schölkopf, B., "A tutorial on support vector regression", Statistics and Computing 14: 199-222, 2004.
15. Vapnik, V., Chervonenkis, A., "A note on one class of perceptrons, Automation and Remote Control", 25, 1964.
16. Vapnik, V., Chervonenkis, A., "Theory of Pattern Recognition" [in Russian]. Nauka, Moscow, 1974. (German Translation: Wapnik W. & Tscherwonenkis A., Theorie der Zeichenerkennung, Akademie-Verlag, Berlin, 1979).
17. Vapnik, V.,"Statistical Learning Theory", John Wiley and Sons, New York, 1998.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

INNOVATIVE SOLAR ENERGY TECHNOLOGY SELECTION USING INTIUTIONISTIC FUZZY SETS

DOI: 10.17261/Pressacademia.2017.609
PAP-WCTIE-V.5-2017(47)-p.341-349

Sezi Cevik Onar[1], Basar Oztaysi[2], Cengiz Kahraman[3]
[1]Istanbul Technical University, Faculty of Management, Industrial Engineering Department, cevikse@itu.edu.tr
[2]Istanbul Technical University, Faculty of Management, Industrial Engineering Department, oztaysib@itu.edu.tr
[3]Istanbul Technical University, Faculty of Management, Industrial Engineering Department, kahramanc@itu.edu.tr

ABSTRACT

Solar energy technologies are mainly composed of photovoltaics and concentrated solar power (solar thermal) technologies. Photovoltaic technology converts sunlight into electricity directly whereas solar thermal technology utilizes its heat. Innovative solar energy technologies appear in the market relatively more often than other energy technologies. Selection among these innovative solar energy technologies is a multi-criteria decision making problem under uncertain environment conditions. Intuitionistic fuzzy sets provide an efficient way of capturing uncertainty by processing both membership and non-membership degrees in a set, whose sum of these degrees is not necessarily equal to one. In this paper, we use an interval-valued intuitionistic fuzzy AHP method for the multi-criteria evaluation of some innovative solar energy technologies. An application of solar energy technology selection is also presented through the proposed method.

Keywords: Solar energy, innovative technology, fuzzy, intuitionistic fuzzy

1. INTRODUCTION

There is an increasing emphasis on renewable sources of energy. Among these renewable energy sources, solar energy is becoming more and more important every day. With the recent developments, it has become one of the main contributors to the total renewable energy installed capacity. Yet, this capacity is still limited when it is compared to the other sources such as fuel or gas. Selecting the appropriate technology can increase the capacity of solar energy.

Solar energy technology selection is a mult-icriteria decision making problem that includes both tangible and intangible criteria which necessities usage of linguistic evaluations. The technology evaluation process contains ambiguity, vagueness and subjectivity in the human judgments.

An intuitionistic fuzzy set has three functions to define it: membership, non-membership, and hesitancy functions. Atanassov and Gargov (1989) proposed interval-valued intuitionistic fuzzy sets (IVIF) where both membership and non-membership functions are defined as intervals. These sets are considered as a generalization of both interval valued fuzzy sets and intuitionistic fuzzy sets.

In this paper, the steps of interval valued intuitionistic fuzzy AHP method will be given and then it is applied to the technology selection for solar energy. In the literature, difficulty of development, operational efficiency of system, difficulty of capacity expansion, supply stability, possibility of replacing oil energy, popularity of use, impact on related industries, pollution impact, scenic impact, development cost, duration of construction, production cost, size of the technology, and annual production are among the most used criteria for the evaluation of solar energy technologies. We make a multi criteria selection among several solar energy technology alternatives under fuzziness.

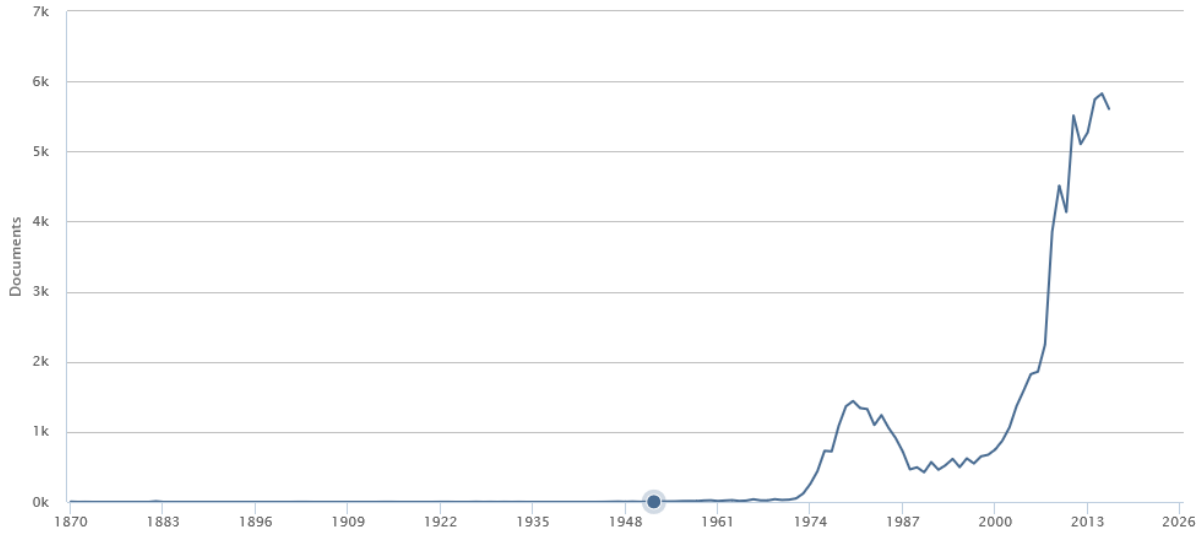
2. LITERATURE REVIEW

The solar power technologies have been developed since 19th century. In 1839, French scientist Edmond Becquerel discovered the potential of photovoltaic effect, which later became the main technique for harnessing solar power (U.S. Department of Energy Energy Efficiency and Renewable Energy). In 1954, first silicon photovoltaic cell has been developed and produced with the efficiency of 4%. Later in 1958 PV panels became the main source of energy for space applications globally. In 1983 worldwide sales of PV cells reached record levels with $250 Million annually. In 1999, cumulative installed capacity of solar power reached 1GW (U.S. Department of Energy Energy Efficiency and Renewable Energy). Total installed capacity of solar energy reached 177GW of photovoltaics and 4.4GW of concentrated solar power systems (Renewable Energy Policy Network for the 21st Century, 2015).

Although solar energy have a long history, in the recent years there is a significant both the generated solar energy capacity and the academic interest on the solar energy. It is possible to understand the year-on-year change for solar energy from the academic databases. A search using the key phrase “Solar Energy” in the “title”, “abstract” or “keywords” has been conducted on 18th of April, 2017 on Scopus database. The results are given in Figure 1.

There are mainly two leading Technologies in solar energy, namely, photovoltaic (PV) and concentrated solar power plants (CSP).

Figure 1. Solar Energy Publications



2.1. Photovoltaic Systems

Photovoltaic panels are working under a principle of creating a negative charge on a special material (i.e. silicon, Si) with the help of radiance from a sun which results in a voltage which then converted to current of electricity. Efficiency of photovoltaic panels are calculated as a ratio of power output from the cell with the total amount of solar radiance captured on the surface of specified area (Arvizu, et al., 2011).

Important characteristic specifications of a photovoltaic solar panel are the material it has been manufactured and the design of the device. According to Direct Solar Energy report, in 2009, 80% of the photovoltaic panels in the market were made with CdTe and Thin Film Si while the rest is made with organic absorber materials (Arvizu, et al., 2011).

Phovoltaic systems are consiting of a PV module and balance of system (BOP) components such as; inverter, storage devices, charge controllers etc. It is possible to classify photovoltaic applications depending on their connection to the grid. Additionally consumer related PV applications are also possible but in much smaller scale (Arvizu, et al., 2011).

2.2. Concentrating Solar Power Systems (CSP)

Concentrating solar power systems create a high beam of sun light to heat a liquid, gas or solid which later being used in thermal electricity generation. Most of the current electricity generation techniques involve heating or burning materials at

some point such as gas, nuclear, biomass or oil. This shows that further improvements in the technology of gas and steam turbines will contribute to creating more viable CSP systems (Arvizu, et al., 2011).

Solar collectors are at the heart of any solar power system and it is possible to categorize collectors into two categories; stationary (non-concentrating) and sun-tracking (concentrating) collectors. As the name implies CSP systems are built upon sun-tracking and concentrating solar collectors. These types of collectors are technologically advanced yet costly when compared to its stationary counterparts as a result offers wide array of advantages such as; higher operating temperature and efficiency. Optical instruments are widely used to provide highly concentrated solar beams to receivers which also adds to the overall cost of the CSP system (Barlev, Vidu, & Stroeve, 2011). Figure 2 shows the criteria used for solar energy technology selection.

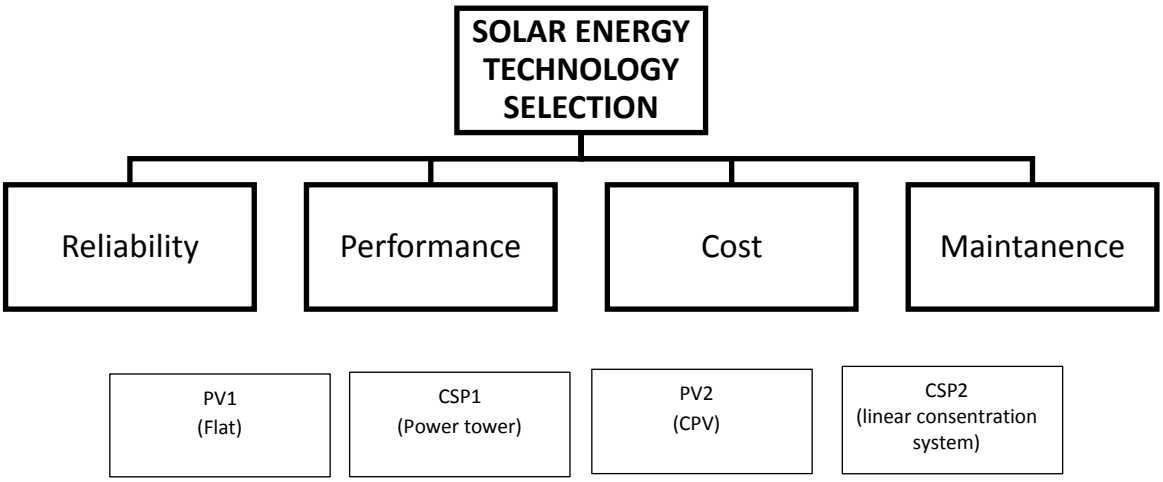


Figure 2: Solar Energy Technology Selection Hierarchy

Reliability covers the issues such as warranty coverage, track record, suitability for environmental conditions such as resistance to extreme weather conditions (Cevik Onar et al., 2015). Performance criterion shows the total energy generation capacity. Cost criterion covers both initial investment costs an operating expenses. The maintanence is the necessity of maintenance activities.

3. DATA AND METHODOLOGY

3.1. Interval Valued Intuitionistic Fuzzy Sets

An interval-valued intuitionistic fuzzy set in $\widetilde{A}$ over X is defined as follows (Atanassov, 1986):

$$\widetilde{A} = \{< x, \mu_{\widetilde{A}}(x), v_{\widetilde{A}}(x) > | x \epsilon X\}, \quad (1)$$

where

$$\mu_{\widetilde{A}} \rightarrow D \subseteq [0,1], v_{\widetilde{A}}(x) \rightarrow D \subseteq [0,1]$$

with the condition $0 \leq \sup\mu_{\widetilde{A}}(x) + \sup v_{\widetilde{A}}(x) \leq 1, \forall x \in X.$

The intervals $\mu_{\widetilde{A}}(x)$ and $v_{\widetilde{A}}(x)$ denote the membership and non-membership functions of the element x to the set $\widetilde{A}$, respectively. Interval-valued Intiutionistic fuzzy set $\widetilde{A}$ is denoted by

$$\widetilde{A} = \left\{< x, \left[\mu_{\widetilde{A}}^{-}(x),\ \mu_{\widetilde{A}}^{+}(x)\right], \left[\upsilon_{\widetilde{A}}^{-}(x),\ \upsilon_{\widetilde{A}}^{+}(x)\right] > | x \in X\right\}, \quad (2)$$

where

$0 \le \mu_{\tilde{A}}^{+}(x) + \upsilon_{\tilde{A}}^{+}(x) \le 1, \qquad \mu_{\tilde{A}}^{-}(x) \ge 0, \upsilon_{\tilde{A}}^{-}(x) \ge 0.$

Eq. (3) defines the hesitancy degree of an interval-valued intuitionistic fuzzy set of $x \in X$ in $\widetilde{A}$.

$$\pi_{\tilde{A}}(x) = 1 - \mu_{\tilde{A}}(x) - v_{\tilde{A}}(x) = \left(\left[1 - \mu_{\tilde{A}}^{+}(x) - \upsilon_{\tilde{A}}^{+}(x)\right], 1 - \mu_{\tilde{A}}^{-}(x) - \upsilon_{\tilde{A}}^{-}(x)\right) \quad (3)$$

For convenience, let $\mu_{\tilde{A}}(x) = \left[\mu_{\tilde{A}}^{-}(x),\ \mu_{\tilde{A}}^{+}(x)\right] = \left[\mu_{\tilde{A}}^{-},\ \mu_{\tilde{A}}^{+}\right]$, $v_{\tilde{A}}(x) = \left[\upsilon_{\tilde{A}}^{-}(x),\ \upsilon_{\tilde{A}}^{+}(x)\right] = \left[\upsilon_{\tilde{A}}^{-},\ \upsilon_{\tilde{A}}^{+}\right]$, so $\widetilde{A} = \left(\left[\mu_{\tilde{A}}^{-},\ \mu_{\tilde{A}}^{+}\right], \left[\upsilon_{\tilde{A}}^{-},\ \upsilon_{\tilde{A}}^{+}\right]\right)$.

Some arithmetic operations with interval-valued intuitionistic fuzzy numbers and $\lambda \ge 0$ are given in the following. Let $\widetilde{A} = ([\mu_{\tilde{A}}{}^{-}, \mu_{\tilde{A}}{}^{+}], [v_{\tilde{A}}{}^{-}, v_{\tilde{A}}{}^{+}])$ and $\widetilde{B} = ([\mu_{\tilde{B}}{}^{-}, \mu_{\tilde{B}}{}^{+}], [v_{\tilde{B}}{}^{-}, v_{\tilde{B}}{}^{+}])$ be two interval-valued intuitionistic fuzzy numbers. Then,

$$\widetilde{A} \oplus \widetilde{B} = \left(\left[\mu_{\tilde{A}}^{-} + \mu_{\tilde{B}}^{-} - \mu_{\tilde{A}}^{-}\mu_{\tilde{B}}^{-}, \mu_{\tilde{A}}^{+} + \mu_{\tilde{B}}^{+} - \mu_{\tilde{A}}^{+}\mu_{\tilde{B}}^{+}\right], \left[v_{\tilde{A}}^{-}v_{\tilde{B}}^{-}, v_{\tilde{A}}^{+}v_{\tilde{B}}^{+}\right]\right) \quad (4)$$

$$\widetilde{A} \otimes \widetilde{B} = \left(\left[\mu_{\tilde{A}}^{-}\mu_{2}^{-}, \mu_{\tilde{A}}^{+}\mu_{2}^{+}\right], \left[v_{\tilde{A}}^{-} + v_{2}^{-} - v_{\tilde{A}}^{-}v_{2}^{-}, v_{\tilde{A}}^{+} + v_{2}^{+} - v_{\tilde{A}}^{+}v_{2}^{+}\right]\right) \quad (5)$$

3.2. Interval Valued Intuitionistic Fuzzy AHP Method

The proposed interval valued intuitionistic fuzzy AHP method is given in the following (modified from Wu et al., 2013).

Firsts step is constructing the fuzzy preference relation matrix composed of linguistic terms assigned from Table 2 for the pairwise comparisons of criteria by each expert.

Table 2: Linguistic Scale and Its Corresponding IVIFS

Linguistic Terms	Membership & Non-membership values
Absolutely Low (AL)	([0.10, 0.25],[0.65, 0.75])
Very Low (VL)	([0.15, 0.30],[0.60, 0.70])
Low (L)	([0.20, 0.35],[0.55, 0.65])
Medium Low (ML)	([0.25, 0.40],[0.50, 0.60])
Equal (E)	([0.50, 0.50],[0.50, 0.50])
Medium High (MH)	([0.50, 0.60],[0.25, 0.40])
High (H)	([0.55, 0.65],[0.20, 0.35])
Very High (VH)	([0.60, 0.70],[0.15, 0.30])
Absolutely High (AH)	([0.65, 0.75],[0.10, 0.25])

Second step is converting the interval-valued intuitionistic fuzzy preference relation matrix on criteria for each expert.

$$\tilde{G}_k = \begin{bmatrix} \tilde{g}_{11}^{(k)} & \cdots & \tilde{g}_{1p}^{(k)} \\ \vdots & \ddots & \vdots \\ \tilde{g}_{p1}^{(k)} & \cdots & \tilde{g}_{pp}^{(k)} \end{bmatrix}, \text{k=1, 2, …, m} \quad (6)$$

where *m* is the number of experts.

Third step is applying the operation in Eq. (6) to aggregate each row of preference relations.

Fourth step is calculating the score judgment matrix $\tilde{S} = \left(\tilde{s}_{ij}\right)_{n\times n}$ and the interval multiplicative matrix $\tilde{A} = \left(\tilde{a}_{ij}\right)_{n\times n}$.

The score judgment matrix of $\widetilde{R}_g$ is represented by the matrix

$\tilde{S} = \left(\tilde{s}_{ij}\right)_{n\times n} = \left[\mu_{g_{ij}}^{-} - v_{g_{ij}}^{+}, \mu_{g_{ij}}^{+} - v_{g_{ij}}^{-}\right]$ as follows:

$$\tilde{S} = \begin{bmatrix} [\mu_{g_{11}}^- - v_{g_{11}}^+, \mu_{g_{11}}^+ - v_{g_{11}}^-] & \cdots & [\mu_{g_{1n}}^- - v_{g_{1n}}^+, \mu_{g_{1n}}^+ - v_{g_{1n}}^-] \\ \vdots & \ddots & \vdots \\ [\mu_{g_{n1}}^- - v_{g_{n1}}^+, \mu_{g_{n1}}^+ - v_{g_{n1}}^-] & \cdots & [\mu_{g_{nn}}^- - v_{g_{nn}}^+, \mu_{g_{nn}}^+ - v_{g_{nn}}^-] \end{bmatrix} \quad (7)$$

The interval multiplicative matrix $\tilde{A} = (\tilde{a}_{ij})_{n\times n} = \left[10^{\left(\mu_{g_{ij}}^- - v_{g_{ij}}^+\right)}, 10^{\left(\mu_{g_{ij}}^+ - v_{g_{ij}}^-\right)}\right]$ is given as follows.

$$\tilde{A} = \begin{bmatrix} \left[10^{(\mu_{g_{11}}^- - v_{g_{11}}^+)}, 10^{((\mu_{g_{1n}}^+ - v_{1n}^-))}\right] & \cdots & \left[10^{\left(\mu_{g_{1j}}^- - v_{g_{1j}}^+\right)}, 10^{\left(\mu_{g_{1j}}^+ - v_{g_{1j}}^-\right)}\right] \\ \vdots & \ddots & \vdots \\ \left[10^{(\mu_{g_{n1}}^- - v_{g_{n1}}^+)}, 10^{((\mu_{g_{n1}}^+ - v_{n1}^-))}\right] & \cdots & \left[10^{(\mu_{g_{nn}}^- - v_{g_{nn}}^+)}, 10^{(\mu_{g_{nn}}^+ - v_{g_{nn}}^-)}\right] \end{bmatrix} = \begin{bmatrix} [\tilde{a}_{11}^-,\ \tilde{a}_{11}^+] & \cdots & [\tilde{a}_{1n}^-,\ \tilde{a}_{1n}^+] \\ \vdots & \ddots & \vdots \\ [\tilde{a}_{n1}^-,\ \tilde{a}_{n1}^+] & \cdots & [\tilde{a}_{nn}^-,\ \tilde{a}_{nn}^+] \end{bmatrix} \quad (8)$$

Thus, the score judgment matrix $\tilde{S} = (\tilde{s}_{ij})_{n\times n}$ is converted to the matrix $\tilde{A}$ whose values are between 0 and 10.

Fifth step is determining the priority vector of the interval multiplicative matrix $\tilde{A} = (\tilde{a}_{ij})_{n\times n}$ by calculating the $\widetilde{w}_i$ interval for each criterion using Eq. (9)

$$\widetilde{w}_i = \left[\frac{\sum_{j=1}^n \tilde{a}_{ij}^-}{\sum_{i=1}^n \sum_{j=1}^n \tilde{a}_{ij}^+}, \frac{\sum_{j=1}^n \tilde{a}_{ij}^+}{\sum_{i=1}^n \sum_{j=1}^n \tilde{a}_{ij}^-}\right] = [w_i{}^-, w_i{}^+], i = 1, \dots, n \quad (9)$$

Sixth step is constructing the possibility degree matrix $P = (p_{ij})_{n\times n}$ by comparing the obtained weights in Step 5. To do this, use Eq. (10):

$$P(w_i \geq w_j) = p_{ij} = \frac{min\left\{L_{w_i} + L_{w_j}, max(w_i{}^+ - w_j{}^-, 0)\right\}}{L_{w_i} + L_{w_j}} \quad (10)$$

where $L_{w_i} = w_i^+ - w_i^-$ and $L_{w_j} = w_j^+ - w_j^-$ and $p_{ij} \geq 0$, $p_{ij} + p_{ji} = 1$, $p_{ii} = 1/2$.

Seventh step is prioritizing the $P = (p_{ij})_{n\times n}$ by Eq. (11):

$$w_i = \frac{1}{n}\left[\sum_{j=1}^n p_{ij} + \frac{n}{2} - 1\right] \quad (11)$$

Eighth step is normalizing the weights vector obtained in Step 7 and obtain the normalized weights w_i^T of the alternatives, as given by Eq. (12):

$$w_i^T = \frac{w_i}{\sum_{i=1}^n w_i} \quad (12)$$

Ninth step is obtaining the compromised decision matrix from experts, composed of alternatives performances with respect to criteria utilizing Table 2.

$$\widetilde{D}_k = \begin{bmatrix} \tilde{d}_{11} & \cdots & \tilde{d}_{1p} \\ \vdots & \ddots & \vdots \\ \tilde{d}_{p1} & \cdots & \tilde{d}_{pp} \end{bmatrix} \quad (13)$$

Then, generate the weighted decision matrix (P) by using the weights of criteria obtained in Step 8.

$$\tilde{P} = \widetilde{D}_k \times\ w_c \quad (14)$$

Obtain the total scores by summing the values in each row and select the alternative with the defuzzified largest score.

Last step is applying sensitivity analyses to check the robustness of the decision against the small changes in the values of problem parameters.

4. FINDINGS AND DISCUSSIONS

The proposed R&D strategy selection model has been applied to solar energy technology selection in Turkey. The planned capacity of the energy generation facility is three megawatts and the land is located in the central Anatolia. There are four solar energy technology alternatives. Two of them are CSP systems. First CSP system is a power tower system and the second CSP system is a linear concentration system. Similarly, two types of photo voltatic (PV) system is under

considiration. The first PV system is the classical flat-plate system wheras the second PV system is a concentrator photovoltaics (CPV) where lenses and mirrors are used to focus sunlight.

Three experts in the company have individually evaluated the main criteria with respect to the goal, which is to select the best solar energy strategy. These evaluation matrices are given in Table 3.

Table 3: Evaluation Matrices for the Main Criteria

Criteria	Reliability	Performance	Cost	Maintenance
Reliability	EE	MH	ML	VH
Performance		EE	ML	VH
Cost			EE	AH
Maintenance				EE

Criteria	Reliability	Performance	Cost	Maintenance
Reliability	EE	MH	E	H
Performance		EE	MH	VH
Cost			EE	VH
Maintenance				EE

Criteria	Reliability	Performance	Cost	Maintenance
Reliability	EE	ML	ML	EE
Performance		EE	ML	VH
Cost			EE	AH
Maintenance				EE

The evaluations for solar energy technologies with respect to the main criteria are presented in Table 4.

Table 4: Evaluations for Solar Energy Technologies with Respect to the Main Criteria

wrt Reliability

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	MH	H	VH
PV2		E	MH	MH
CSP1			E	MH
CSP2				E

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	E	MH	MH
PV2		E	MH	MH
CSP1			E	E
CSP2				E

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	VH	H	VH
PV2		E	L	ML
CSP1			E	H
CSP2				E

wrt Performance

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	ML	ML	ML
PV2		E	ML	ML
CSP1			E	MH
CSP2				E

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	ML	L	L
PV2		E	ML	ML
CSP1			E	E
CSP2				E

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	VL	L	AL
PV2		E	MH	L
CSP1			E	L
CSP2				E

wrt Cost

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	VH	MH	AH
PV2		E	L	ML
CSP1			E	MH
CSP2				E

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	MH	MH	VH
PV2		E	E	MH
CSP1			E	E
CSP2				E

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	MH	VH	AH
PV2		E	MH	MH

wrt maintenance

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	MH	H	H
PV2		E	MH	MH
CSP1			E	VH
CSP2				E

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	MH	MH	MH
PV2		E	MH	E
CSP1			E	E
CSP2				E

Criteria	PV1	PV2	CSP1	CSP2
PV1	E	MH	H	H
PV2		E	H	E

CSP1			E	E
CSP2				E

CSP1			E	ML
CSP2				E

Table 5 summarizes the aggregated TIFN evaluations and the defuzzified weights of the main criteria.

Table 5: Aggregated TIFN Evaluations and the Defuzzified Weights

	Reliability	Performance	Cost	Maintenance	weight
Reliability	(0.5,0.5,0.5,0.5)	(0.43,0.54,0.31,0.46)	(0.34,0.44,0.5,0.56)	(0.55,0.63,0.25,0.37)	0.266
Performance	(0.34,0.48,0.4,0.52)	(0.5,0.5,0.5,0.5)	(0.34,0.48,0.4,0.52)	(0.6,0.7,0.15,0.3)	0.287
Cost	(0.5,0.57,0.31,0.43)	(0.43,0.54,0.31,0.46)	(0.5,0.5,0.5,0.5)	(0.65,0.73,0.11,0.27)	0.322
Maintenance	(0.3,0.39,0.55,0.61)	(0.15,0.3,0.6,0.7)	(0.12,0.27,0.63,0.73)	(0.5,0.5,0.5,0.5)	0.125

Cost is the most important criteria whereas maintenance is the least important criteria for the solar energy investment decision. Table 6 gives the scores of solar energy technologies with respect to the main criteria. The total score of the alternatives are also given in this Table. According to the results, although it shows the lowest performance classical flat-plate system (PV1) gets the highest score due to its cost advantage. A decrease in the cost of the other technologies will have a significant impact on the solar energy technology selection.

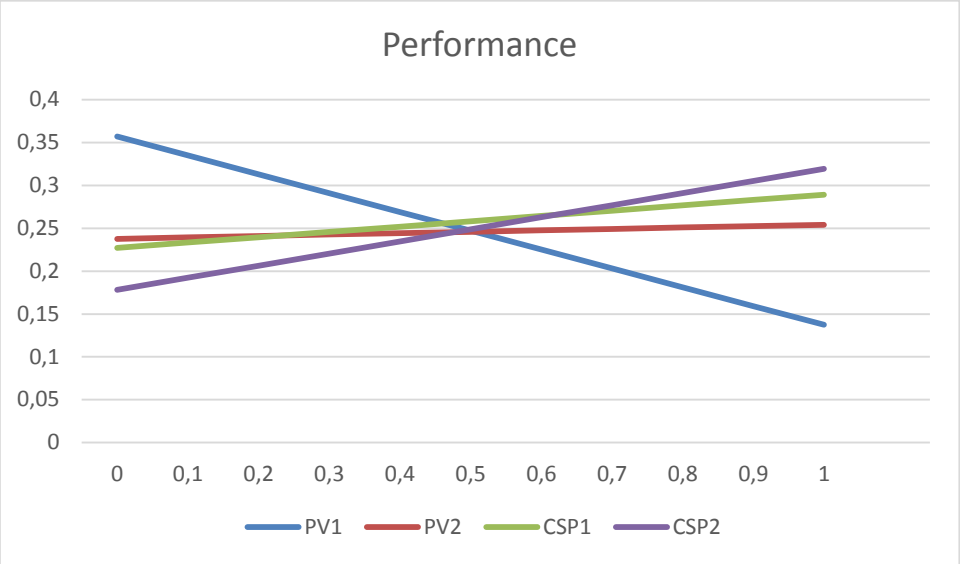
Table 6: Scores of Solar Energy Technologies with respect to the Criteria

Main Criteria	Alternatives	Local score	Global score
Reliability (0.266)	PV1	0.347	0.093
	PV2	0.240	0.064
	CSP1	0.239	0.064
	CSP2	0.174	0.046
Performance (0.287)	PV1	0.137	0.039
	PV2	0.254	0.073
	CSP1	0.289	0.083
	CSP2	0.319	0.092
Cost (0.304)	PV1	0.369	0.119
	PV2	0.224	0.072
	CSP1	0.227	0.073
	CSP2	0.180	0.058
Maintenance (0.125)	PV1	0.346	0.043
	PV2	0.270	0.034
	CSP1	0.202	0.025
	CSP2	0.181	0.023

	Total Score
PV1	0.294
PV2	0.242
CSP1	0.245
CSP2	0.219

In order to check the robustness of the model, we conduct a sensitivity analysis based on the *performance* main criteria. Figure 3 illustrates the results of the sensitivity analysis. This shows that increasing the importance of the performance criteria changes the results. When the performance becomes very important CSP2 is a better alternative for this area.

Figure 3: Sensitivity Analysis with respect to Performance Criteria Weight



5. CONCLUSION

Innovative solar energy technologies should be evaluated by using a multi-criteria decision making method, which can be handle linguistic evaluations under uncertainty. Innovativeness requires new criteria to be used in the evaluation process and causes a harder evaluation process because of uncertainty conditions of new technologies. The proposed method can consider these uncertainties through intuitionistic fuzzy sets. AHP based IVIF multi-criteria method successfully evaluated the solar energy technologies. Consistency of the pairwise comparison matrices and sensitivity of the given decisions have been examined. The obtained results are consistent and robust.

For further research, other extensions of fuzzy sets such as hesitant fuzzy sets, type-2 fuzzy sets, or neutrosophic fuzzy sets can be used in the proposed method for comparison purposes.

REFERENCES

Atanassov K.T. (1986), Intuitionistic Fuzzy Sets, Fuzzy Sets and Systems, 20, 87-96, Elsevier.

U.S. Department of Energy Energy Efficiency and Renewable Energy. (n.d.). *The History of Solar.*

Renewable Energy Policy Network for the 21st Century. (2015). Renewables 2015, Global Status Report. Paris.

U.S. Department of Energy Energy Efficiency and Renewable Energy. (n.d.). The History of Solar.

Arvizu, D., Balaya , P., Luisa, C. F., Hollands, K. T., Jager-Walnau, A., Kondo, M., . . . Zilles, R. (2011). Special Report on Renewable Energy Resources and Climate Change Mitigation. Intergovernmental Panel on Climate Change. Cambridge University Press.

Barlev, D., Vidu, R., & Stroeve, P. (2011, May). Innovation in concentrated solar power. Solar Energy Materials & Solar Cells, 95, 2703-2725.

Wu, J., Huang, H.-B., and Cao,Q.-W. (2013). Research on AHP with interval-valued intuitionistic fuzzy sets and its application in multi-criteria decision making problems, Applied Mathematical Modelling, 37, 9898–9906.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

ACADEMIC ENTREPRENEURIAL INTENTIONS IN SCIENCE AND ENGINEERING: VALIDATION OF THE SCALE IN TURKEY

DOI: 10.17261/Pressacademia.2017.610
PAP-WCTIE-V.5-2017(48)-p.350-358

Ufuk Gur[1]
[1]Yeditepe University, ufukozgulgur@outlook.com

ABSTRACT

Academic entrepreneurship is defined as entrepreneurial activities engaged by faculty in order to commercialize research results and deliver societal and economic benefits. It is closely related to the "Entrepreneurial University" which extends the mission of universities beyond teaching and research. Higher education institutions and their affiliated faculty members in sciences and engineering disciplines are important actors of innovation ecosystems. This study investigates academic entrepreneurial intentions distinguished by soft level intentions (industry collaboration) and hard level intentions (spin-off formation). The data was collected between December 2015, and March 2016 from a national online survey of Turkish academics in science and engineering faculties at approximately 90 universities of which 402 full responses were gathered. Major methodology included multivariate analysis technique namely as Structural Equation Modelling. Within the scope of this study, the constructs of academic entrepreneurship intentions were created and applied comprehensively in sciences and engineering disciplines in Turkey for the first time.

Keywords: Academic entrepreneurship, entrepreneurial university, technology transfer, innovation, entrepreneurship

1. INTRODUCTION

Academic entrepreneurship is the knowledge phenomenon of the new economy in which scientific excellence is transformed into commercialized assets for societal and economic benefits. Economic value created at universities is the focus of economic development policies; universities are recognized as the regional engines of innovation and growth providing new technologies and business ventures (Laukkanen, 2003). Sciences and engineering disciplines form knowledge intensive industries and it requires a closer integration of university and industry with the aim of using specific high technology-capital intensive infrastructures and graduate level credentials of human capital from higher education institutions. Firms have emerged usually through the licensing option however the new paradigm of entrepreneurial universities and entrepreneur academics bring a new dimension to the industry. The emergence of knowledge-intensive entrepreneurship is dependent on the active involvement of scientists in this respect (Jain, George, & Maltarich, 2009). Many corporations in technology business had their origins as TLO start-ups including Genentech in biotechnology, Cirrus Logic in semiconductors, and Lycos in internet search engines representing an important mechanism for technology transfer transformed into economic activity (Di Gregorio & Shane, 2003).

2. LITERATURE REVIEW

2.1. Definitions of Academic Entrepreneur

Academic entrepreneur can narrowly be defined as the faculty at a university who creates a new organization and bring his or her innovation/invention/solution to market as a commercial opportunity (see Table 1). Spin-off activity at universities in the form of business ventures based on academic research (Shane, 2004) is a concept which came out with the systematic

analysis of MIT model by Roberts (1991). Actually "*spinoff behavior is a reflection of individial actions and therefore is largely due to personality, ability, and willingness of the individual to engage successfully in entrepreneurial behavior*."(O'Shea, Allen, O'Gorman, & Roche, 2004, p. 16). Samson and Gurdon (1993) defines academic entrepreneur as a lecturer or researcher affiliated with a university while pursuing a role in a start-up venture.

Table 1: Definitions of Academic Entrepreneur

Louis et al, 1989	where academic entrepreneurship is defined as "the attempt to increase individual or institutional profit, influence, or prestige through the development and marketing of research ideas or research based products
Roberts, 1991	the founding of a new company by a researcher who previously worked at a laboratory or academic department where the technology originated
Samsom and Gurdon (1993)	an academic whose primary occupation, prior to playing a role in a venture start-up, and possibly concurrent with that process, was that of a lecturer or researcher affiliated with a Higher Education Institute.
Radosevich (1995)	inventor–entrepreneurs who are or were laboratory employees and who actively seek to commercialize their own inventions, surrogate–entrepreneurs who are not the inventors but who acquire rights to the federally–sponsored technology.
Dickson et al, 1998	Academic entrepreneur with entrepreneurial aspirations in addition to academic work; the entrepreneurial scientist who is full time involved in a business venture dedicating to scientific interests, scientific entrepreneurs who are involved in a firm both dedicated to business and scientific interests.
Etzkowitz, 1998	the entrepreneurial scientist more broadly as someone with "an entrepreneurial perspective in which results are scanned for their commercial as well as intellectual potential"
Shane, 2004	A new company founded to exploit a piece of intellectual property created in an academic institution
Murray, 2004	Entrepreneurial activities by academics are complex and can vary in "the range from limited interaction, through extensive research collaboration at formal and informal levels, to scientists as fully-fledge entrepreneurial founders
Perkmann and Walsh (2007)	Development and commercial exploitation of technologies pursued by academic inventors through a company they (partly) own (spin-off companies).
Jain et al., 2009	Any form of technology transfer which has some potential commercial benefit can be defined as academic entrepreneurship.
Gurau et al., 2012	• academic manager/entrepreneur mainly responsible for founding/leading the venture and day to day management, • academic project manager responsible for spesific scientific projects in existing firms and • academic scientific advisor of one or more firms, namely as the most senior one relative to the previous roles, mainly responsible for scientific advice and mapping opportunities.
Abreu and Grinevich, 2013	In addition to previous, Non commercial activities This category includes providing informal advice, giving public lectures, organising exhibitions, and publishing books for a general audience.
Perkmann et al, 2013	a sub-output of "academic engagement" which is wider than commercial exploitation of a spesific technology.

Dickson and others claimed in 1998 (Gurău, Dana, & Lasch, 2012; Lundqvist & Williams Middleton, 2013) that academic entrepreneurs possess three different profiles: academic entrepreneur with entrepreneurial endeavors in addition to academic work; the entrepreneurial scientist who is full time involved in a business venture dedicating to scientific interests, scientific entrepreneurs who is involved in a firm both dedicated to business and scientific interests. Gurău et al. (2012) in their study, identified three main categories of academic entrepreneurship based on the level of involvement in managerial and scientific roles; academic manager/entrepreneur mainly responsible for founding/leading the venture and day to day management, academic project manager responsible for spesific scientific projects in existing firms and academic scientific advisor of one or more firms, namely as the most senior one relative to the previous roles, mainly responsible for scientific advice and mapping opportunities. All roles network with scientific and business community.

Klofsten and Jones-Evans (2000) defined the academic entrepreneurship activities, in addition to teaching and research roles of faculty, as following: large scale science projects obtained through public grants or industry support; contracted research for external organizations, sales of consulting for scientific or technological expertise; patenting and licensing research results to industry; formation of new firms exploiting university research; teaching to non-university based individuals and organizations; commercial sales of products developed in the university; provision of testing and calibration facilities to non-university based individuals and organizations. This broad definition extends the role of academic entrepreneurship to a new level not limited to firm formation. Louis, Blumenthal, Gluck, and Stoto (1989) also defined academic entrepreneurship in a similar form: large-scale science obtained through funds, consulting for knowledge, soliciting funds from industry, patenting and firm formation. Philpott, Dooley, O'Reilly, and Lupton (2011) defined the forms of academic entrepreneurship as following adopting from Jones-Evans and Louis et al.: creation of a technology park, spin-off formation, patenting and licensing, contract research, industry training courses, consulting, grantsmanship, publishing academic results, producing highly qualified graduates in contrast to Samsom and Gurdon (1993) limited definition of taking role in a venture start-up.

The process model of academic entrepreneurship by Wood (2011) argues whether academic entrepreneurship is initiated by the efforts of a technology transfer office (TTO), concluding only if the faculty are actively interested in commercialization or they have to do so by the policies. Academic entrepreneurship can be traced back to formation of research laboratories to obtaining funding for future research in a resource limited environment. Louis et al. (1989) had referred to patenting as an interest in commercial applications of research however academic entrepreneurship is not necessarily quantifiable and it can occur at many levels (Rasmussen, Moen, & Gulbrandsen, 2006). Knockaert, Foo, Erikson, and Cools (2015) referred to academic entrepreneurship in a broader sense including the overall patenting and licensing activity and university industry collaboration. Perkmann et al. (2013) defined academic entrepreneurship as a sub-output of "academic engagement" which is wider than commercial exploitation of a spesific technology. One can claim that academic engagement is correlated with scientific productivity yet it is a question whether engagement is an antecedent of commercialization behaviour.

We can define the "academic entrepreneur" as a faculty at a university who brings his or her innovation/invention/solution to market or society as a commercial or non-commercial opportunity. Academic entrepreneurship is one of the channels through which scientific knowledge reaches the market. The "Entrepreneurial scientist" discovers the frontiers of knowledge and transfer them into commercial gains, societal benefits as well as intangible benefits such as recognition and prestige. "Entrepreneurial scientist"[1] (Etzkowitz, 2013) can be attached to the notion of academic entrepreneurship as the potential entrepreneur may engage in wealth creation and prestige seeking behavior while transferring the polyvalent[2] knowledge and science into practical and financial business (Etzkowitz, 2013; Etzkowitz & Viale, 2010).

In broader sense, as Abreu and Grinevich (2013) suggested, academic entrepreneurship should be extended to the overall set of activities beyond licensing, patents and spinoffs. However spinouts are more common in life sciences due to the nature of product development with the long time horizon of market entry. In social sciences consultancy and contract research are more common. They have conceptually framed academic entrepreneurial activities as; formal commercial activities including licensing and spinoffs, informal commercial activities beyond patent including consultancy, contract research, joint research projects; non-commercial activities providing informal advice, giving public lectures, organizing exhibitions, and publishing books for general audience.

Academic entrepreneurial styles can be adopted according to different degrees of involvement developed by Etzkowitz (2013):

- Direct interest in the formation of a spin-off firm and leading role.

[1] "Seeking tangible rewards as well as prestige – weaving knowledge, money and power into a single framework – entrepreneurial scientists are creating a new and potent element in the ethos of science." (Etzkowitz, 2013)

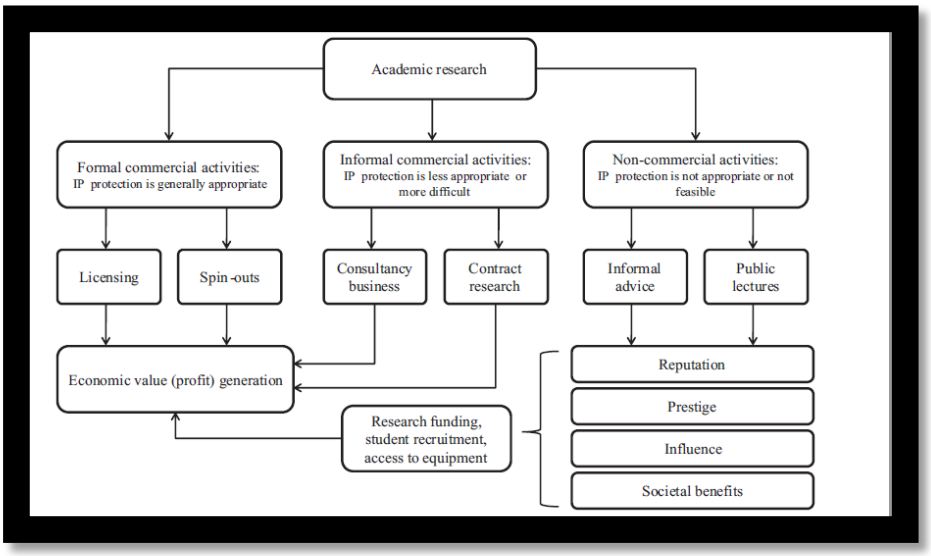
[2] Etzkowitz and Viale (2010) explains polyvalent nature of knowledge being both theoritical and practical, both publishable and patentable

- Interest in commercialization of discoveries and supporting role as a member of scientific advisory board.
- Indirect interest in economic implications of research and handing over the role to technology transfer offices.
- No interest in firm formation but supporting firm formation for advancement in research.

2.2 Academic Entrepreneurial Intentions

With regards to the Theory of Planned Behavior by Ajzen (1991) and Armitage and Conner (2001), the best predictive determinant of entrepreneurial activity is the "entrepreneurial intention". Within the scope of academic entrepreneurship, academic entrepreneurial intention can be defined as the intention for getting involved in entrepreneurial activities including formal commercial activities, informal commercial activities, and on-commercial activities adressed by Abreu and Grinevich (2013) (see Figure 1).

Figure 1: Academic Entrepreneurship Activities, Adopted from Abreu, M., & Grinevich, V. (2013). The nature of academic entrepreneurship in the UK: Widening the focus on entrepreneurial activities.



Research Policy, 42(2), 408-422. doi: 10.1016/j.respol.2012.10.005

The available measurement constructs for entrepreneurial intentions include Entrepreneurial Decision Scale (Chen, Greene, & Crick, 1998), Entrepreneurial Intentions Scale (Krueger, Reilly, & Carsrud, 2000), Individual Entrepreneurial Intent Scale (Thompson, 2009), and Entrepreneurship Intentions Questionnaire (Liñán & Chen, 2009). Although those constructs are sufficient for analyzing adults' and students' intentions for entrepreneurship, the notion of academic entrepreneurial intentions require a narrow focus on specific activities such as intellectual property creation within the context of universities. Academic entrepreneurial intentions have been analyzed by Prodan and Drnovsek (2010) previously. However, the construct did not again provide the necessary narrow focus to target faculty. The most promising novel contribution was provided by the study of Huyghe and Knockaert (2014) (see Table 2).

Table 2: Academic Entrepreneurial Intentions

Huyghe and Knockaert (2014)	How likely is it that, in the foreseeable future, You will engage in the founding of a university spin-off?, You will engage in the establishment of a company based upon an idea and/or technology developed at the university?, and You will participate in the founding of a firm to commercialize your research? You will apply for a patent resulting from your research at the university?,

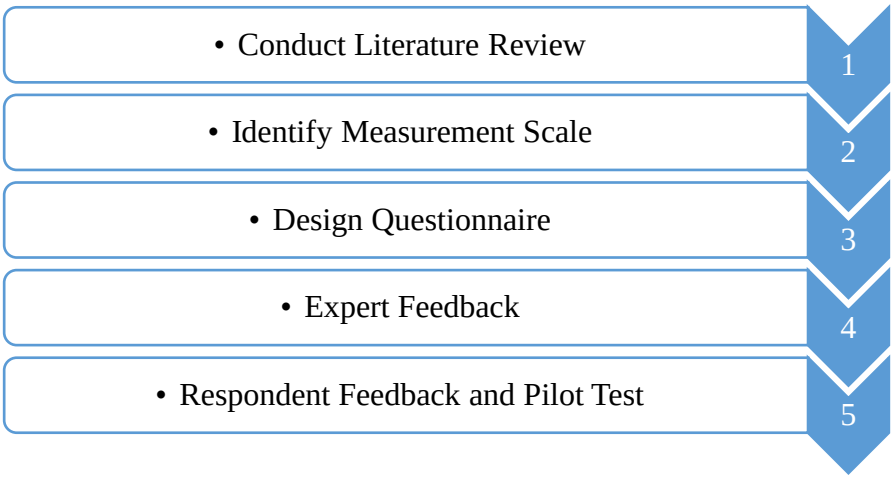
	You will license some of your technological developments to the industry?, and You will become the owner of intellectual property rights (patent, copyright, trademark,…)? You will engage in collaborative research with industry? and You will engage in contract research or consulting activities with industry?

Aiming to offer a comprehensive measurement construct for academic entrepreneurial intentions within Turkish setting, adopted version of Huyghe and Knockaert was used and additionally, the soft side of the academic entrepreneurial activities such as publishing, placement of students in the industry, attendance to scientific meetings, and training in the industry was included as adressed by De Silva (2012).

3. DATA AND METHODOLOGY

As a second generation multi-variate analysis technique, Structural Equation Modelling method was adopted for the purpose of this study which relies on theory based testing of a measurement model including multiple variables (Hair, Black, Babin, & Anderson, 2010). Exploratory factor analysis results have also been provided in order to assess construct and factor reliability.Based on currently available literature, and previous research, it was decided to develop a new construct for the purpose of the specific research study based on the scale development principles and steps suggested by Karakoç and Dönmez (2014) (Figure 2).

Figure 2: Principles of Scale Development



Entrepreneurial intent constructs in entrepreneurship studies focus on self-efficacy, entrepreneurial and proactiveness abilities of general adults however a more specific construct needs to be designed for targeting academics of whom entrepreneurial activity is more different than entrepreneurial trajections of adult careers. In order to be responsive to the objectives of this specific research, a more comprehensive sets of items were generated based on several constructs and theory used in previous studies. Initial item generation has been completed based on those constructs and theoretical foundations.

Questionnaires in the form of measurement instruments have been developed following the effective questionnaire development procedures adressed by Peterson (2000).Secondly, expert feedback has been retrieved from following respondents and relevant adjustments have been made on items accordingly:

-PhD Candidate with MBA in Knowledge and Innovation Management

-Professor of Entrepreneurship

-Professor of Entrepreneurial Finance

-Top Executive at Technology Transfer function

Following expert feedback, respondent reviews have been conducted by actual respondents for item evaluation in terms of objectivity, ethical considerations, relevance and cognitive penetrability. The questionnaire scale was designed on likert type scale[3] in order to measure probability of action towards the topic of interest with graded responses on each statement.

Based on suggested adjustments, final versions of the constructs have been used in web survey tool. Web-based survey method has been adopted as the respondent profile is qualified in terms of education and familiarity with technology. Furthermore, internet based surveys provide decreased costs, and faster response rates (Reynolds, 2006).

4. FINDINGS AND DISCUSSIONS

4.1 Exploratory Factor Analysis

Based on the analysis with the field data, final results for KMO and Barlett' s test showed that the construct is appropriate for conducting exploratory factor analysis. The analysis resulted in three dimensions explaining 64.20% of total variance with cronbach alpha reliability of ,910 for all items (Table 3). Principle components analysis has been conducted and factor based reliability analysis are shown below (Table 4).

Table 3: KMO and Bartlett' s Test

KMO	Barlett' s	Sig.	Dimensions	Variance Explained	Cronbach' s Alpha
,893	2262,880	,000	• Soft Entrepreneurial Intentions • Hard Entrepreneurial Intentions	64,20%	,910

Table 4: Factor Loadings from Principal Component Factor Analysis with Varimax Rotation for a Two-Factor Solution for Academic Entrepreneurial Intentions Questionnaire (n=404)

Factors / Items	Soft Entrepreneurial Intentions α=0,882	Hard Entrepreneurial Intentions α=0,875
109	0,789	
111	0,785	
108	0,760	
113	0,745	
110	0,726	
112	0,688	
103		0,861
104		0,847
102		0,806
106		0,622
105		0,621
107		0,511

[3] A Likert scale is a summated rating scale used for measuring attitudes. The method was developed by Rensis Likert in 1932.

Note. Loadings <0.40 are omitted.

4.2 Face and Content Validity

As one of the sources of validity evidence (Cook & Beckman, 2006), face and content validity is related to the construct' s ability to measure the intended topic. Starting from item generation, every step must be taken carefully in order to prove the construct' s face and content validity. As in similar studies (Axler, 2015; Kilian, Schubert, & Bjørn-Andersen, 2015), in-depth review of literature, pre-tests with experts and respondents, and relevant modifications ensured face and content validity of the constructs.

4.3 Goodness of Fit

Hair et al. (2010) framed the rules of thumb for structural equation modelling starting with measurement model specifications. In order to test structural model hypotheses, goodness of fit indices of measurement model constructs should meet criteria values. As Hair et al. (2010) suggested fundamental measures of goodness of fit indices may represent chi-Square, degree of freedom, statistical significance of chi square, RMSEA as one type of absolute indices; Normed Fit Index (NFI), Tucker Lewis Index (TFI), Comparitive Fit Index (CFI), or Relative non-centrality index (RNI) as one type of incremenal fit indices, and Adjusted Goodness of fit index (AGFI) or Parsimony Normed Fit İndex (PNFI) as part of parsimony fit indices. They claimed that reporting chi square, degrees of freedom, RMSEA, CFI or TLI, provide sufficient evidence to prove a model' s acceptability. Var 112 loaded below 0,50 threshold value and was eliminated. Variables 103 and 102 were eliminated due to cognitive bias potential with extreme covariation between error terms. Final model showed goodness of fit with no more than three modifications (Table 5).

Table 5: Goodness of Fit Indices for Academic Entrepreneurial Intentions Construct

Models	χ^2	df	χ^2/ df	RMSEA	CFI	sig
Basic Model	535,381	53	10,102	,150	,839	,000
Model 1 (var 112 eliminated)	487,641	43	11,340	,160	,845	,000
Model 2 (var 103 eliminated)	195,329	34	5,745	,109	,933	,000
Model 3 (Var 102 eliminated)	89,841	26	3,455	,078	,971	,000

4.4 Convergent Validity

As a means of testing construct validity, additional to confirmatory factor analysis with goodness of fit (GOF) indices, factor loadings, composite reliability[4] (CR) and average variance extracted[5] (AVE) are reported for convergent validity (Fornell & Larcker, 1981). Results for analyses were reported below.

Table 6: Factor Loadings, AVE and CR

Construct	**Items**	**Factor Loading (>0.50)**	**AVE (>0.45)**	**CR (>0.70)**
Soft Entrepreneurial Intentions	var110	0,663	0,62	0.890
	var113	0,733		
	var108	0,879		
	var111	0,759		
	var109	0,884		
Hard Entrepreneurial Intentions	var107	0,732	0,62	0.865
	var105	0,849		
	var106	0,929		
	var104	0,604		

4.5 Discriminant Validity

Fornell and Larcker (1981) reported that discriminate validity exists when the level of square root of AVE is greater than the intercorrelations between constructs. Hair et al. (2010) suggested that intercorrelations between the constructs below 0.90

[4] The automated formula on the link was used for composite reliability calculations using factor loadings of the AMOS output: http://www.thestatisticalmind.com/calculators/comprel/comprel.htm

[5] AVE was calculated based on the Formula; total of square factor loadings divided by number of items of the latent variable.

value indicate no multicollinearity. In absence of sufficient discriminant validity and in the presence of multi-collinearity issues, some scholars suggested (Cohen, Cohen, West, & Aiken, 2013; Farrell, 2010) that excluding collinear variables from the model is a solution for enhancing model' s validity. Results are reported below for the constructs.

Table 7: Discriminant Validity

	Soft Entrepreneurial Intentions	Hard Entrepreneurial Intentions
Soft Entrepreneurial Intentions	√AVE 0,78	
Hard Entrepreneurial Intentions	0,708	√AVE 0,78

Note. *Diagonal elements report the square root AVE and other matrix entries report the correlation estimation between them.

5. CONCLUSION

Within the scope of this study, the constructs of academic entrepreneurship intentions were created and applied comprehensively in sciences and engineering disciplines in Turkey for the first time. The results have shown that academic entrepreneurial intentions can be analyzed on soft and hard levels. Soft academic entrepreneurial intentions include a wide range of activities such as publishing and research collaboration with industry. Hard academic entrepreneurial intentions are based on spin-off formation or co-foundation. The construct can be used for analyzing science and engineering faculty' s entrepreneurial intentions at entrepreneurial university settings for further research.

REFERENCES

Abreu, M., & Grinevich, V. (2013). The nature of academic entrepreneurship in the UK: Widening the focus on entrepreneurial activities. *Research Policy, 42*(2), 408-422. doi:10.1016/j.respol.2012.10.005

Axler, R. E. (2015). *Commercialization, Collaboration and Conflict of Interest: An Institutional Work Analysis of Academic Entrepreneurship in Canada.* University of Toronto.

Chen, C. C., Greene, P. G., & Crick, A. (1998). Does entrepreneurial self-efficacy distinguish entrepreneurs from managers? *Journal of Business Venturing, 13*(4), 295-316.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*: Routledge.

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: theory and application. *The American journal of medicine, 119*(2), 166. e167-166. e116.

De Silva, L. (2012). Academic entrepreneurship in a resource constrained environment.

Di Gregorio, D., & Shane, S. (2003). Why do some universities generate more start-ups than others? *Research Policy, 32*(2), 209-227.

Etzkowitz, H. (2013). Anatomy of the entrepreneurial university. *Social Science Information, 52*(3), 486-511. doi:10.1177/0539018413485832

Etzkowitz, H., & Viale, R. (2010). Polyvalent Knowledge and the Entrepreneurial University: A Third Academic Revolution? *Critical Sociology, 36*(4), 595-609. doi:10.1177/0896920510365921

Farrell, A. M. (2010). Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu (2009). *Journal of Business Research, 63*(3), 324-327.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 39-50.

Gurău, C., Dana, L. P., & Lasch, F. (2012). Academic entrepreneurship in UK biotechnology firms. *Journal of Enterprising Communities: People and Places in the Global Economy, 6*(2), 154-168. doi:10.1108/17506201211228958

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). Multivariate Data Analysis Seventh Edition Prentice Hall.

Huyghe, A., & Knockaert, M. (2014). The influence of organizational culture and climate on entrepreneurial intentions among research scientists. *The Journal of Technology Transfer, 40*(1), 138-160. doi:10.1007/s10961-014-9333-3

Jain, S., George, G., & Maltarich, M. (2009). Academics or entrepreneurs? Investigating role identity modification of university scientists involved in commercialization activity. *Research policy, 38*(6), 922-935.

Karakoç, F. Y., & Dönmez, L. (2014). Ölçek Geliştirme Çalışmalarında Temel İlkeler. *Tıp Eğitimi Dünyası, 40*(40).

Kilian, T., Schubert, P., & Bjørn-Andersen, N. (2015). *Benefits and Barriers of University Industry Collaborations from a Researcher's Perspective: Development of Formative Scales and Cluster Analysis.* Paper presented at the 23rd European Conference on Information Systems (ECIS) 2015.

Klofsten, M., & Jones-Evans, D. (2000). Comparing Academic Entrepreneurship in Europe – The Case of Sweden and Ireland. *Small Business Economics, 14*(4), 299-309. doi:10.1023/A:1008184601282

Knockaert, M., Foo, M. D., Erikson, T., & Cools, E. (2015). Growth intentions among research scientists: A cognitive style perspective. *Technovation, 38*, 64-74. doi:10.1016/j.technovation.2014.12.001

Krueger, N. F., Reilly, M. D., & Carsrud, A. L. (2000). Competing models of entrepreneurial intentions. *Journal of Business Venturing, 15*(5), 411-432.

Laukkanen, M. (2003). Exploring academic entrepreneurship: drivers and tensions of university-based business. *Journal of Small Business and Enterprise Development, 10*(4), 372-382. doi:10.1108/14626000310504684

Liñán, F., & Chen, Y. W. (2009). Development and Cross-Cultural application of a specific instrument to measure entrepreneurial intentions. *Entrepreneurship Theory and Practice, 33*(3), 593-617.

Louis, K. S., Blumenthal, D., Gluck, M. E., & Stoto, M. A. (1989). Entrepreneurs in Academe: An Exploration of Behaviors among Life Scientists. *Administrative Science Quarterly, 34*(1), 110-131. doi:10.2307/2392988

Lundqvist, M. A., & Williams Middleton, K. L. (2013). Academic entrepreneurship revisited-university scientists and venture creation. *Journal of Small Business and Enterprise Development, 20*(3), 603-617.

O'Shea, R., Allen, T. J., O'Gorman, C., & Roche, F. (2004). Universities and Technology Transfer: A Review of Academic Entrepreneurship Literature. *Irish Journal of Management, 25*(2).

Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., . . . Sobrero, M. (2013). Academic engagement and commercialisation: A review of the literature on university–industry relations. *Research Policy, 42*(2), 423-442. doi:10.1016/j.respol.2012.09.007

Perkmann, M., & Walsh, K. (2007). University–industry relationships and open innovation: Towards a research agenda. *International Journal of Management Reviews, 9*(4), 259-280.

Peterson, R. A. (2000). *Constructing effective questionnaires* (Vol. 1): Sage Publications Thousand Oaks, CA.

Philpott, K., Dooley, L., O'Reilly, C., & Lupton, G. (2011). The entrepreneurial university: Examining the underlying academic tensions. *Technovation, 31*(4), 161-170. doi:10.1016/j.technovation.2010.12.003

Prodan, I., & Drnovsek, M. (2010). Conceptualizing academic-entrepreneurial intentions: An empirical test. *Technovation, 30*(5-6), 332-347. doi:10.1016/j.technovation.2010.02.002

Radosevich, R. (1995). A model for entrepreneurial spin–offs from public technology sources. *International Journal of Technology Management, 10*(7), 879-893.

Rasmussen, E., Moen, Ø., & Gulbrandsen, M. (2006). Initiatives to promote commercialization of university knowledge. *Technovation, 26*(4), 518-533. doi:http://dx.doi.org/10.1016/j.technovation.2004.11.005

Reynolds, R. A. (2006). *Handbook of research on electronic surveys and measurements*: IGI Global.

Roberts, E. B. (1991). *Entrepreneurs in high technology*: New York: Oxford University Press.

Samsom, K. J., & Gurdon, M. A. (1993). University scientists as entrepreneurs: a special case of technology transfer and high-tech venturing. *Technovation, 13*(2), 63-71.

Shane, S. A. (2004). *Academic entrepreneurship: University spinoffs and wealth creation*: Edward Elgar Publishing.

Thompson, E. R. (2009). Individual entrepreneurial intent: Construct clarification and development of an internationally reliable metric. *Entrepreneurship Theory and Practice, 33*(3), 669-694.

Wood, M. S. (2011). A process model of academic entrepreneurship. *Business Horizons, 54*(2), 153-161. doi:10.1016/j.bushor.2010.11.004

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

SIMULATION OF THE RADIO-INDUIT DAMAGES BY PHOTONS ON THE DNA MOLECULE IN ITS ENVIRONMENT

DOI: 10.17261/Pressacademia.2017.611
PAP-WCTIE-V.5-2017(49)-p.359-367

H.Oudira [1]; D. Lotfi [1]; A.Saifi [2]
[1]Faculty of Technology, Department of Electronics University Mouhamed Boudiaf of M'Sila , Algeria. Oudira_houcine@yahoo.fr
[2]Faculty of Technology, Department of Electronics University Mouhamed Boudiaf of M'Sila , Algeria. lotfi_djouane@yahoo.fr
[3]Faculty of Engineering Department of Electronics University of Constantine 1 , Algeria. Saifi19@yahoo.fr

ABSTRACT

Monte Carlo codes and equation of diffusion were successively, used to simulate the physical, and physico-chemical, and chemical steps that to estimate the simple and double strand breaks produced on a linear and hydrated DNA model by an incident photons beam. The increase of the biological effectiveness, between 150 and 550 eV, is mainly due to the variation of photon cross section upon the K-ionisation threshold in O,N,C atoms of DNA. In this study we present yields of main radiolysis species (e^-_{aq}, H, OH) as function of time, then parameterised by the concentration of solute such formicate ($HCOO^-$). Moreover, we show that, with the same number of absorbed photons, the evolution of breaks as a function of energy reaches a maximum around 450 eV.

Keywords : DNA, Monte Carlo, diffusion, yields, solute, photons.

1. INTRODUCTION

The interaction of a mono-energy photon beam with matter induces point energy deposition. The latter then participate in the formation of a group of radicals which, after diffusion and reaction with the molecules of biological interest (cytosine, adenine, guanine, thymine), cause lesions and consequently genetic alterations in the Molecule of DNA [1]. We consider in this work:

- A break simple stand (C.S.B) is recorded when a direct ionization takes place on the grouping sweetens
- A break stand-in stand (CDB) is defined by a rupture simple sprig on every chain of the stand-in helix at the inside of a maximum distance of 10 pairs of cousin [2].

The direct effect of the photon incidents are analyzed by the methods of Monte Carlo step by step [3], while the indirect effect is treated by of the equation of diffusion to calculate the concentrations and the radiochemical yields of the different radiolysis species. Although it is less precise, it has the double advantage of reducing computation time and memory space [4].

2. MODEL DESCRIPTION

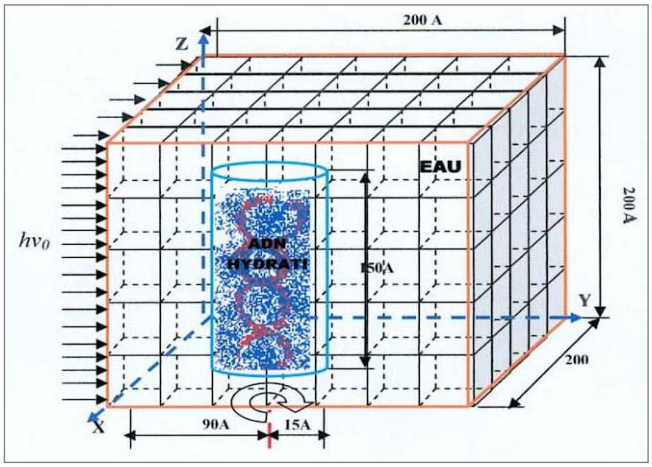
To simulate the biologic environment, we consider linear DNA fragment of B type constituted of 41 pairs of bases, whose properties are summarized in the table 1 [5].

Table 1: Helical Properties of the DNA

Description of the fragment of DNA [5]	large size, END
Sense of rotation of the helix	right
Angle of rotation of a tour (in degree)	36
Axis of the helix	The pairs of basis
Big furrow	large
Small furrow	narrow
Pas(nm)	0.338

The hydration envelope is adjusted with the help of modelling molecular Biosym software [6]. The whole is relaxed by the molecular minimization algorithm using the method of the biggest slope of the system potential energy, and inserted in the centre of a three-dimensional grid containing the water bulk; and the formate concentration varying between 0 and 2 moles by liter of water. The incidental photons are uniformly issued and perpendicularly to a plane surface, to a distance of 8 nm of the DNA center, privileging the direct effect of the irradiation thus in this range of energy (150-550 eV) as it can noted in figure 1. It is necessary to note that at the time of simulation of the ionization and excitation processes, it is taken account of the heterogeneity of the middle as well as the diffusion of the radiolysis spices created to the neighbourhood of the volume targets.

Figure1: Geometry of the Irradiation of the Fragment of DNA



3. MATIMATICAL MODELS

3.1. Stochastic Model of Monte Carlo

The Monte Carlo method of the type step by step is adapted perfectly to the study of the physical and physic-chemical evaluation phases of the energy deposited in the target volume, to the extent that we consider the path λ that an electron can traverse in a heterogeneous milieu and the probability of interaction with a given entity.

$$\lambda = -\overline{\lambda(E)}.\log R \qquad (1)$$

$\overline{\lambda(E)}$ is the range and is given by the relation:

$$\overline{\lambda(E)} = \frac{1}{\sum_{ij} N_i \sigma_{ij}(E)} \qquad (2)$$

Where: R is an arbitrary number equally distributed between 0 and 1.

N_i is the number of atoms or molecules of the type *i* by unit of volume.

$\sigma_{ij}(E)$ is the effective section of interaction of the type *j* on the atom or the molecule of the type *i* for an incidental particle of energy *E.*

$$P_i = \sigma_i(E).N_i.\overline{\lambda(E)} \qquad (3)$$

P_i is the probability for that an incidental particle interacts with an atom or a molecule of the *i* component, knowing that there is interaction.

If there are k_i types of different interactions from the incidental particle with the *i* component, the *j* type of interaction having an effective section total $\sigma_{ij}(E)$ by atom or by molecule, the probability so that the interaction is of the type *j*, knowing that the interaction takes place with the *i* component, is thus：

$$P_{ij} = \frac{\sigma_{ij}(E)}{\sum_{l=1}^{k} \sigma_{il}(E)} \quad (4)$$

During simulation, the point of interaction being fixed, we will have an interaction with the *i* component if *i* verifies:

$$\sum_{l=1}^{i-1} P_l < R_1 \leq \sum_{l=1}^{i} P_l \quad (5)$$

and the interaction will be of the type j on the *i* component if *j* verifies:

$$\sum_{l=1}^{j-1} P_{il} < R_2 \leq \sum_{l=1}^{j} P_{il} \quad (6)$$

R_1 and R_2 are arbitrary numbers equally distributed between 0 and 1.

At the end of the physical stage (10^{-15} Second) the irradiated molecules are in an excited or ionized state [7] whose fate is drawn by probabilities indicated above.

at 10^{-12} second, the electronic trace produced, in its neighbor, the chemical species such as：

e_{aq}^{-} ,H, OH, H2, H2 O2, and H_3O^+.

With regard to, the cross sections of biological interest molecules [8], we adopted the correction introduced by Laverne and Pimblott [9].

$$\sigma_{DNA} = \frac{A_{DNA}}{A_{water}} . \sigma_{water} \quad (7)$$

Where A_{DNA} and A_{water} are respectively the molar masses of the DNA and water.

Each atom of the DNA is differentiated from the other atoms by introduction of an empirical grandeur known as ray of Van Der Waals (RVDW). The total cross section σ_i of an atom of VDW RVDW's ray is written:

$$\sigma_i(E) = \frac{R_{VDW}^i}{\sum_{Base} R_{VDW}^i} \frac{A_{Base}}{A_{water}} \sigma_{xater} \quad (8)$$

Thus, we exploited the rays of VDW presented by Burkert [10] for the simulation of the electrons transport in the biological environment.

3.2. Diffusion Equation

To reduce the calculating time machine, relating to the study of the chemical phase, we chose a deterministic approach based on the resolution of the diffusion equation [11].

$$\frac{\partial C_i}{\partial t} = D_i \nabla^2 C_i + S \quad (9)$$

- *Ci* is the concentration of the species of the type *i* (mole dm^{-3})

- $D_i \nabla^2 C_i$ is the product of the Laplacian of C_i by the constant of diffusion D_i ($cm^2\ s^{-1}$) of the species of the type *i*;

- S is the term source (mole $dm^{-3}\ s^{-1}$), considered as the algebraic sum of terms representing, following the chemical reaction taken in consideration, the impoverishment or the production of the species of the type *i*:

$$S = \left(\sum_w k_{iw} C_w + \sum_l \sum_j k_{lj} C_l C_j \right) - \left(k_i C_i + \sum_j k_{ij} C_i C_j \right) \quad (10)$$

$k_{iw} C_w$, $k_{lj} C_l C_j$ represent respectively the production of species *i* following the dissociation of w, and the reaction of the species *l* with the species *j*. $k_i C_i$ and $k_{lj} C_l C_j$ represent respectively the impoverishment of species *i* following its dissociation, and its reaction with the particles of the type *j*; k_{iw} k_i, k_{lj} and k_{ij} indicate the constants of reaction's speed; k_{iw} and k_i are expressed in s^{-1}, while k_{lj} and k_{ij} are in dm3 $mole^{-1}\ s^{-1}$.

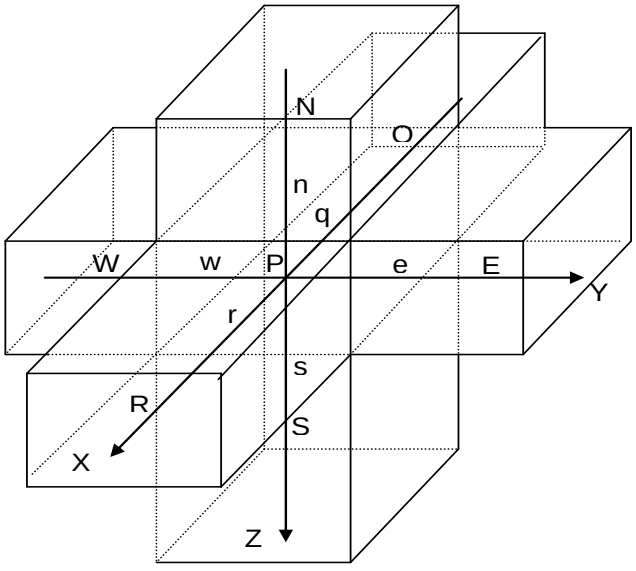
3.2.1 Discretizing the Diffusion Equation

To solve the equation (9), it is necessary to discretize it, that is to say, to replace it by simple algebraic equations expressing the same physical information that can be solved numerically. For this, converts the equation (9) in an integral equation by applying the weighted residuals method:

$$\iiint \int \frac{\partial C_i}{\partial t} W(X,Y,Z,t) dXdYdZdt = \iiint \int \left(D_i \nabla^2 C_i + F \right) W(X,Y,Z,t) dXdYdZdt \quad (11)$$

The choice of the test function W (X, Y, Z, t), and integration of equation (11) cause algebraic equations such that the concentration C_i in a point P (C_{iP} noted), the center of a elementary volume within the computational domain, is connected to those of the points E, W, R, Q, S, and N center of neighboring elementary volumes considered (Figure 2).

Figure 2: Immediate Neighboring of an Elementary Volume Inside the Three-Dimensional Grid



To access this formulation we have used the continuous finite element method [12], [13] and the method of centered differences [10]. Thus, the explicit method can replace the equation (11) by the following algebraic equation:

$$A_P C_{iP} = A_E C_{iE} + A_W C_{iW} + A_R C_{iR} + A_Q C_{iQ} + A_S C_{iS} + A_N C_{iN} + K \quad (12)$$

3.2.2. Calculation of Equation (12) Coefficients:

For a point P situated inside the three-dimensional grid, integrating (11) on the following ranges:

$$\left[t\,;\,t+\Delta t\right],\left[X_P-\frac{\Delta X}{2}\,;\,X_P+\frac{\Delta X}{2}\right],\left[Y_P-\frac{\Delta Y}{2}\,;\,Y_P+\frac{\Delta Y}{2}\right],\left[Z_P-\frac{\Delta Z}{2}\,;\,Z_P+\frac{\Delta Z}{2}\right]$$

provides the coefficients A_P, A_E, A_W, A_R, A_Q, A_S, A_N and K relating to the equation (12) and whose values:

$$A_E = D_i\frac{\Delta X\Delta Z}{\Delta Y} = A_W$$

$$A_R = D_i\frac{\Delta Y\Delta Z}{\Delta X} = A_Q$$

$$A_S = D_i\frac{\Delta X\Delta Y}{\Delta Z} = A_N$$

$$A_P = \frac{\Delta X\Delta Y\Delta Z}{\Delta t} + A_E + A_W + A_R + A_Q + A_S + A_N$$

$$k = \frac{\Delta X\Delta Y\Delta Z}{\Delta t}C_{iP}^0 + F_P^0\Delta X\Delta Y\Delta Z$$

ΔX , ΔY, ΔZ are spaces step in the three directions of the three-dimensional grid.

Δt is the time after which the changes in the concentration of a given radiolytic species are evaluated. The boundary conditions on the extreme points of the domain used to associate their coefficients whose value varies according to the position of the elementary considered volume [10]. Following integration of equation (11) on all the elementary volumes of space concentrations, we obtain another matrix system by block, whose resolution by the method of generalized Thomas matrix calculation (called TDMA) is immediate [14]. Also, we give a detailed attention to the calculation of the space step Δx whose value is reported with the temporal step Δt to ensure the stability of the solutions of the equation of diffusion, to respect the electronic balance of the milieu as much as that are possible. This relation uses the constant of diffusion D of the fastest particle $\left(H_{aq}^{+}\right)$.

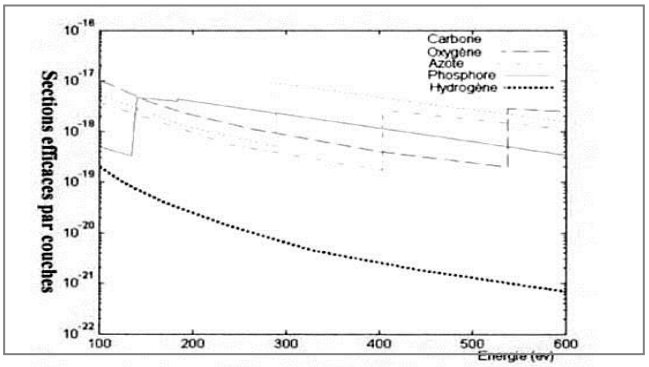
$$\left(\Delta x\right)^2 \geq \sqrt{4D}.\Delta t \qquad (13)$$

4. DATA OF THE PROBLEM TO BE TREATED

4.1. Photoelectric Effective Sections

In the range of considered energy (150-550 eV), almost 99.5% of the interactions are photoelectric [8]; also, at the time of the calculation of the free course and the different probabilities of interaction, we used the efficient sections established by Cullen [15], whose representation in Figure3 is corrected for the mean atomic composition of each element in a nucleotide of the nucleus, DNA.

Figure 3: Photoelectric Efficient Sections Expressed by Nucleotide



The middle atomic composition of our fragment of DNA est : Carbon 10, Oxygen 6, Nitrogen 4, Phosphor 1, hydrogen 11,.

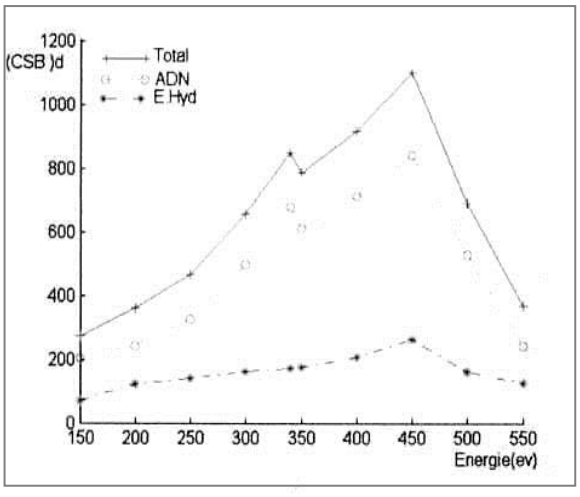
5. RESULT AND DISCUSSION

In this study, we were interested on the one hand, to the photoionizations produced on the carbon, nitrogen and oxygen atoms in view of the important proportions that they represent in the composition of the DNA and its envelope, On the other hand, to the outcome of the radicals created and their contributions to the formation of single and double-strand breaks as a function of energy.

5.1 Direct Effect

Between 0 and 600 eV, the free mean path is less than 1 nm. This results lead to close interactions at the level of DNA and its hydration envelope. Indeed, as we can see in figure 4, between 250 eV and 340 eV the number of direct CSBs goes from one to two times of its initial value. This is due to the predominance of the number of carbon atoms per nucleotide, and to their cross-sections raised at this energy level. Moreover, between 350 eV and 450 eV, the effect of the ionization threshold-K of the conjugated nitrogen at a relatively large proportion of atoms per nucleotide results in the creation of a maximum of CSB. Below the ionization threshold of carbon (290 eV) and above the oxygen threshold (540 eV), the photoionization is mainly done with liquid water, whose relatively small cross-sections confirm the number Restricted fractures at the level of the curve shown in figure 4.

Figure 4: Simple Strand Breaks (CSB)d According to the Energy



The photoelectrons and Auger electrons released in the surrounding water are the source of multiple ionizations leading to the formation of the H, OH and e^-_{aq} radicals which indirectly react with the DNA subunits and its envelope of hydration to produce single and double strand breaks (Figure 5).

The analysis of figure5 shows that the single-strand breaks are greater in the hydration layer compared to those recorded at the DNA level, which has inspired us to undertake the study of the yields of the three main spices as a function of time (Figure 6).

It is deduced that the more the number of radicals is, the more the DNA subunits are attacked for a concentration Γ and Cs low. The maximum protection of the DNA is reached for Γ = 20 molecules of crystal water per nucleotide and Cs = 0.7 liter / mol of formate solute. Finally, we compare our results with those obtained by a purely stochastic method (Figure 7). An acceptable agreement between the breaks recorded by the two approaches is evident.

Figure 5: Indirect Simple Strand Breaks (CSB)i According to the Energy

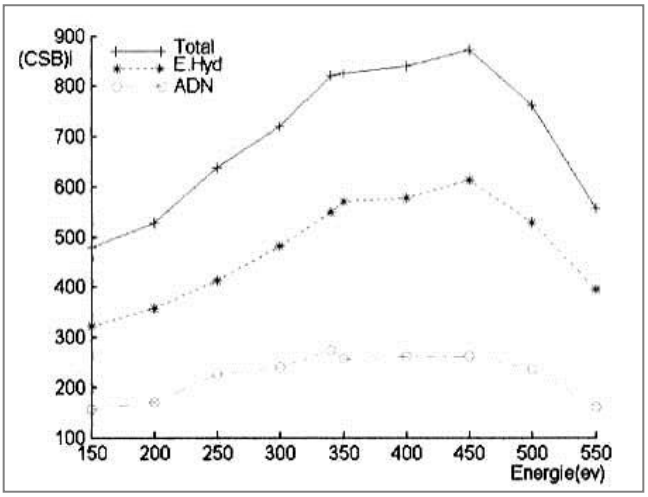


Figure 6: Radiochemical Yields of the Main Radiolytic Species as a Function of Time for Photons Energy of 290 eV

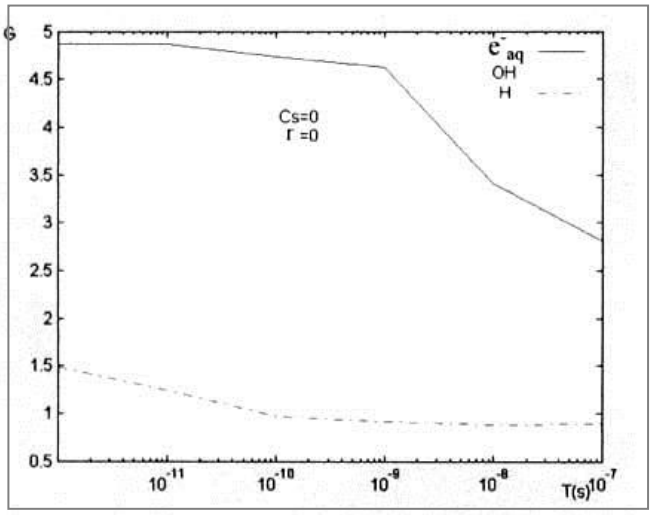
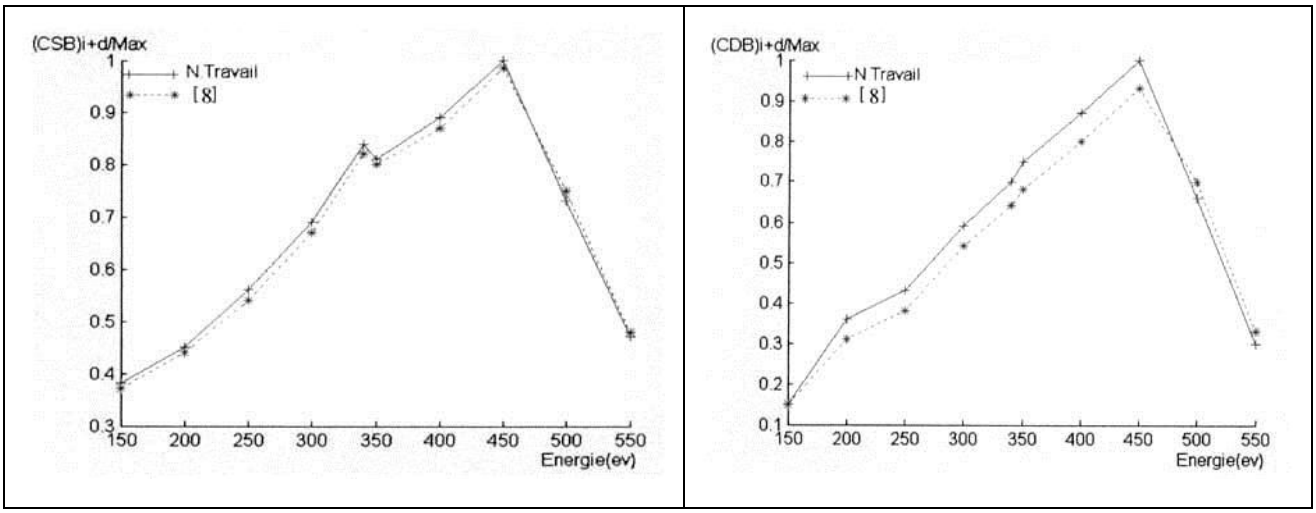


Figure 7: Normalized variations in relation to the peak value, of the damages recorded on a linear DNA model to 10^{-8} s for 2000 absorbed photons. Comparison with the results of a merely stochastic model.



5. CONCLUSION AND PERSPECTIVES

Our conclusion to the evaluation of radio-induced damage at the level of the DNA molecule as a linear model is distinguished from those carried out numerically by Vrigneaud [8], in that it uses a deterministic method to study the Indirect effect of irradiation on the biological environment. This has the consequence in the elaborated software, the taking into account a maximum numbre of chemical reactions, a finest spatial step, while retaining the advantage of a lower calculation time. Moreover, the computer memory space helping, the extension of this method to the study of a more elaborate model such as the nucleosome is conceivable to approach a little more to the actual course of events.

REFERENCES

[1] - Saifi A.., Abdelmoumne M.E.K, Selmi S.,, "Survey of the products of the interaction of two distinct distributions of particles radiolytiques created at the time of the radiolyse of water by electrons ", Sc. *& Tech.*, Univ. of Constantine, N°19 (2003), pp.36-44.

[2] - Vrigneaud J.M., "Development of a biophysical model for the radio damage assessment led in the chromosomal fiber", Thesis of university Doctorate, CPAT of Toulouse, (2000), p. 98.

[3] - Terrissol M., "Method of simulation of the transportation of electrons of included energies between 10eV and 30keV", Thesis of be-sciences Doctorate, CPAT of Toulouse, (1978).

[4] - Saifi A., "Resolution of the diffusion equation by a semi-implicit method. Application to the case of the water radiolyse by electrons", Radiop., Vol.30, n°4, (1995), pp.557-573.

[5] - Demonchy M., "Modelling of the primary effect of the radiances on the DNA in his/her/its environment", Thesis of university Doctorate, CPAT of Toulouse, (1997), p.46.

[6] - Biosym / Molecular simulation, to use guides 95.O, San Diego, THAT.(1995).

[7]- Demonchy M., Ph.D, Modélisation de l'effet primaire des rayonnements sur l'ADN dans son environnement, T (CPAT de Toulouse 1997).

[8] - Vrigneaud J.M., "Development of a biophysical model for the radio damage assessment led in the chromosomal fiber", Thesis of university Doctorate, CPAT of Toulouse, (2000), p. 98.

[9]- La Verne.J.M, Pumblott.S.M., Effect elastic collisions on energy deposition by electrons in water, J.Phys. chem..A, 1997, 101, 4504-4510.

[10]- Burkert.U, Allinger.N.L., Molecular Mechanics, ACS Monograh177, American chemical Society, Washington, D.C,1982.

[11]- Oudira H., Djamai D., Saifi A., Application d'un modèle déterministe à l'étude de l'influence des molécules radioprotectrices sur les rendements des cassures simple et double brin de la molécule d'ADN, Radioprotection, 2008, 43, 389-408

[12]- CHATWITI .A. Ph.D, Etude de l'évolution temporelle des électrons dans un gaz faiblement ionisé, (Toulouse 1988).

[13]- ZIENKIEWICZ .O.C.,(1988) " La méthode des éléments finis :McGraw-Hill,1971.

[14]- Boumahrat M., Gourdin A., (1983)."Méthodes numériques appliquée", Alger, OPU

[15] - Cullen D.E., "Tables and graphs of μ photon interaction cross-country section derived form the LLNL evaluated photon dated library, Z=1-50" Rad. *Res.*,Vol.6, part TO, Rev 4, Laurence Livemore Library, (1989).

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

AUTOMATIC HATE SPEECH DETECTION IN ONLINE CONTENTS USING LATENT SEMANTIC ANALYSIS

DOI: 10.17261/Pressacademia.2017.612
PAP-WCTIE-V.5-2017(50)-p.368-371

Xhemal Zenuni[1], Jaumin Ajdari[2], Florije Ismaili[3], Bujar Raufi[4]
[1]South East European University. xh.zenuni@seeu.edu.mk
[2]South East European University. j.ajdari@seeu.edu.mk
[3]South East European University. f.ismaili@seeu.edu.mk
[4]South East European University. b.raufi@seeu.edu.mk

ABSTRACT

Internet in general and social media in particular have greatly facilitated the communication, interaction and collaboration among people and different entities. As generally there is no censorship, these media sometimes are used to proliferate discourses that contain hateful messages targeting ethnic origin, religious or sexual groups, which potentially may degenerate to violent acts against individuals of such groups. Therefore, we explore the idea of building of automatic classifier that can be used for detection of hate speech in public Albanian language pages. A hate speech corpus for Albanian language is created, and then based on Support Vector Machine (SVM) approach, an automatic hate speech detection system is proposed. Such system can be used to detect and analyze hate speech in online contents over time and to enhance our knowledge on how they affect opinion creation in society.

Keywords: Hate speech detection, text classification, support vector machines, NLP, Albanian language

1. INTRODUCTION

The continous growth of social media and other Internet services, such as Facebook, Twitter, microblogging or Web services among others has greately facilitated the information exchange, interaction and collaboration among people and different entities. However, the widespread adoption of social media and other online services offers new opportunities to dissiminate hateful messages. Up to date, there is very little research and evidence how the diffusion of hate speech in online contents could trigger hate crimes, yet this potential is recently recognised. For example, Facebook and Twitter pledge to remove hate speech contents within 24 hours after they are reported (Kottasova, 2016). On the other side, EU despite its security and political situations, launched a “code of conduct” to establish public commitments for the biggest Internet companies that the valid hate speech contents will be removed and yet the right to freedom of expression will be preserved (Commision, 2016).

In this context, automatic detection of abusive and hate speech in online contents becomes important topic and task. An automatic detection method could scan large amount of text, analyze and categorize it as hateful or not. The trends of hateful messages could not only be reported to relevant authorities, but it could provide a solid ground to researchers to understand how hateful messages in online contents affect the social processes.

But as noted in (Thomas Davidson, 2017), effective automatic hate speech detection is challenging and very difficult task. The difficulties mostly come from the complexity of natural language processing. The ambiguity and language variability represents a real challenge to be solved. On the other hand, when building more complex and effective automatic machine learning text classifier, the training data becomes crucial.

In this paper, we aim to develope a method to detect hate speech in public online contents in Albanian language, while also addressing the above mentioned challenges. We have collected data from public Facebook pages in Albanian language, and

labeled them as hate speech or not. Than a classifier based on SVM (support vector machines) is trained to differentiate between these categories. To our best knowldge, the contribution in this paper is two-fold:

- It represents the first attempt to create a hate speech corpus in Albanian language
- We make the first attempts to create a hate speech text classifier for Albanian language based on supervised machine learning approach

2. LITERATURE REVIEW

Bag-of-word approaches like in (Kwok & Wang, 2013) are simpler to implement, especially if the classifier is targeting racial hate of speech, but such approaches are insufficieent for accurate classification as it leads to high rates of false positives.

Syntactic features have been explored in (Gitari, Zuping, Damien, & Long, 2015). The experimental results has shown improvements both on precision and recall when used semantic, hate and theme-based features. Chen (Chen, Zhu, Zhou, & Xu, 2012) utilize the profanties, obscenities and pejorative terms as features, weighted accordingly and produced a set of rules to model offensive content, which improved the precision on standard machine learning approaches.

Leveraging morpho-syntactical features, sentiment polarity and word embedding lexicons, Vigna (Vigna, Cimino, Dell'Orlleta, Petrocchi, & Tesconi, 2017) proposed two hate speech classifiers for Italian language based on Support Vector Machines (SVM) and on Reccurent Neural Network named Long Short Term Memory.

Other suppervised approaches to hate speech classification have been proposed as well. Neural language models have potential (Djuric, Zhou, & Morris, 2015), but in all cases the training set data is important. Moreover, the accuracy of hate speech classsifiers could be improved by non-linguistic features, like the gender, ethnicity or age of the author, but this information is often unreliabale or unavailable (Waseem & Hovy, 2016).

3. HATE SPEECH CORPUS

To our best knowledge, there is no previous work on building a hate speech corpus for Albanian language. Therefore, during a period of time, we collected data from Facebook pages in Albanian language and prepared a hate speech corpus that could be used by a classifier. This section reports on data collection, annotatin phase, preprocessing and feature selection in data.

3.1. Data Collection

We explored the Graph API (https://developers.facebook.com/docs/graph-api) provided from Facebook to reterieve and build a corpus of comments from two public pages that publish posts on variety of topics on different political and social events, and which we suspected to find a lot of comments contating hateful speeches. On the other side, we also looked forward posts that contained a significant number of comments. Table 1 summarizes the pages that where crawled and the number annotated posts and comments

Table 1: Dataset Description and Annotations

Facebook pages	Annotated Post	# of Commnets
jetaoshqef	108	4737
tvklan	19	149

3.2. Data Annotation

Two annotators were asked to analyze the content of the crawled comments and to categorize them as *hate* or *no hate*. Overall, 4886 comments received two annotations, and as *hate* were considered only the comments that were categorized as such from both annotators. In total, 2764 comments were categorized as containing hateful content and on other comments either both annotators agreed that the message is not hateful or no concensus was reached.

3.3 Data Preprocesing

In order to prepare the data for the supervides learning algorithm, several pre – processing steps were undertaken. First, the collected text was transformed to lowercase with the objective to improve syntactic matching. Then, extra white spaces, puncuation marks, digits and emoji were removed from the text as they were not considered important in the classifiction process. Finally, we removed from the text the words which we find redundant for text classification (such as conjuctions) and consequently reduced the size of document-term matrix.

4. TEXT CLASSIFICATION MODEL

We tested the Support Vector Machines (SVM) as supervised learning technique used for text classification. As algorithm it captures sparse and discrete features in text classification, which makes it good candidate in our case. On the other side, as noted in (Joachims, 1998) there are theoretical evidence that SVM is an extremely strong performer when having high dimensional input space, few irrelevant features and especilly when most of text classification problems are linerely separable.

We implemented the approach in R System (http://www.rsystems.com/) based on RTextTool. RtextTool is an easy to use tool that can be used for end-to-end implementation by interfacing with existing pre-processing routines and machine learning algorithms. The supporting features include the process from document-term matrix creation, data pre – porcessing, training, classification, up to analytical reports which help users to understand the classification of the employed model.

The speech corpus was divided in two parts. 4000 records were used as training set, and the rest of 886 records were used as testing set. And than based on this dataset, RTextTool functons were used to implement the text classification workflow.

5. FINDINGS AND DISCUSSIONS

While there are many techniques to evaluate the performance of the algorithm, precission, recall and F-score are considered standard evaluation metrics in classification tasks. Accuracy measures In the context of the hate speech system, the accuracy tells what propotion of hate speech comments, are actually hate speech content. Recall tells what percentage of hate speech comments did the algorithm correctly classify, and F-score produces a weighted average of precision and recall.

Table 2 reports the results for the conducted experiement. And the numbers were generated through *create_analytics( )* function contained in RTextTool.

Table 2: Evaluation of Classification Model

Classifier	Precission	Recall	F-Score
SVM	.61	0.57	0.58

6. CONCLUSION

This paper presents the first efforts in bulding an automated hate speech classifier for Albanian language texts. The first experiments show that binary classificaion based on Suport Vector Machines are a promising approach toward building an automated hate speech detection system for online text contents for Albanian language. We are encouraged by initial results, however for the hate classfier of Albanian language to achieve results comparable with similar approaches, it needs richer hate speech corpus and and to explore other language processing features of Albanian language, which for the time being are lacking. However, we believe this work represents the basis toward a building an automated system that could be used to track and monitor online content.

As future work, we intend to extend the annotaed hate speech corpus from different Facebook sites and crawl more comments. This will make richer the current training set, which we believe it will consequently increase the evaluation metrics employed in standard classification tasks. Another important aspect will be to see how other similar supervised learning models will work under the same speech training set.

REFERENCES

Chen, Y., Zhu, S., Zhou, Y., & Xu, H. (2012). Detecting Offensive Language in Social Media to Protect Adolescent Online Safety. *Proceedings of the Fourth ASE/IEEE International Conference on Social Computing.* Amsterdam.

Commision, E. (2016). *CODE OF CONDUCT ON COUNTERING ILLEGAL HATE SPEECH ONLINE.*

Djuric, N., Zhou, J., & Morris, R. (2015). Hate Speech Detection with Comment Embeddings. *Proceedings of the 24th International Conference on World Wide Web*, (s. 29-30).

Gitari, N., Zuping, Z., Damien, H., & Long, J. (2015). A Lexicon-based Approach for Hate Speech Detection. *International Journal of Multimedia and Ubiquitous Engineering*, 2015-230.

(tarih yok). *http://www.rsystems.com/.*

https://developers.facebook.com/docs/graph-api. (tarih yok).

Joachims, T. (1998). Text Categorization with Support Vector Machines: Learning with Many Relevant Features. *European Conference on Machine Learning*, (s. 137-142).

Kottasova, I. (2016). *Facebook and Twitter pledge to remove hate speech within 24 hours.* http://money.cnn.com/2016/05/31/technology/hate-speech-facebook-twitter-eu/.

Kwok, I., & Wang, Y. (2013). Locate the Hate: Detecting Tweets against Blacks. *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, (s. 1621-1622).

Thomas Davidson, D. W. (2017). Automated Hate Speech Detection and the Problem of Offensive Language. *In the Proceedings of ICWSM 2017.*

Vigna, D. V., Cimino, A., Dell'Orlleta, F., Petrocchi, M., & Tesconi, M. (2017). Hate Me, Hate Me Not: Hate Speech Detection on Facebook. ITASEC.

Waseem, Z., & Hovy, D. (2016). Hateful symbols or hateful people? Predictive features for hate speech detection on Twitter. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, (s. 88-93).

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

APPLICATION OF GEOGRAPHICAL INFORMATION SYSTEM (GIS) FOR SPATIAL DISTRIBUTION OF ELECTRICAL TRANSFORMERS IN IKORODU DISTRICT, LAGOS STATE, NIGERIA

DOI: 10.17261/Pressacademia.2017.613
PAP-WCTIE-V.5-2017(51)-p.371-378

Akinwumi A. Akinpelu[1], Olufemi F. Iyiola[2] and Abiodun O. Abisuga[3]
[1&3] Department of Building Technology, Lagos State Polytechnic, Lagos, Nigeria. aaakinpelu@yahoo.com, felixiyiola@yahoo.com
[2]Department of Surveying & Geoinformatics, Federal School of Surveying, Oyo, Nigeria. abisuga_olatunji@yahoo.com

ABSTRACT

Epileptic power supply is a major problem which has great impact on socioeconomic development of Nigeria. Power Holding Company of Nigeria (PHCN) is responsible for power distribution in the country and one of the major challenges is the haphazard distribution of their facilities, most especially, electrical transformers. In addition, spatial locations of electrical transformers and other attribute data are kept using manual approach and in some cases, the records are not available. This study is aimed at using Geographic Information System (GIS) to take inventory of PHCN electrical transformers in Ikorodu district of Lagos, in order to assess their spatial distribution. The research method adopted includes converting the existing map of the study area into digital format through scanning and digitizing, collection of spatial locations of electrical transformers' with GPS receivers and attribute data through social survey. ArcGIS software was employed to create spatial database for forty two (42) electrical transformers located in the study area and various spatial analyses such as query, overlay, buffering, etc. were performed to view the spatial distribution of PHCN facilities. The spatial database created can be easily updated from time to time at a fast rate and reduced cost, and without passing through the whole process of map production. Appropriate recommendations were made for socioeconomic development of the country and effective management of the PHCN facilities. The outcome of the study can be adopted in other developing countries with similar challenges.

Keywords: Geographic information system (GIS), global positioning system (GPS), spatial location, transformers.

1.INTRODUCTION

1.1.Background to Electricity Network in the Last Few Decades

Electric power industries have been developing transmission systems to follow up with the rapid growth of the power demand. On the other hand, the distribution of transformers has been getting restricted, because of development of rural areas and the growing concern over environmental issues. Generally, Electric network can be described as a set of devices used to transport electric power from the power plants to the consumers.These various devices include Overhead transmission lines, underground cables (found mostly in developed countries), switching equipments (circuit breakers and disconnectors), substations (i.e. subsystems equipped with transformers that allow power transfer between systems operating at different voltage levels) and reactive power compensation devices (such as capacitors and reactors).GIS can effectively be used to manage and monitor information on the distribution of transformers to end-users including information describing their spatial and non-spatial attributes such as geographical location and electricity use.

Ikorodu Division is one of the five divisions in Lagos state. With an area approximately 161, 954 square kilometers. It is the second largest local government in Lagos State. It lies about 36 kilometers north-east of the city of Lagos and 26 kilometers from Ikeja, the State capital. It is geographically located on latitudes $3^0$26'31.82"E and $3^0$43'5.13"E and longitude $6^0$41'51.13N and $6^0$31'20.95"N. Like most parts of Lagos State, Ikorodu Local Government Area is a veritable lowland region with relatively

flat undulating features, stretching about 18 kilometers from east to the west, along the Lagos lagoon front. With the opening of Lagos-Ikorodu highway in 1953, Ikorodu division has been an important gateway to Nigeria Hinterland. The climate of Ikorodu Division is similar to that of the entire country, Nigeria, which lies in the world's tropical region. There are two separate seasons, the rainy season which runs from April to October and dry season from November to March. Ikorodu is in the rain forest belt with common vegetation in Nigeria's southern part interspersed with plenty of cassava, yam, and cocoa, kolanut and maize plantations, because of the sandy nature of the soil. The inhabitants of Ikorodu Division are Ijebus, a Yoruba sub- group. The dialect of the people is Ijebu, a Yoruba dialect. Today, the general Yoruba language (or Lagos dialect) is common, particularly among the youths. Farming, hunting and fishing have been the chief occupations of the people. Food crops which are popular in the area are vegetable, melon and beans etc. Cocoa and kolanut are no longer extensively grown as in the past. There are over 200 notable towns and villages in Ikorodu Local Government Area. The six major towns are Ikorodu (headquarter of the Local Government), Imota, Igbogbo, Ipakodo, Ijede and Ibeshe. Other notable towns include Baiyeku, Isiu, Agura, Igberigbe, Oyin, Oreta, Igbokuta, Mowo, Odo- nla (surrounded by Ikorodu Industrial Estate), Odongunyan and Maya.

1.2.Aim and Objectives of the Study

The aim of the study is to generate a geospatial model for electricity and facilities to provide a better understanding towards effective distribution and conservation of electricity. The objectives of the study are:

1. To carry out an appraisal of the existing transformer distribution network in the study area;

2. To map the existing transformers' distribution facilities in the study area;

3. To estimate the number of existing transformers in the study area; and

4. To conduct gap analysis of transformer requirements in the study area.

2.METHODOLOGY

The methodology adopted for this study are database design and creation. Entities were identified and database created upon which queries and spatial analyses were conducted.

2.1.Data Acquisition

The Data acquisition began with the physical phenomenon to be measured. The first step in the methodology of this work was the design of a geodatabase for the entities of interest in the study area. The entities include the Low Tension Electric Poles, High Tension Electric Poles; high and low tension cables, Distribution substation transformer, power transformers, feeders and Injection Substation. This was closely followed by the collection of the required data from the study area. In this case, there were two types of GIS data involved which included: spatial and Attribute data.

2.2.Spatial Data

The spatial data used for the research study include the following;

(i) Base Map: The base map consists of roads, buildings and facilities digitized from a High-Resolution Image data
(ii) GPS Data: GPS Collection points on the available PHCN (Power Holding Company Nigeria) facilities (Poles, Transformers, etc) were acquired from the field of study using a Hand Held GPS.

2.3.Attribute Data

The Attribute data includes the following information which was acquired from the PHCN (Power Holding Company Nigeria), Ikorodu District Office and consumers;

(i) Consumer Details: Meter Account Number, Address of the consumer, the Line Transmission Pole/Pillar to which a consumer is connected;
(ii) Electrical Network Details: 11kV line diagrams with cable sizes, lengths, distribution substation (DSS) transformers, parameters of the equipments, pillars, poles and low voltage networks
(iii) Transformer Details: The details about transformers and other related facilities for this study, PHCN (Power Holding Company Nigeria) bills were acquired so as to acquire information on the average amount of energy consumption.

2.4 Database Design, Creation, Processing and Analysis

Database design is the process of producing a detailed data model of a database. This data model contains all the needed logical and physical design choices and physical storage parameters needed to generate a design in a Data Definition Language, which can then be used to create a database. A fully attributed data model contains detailed attributes for each entity.

The database in the course of this study supports an object oriented vector data model. In this model, real world entities were represented as objects with their various properties, behavior and relationships. In GIS, a database is often referred to as a spatial database or geodatabase and it is the heart of GIS. This process involved the analysis and modeling of real world entities and their interrelationships in such a way that maximum benefits were derived while utilizing minimum amount of data, Kufoniyi, (1998).

Data were formally entered into their respective tables, created in ArcCatalog, in the Arc Map environment. After the data entry process, the personal geodatabase was saved in ArcMap in a folder for easy recovery, geospatial analyses and queries carried out on them. The data stored in the database forms the information base. Once the layers were digitized, the non-spatial data were then added as attributes to the digitized features. This attribute table was linked to the spatial themes containing geographic information. The database created included location and descriptive information for all the different components of the system.

To estimate the areas in the study area that were of short comings in terms of the distribution and develop an estimation trend in the consumption of electricity, the following were operations carried out during the course of this study; overlay operations, spatial queries, proximity analysis and measurement of electricity consumption trend pattern.

3.RESEARCH FINDINGS

3.1.Nature of Transformer Distribution Network in the Study Area

Figure 1 below shows the map of electricity distribution network in the area of study. The map displays the geographic location of all the facilities in the selected area of interest. The entities identified included the Low Tension Electric Poles, High Tension Electric Poles; high and low tension cables, Distribution substation transformer, power transformers, feeders and Injection Substation.

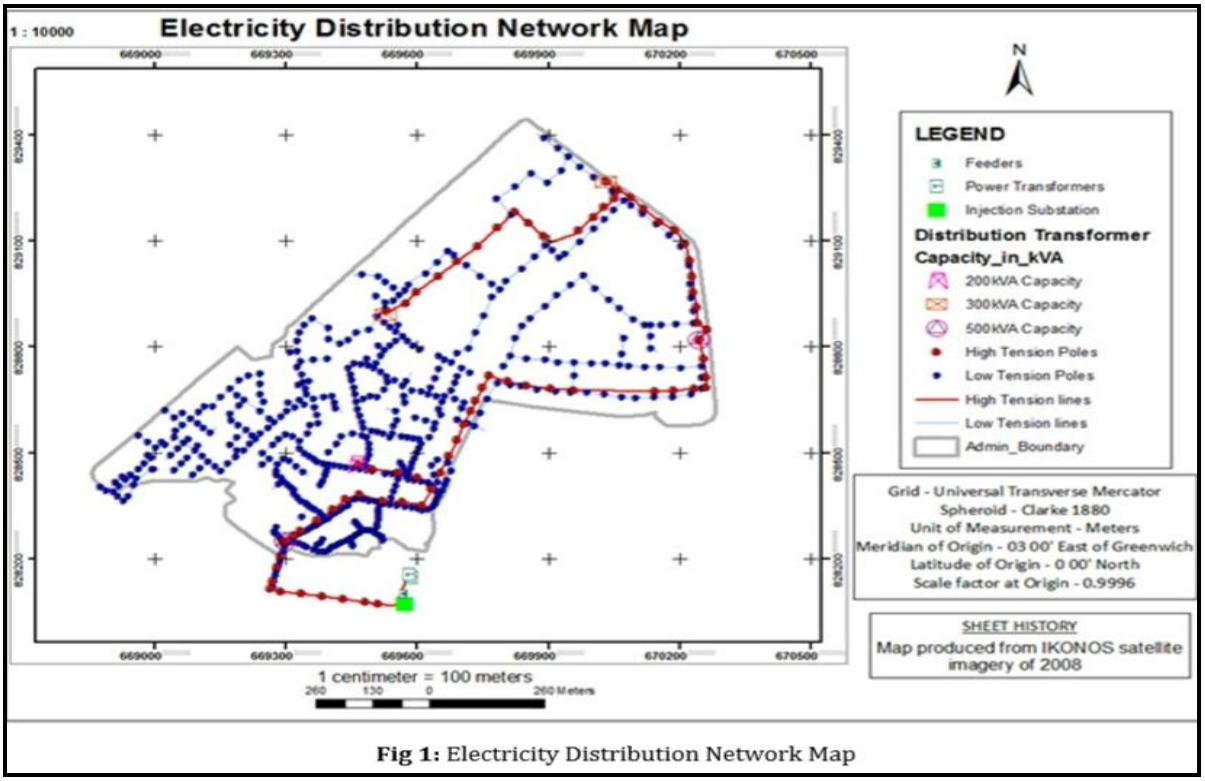


Fig 1: Electricity Distribution Network Map

Figure 2 shows a more detailed map of the electricity distribution facilities connected to the end users - consumers.

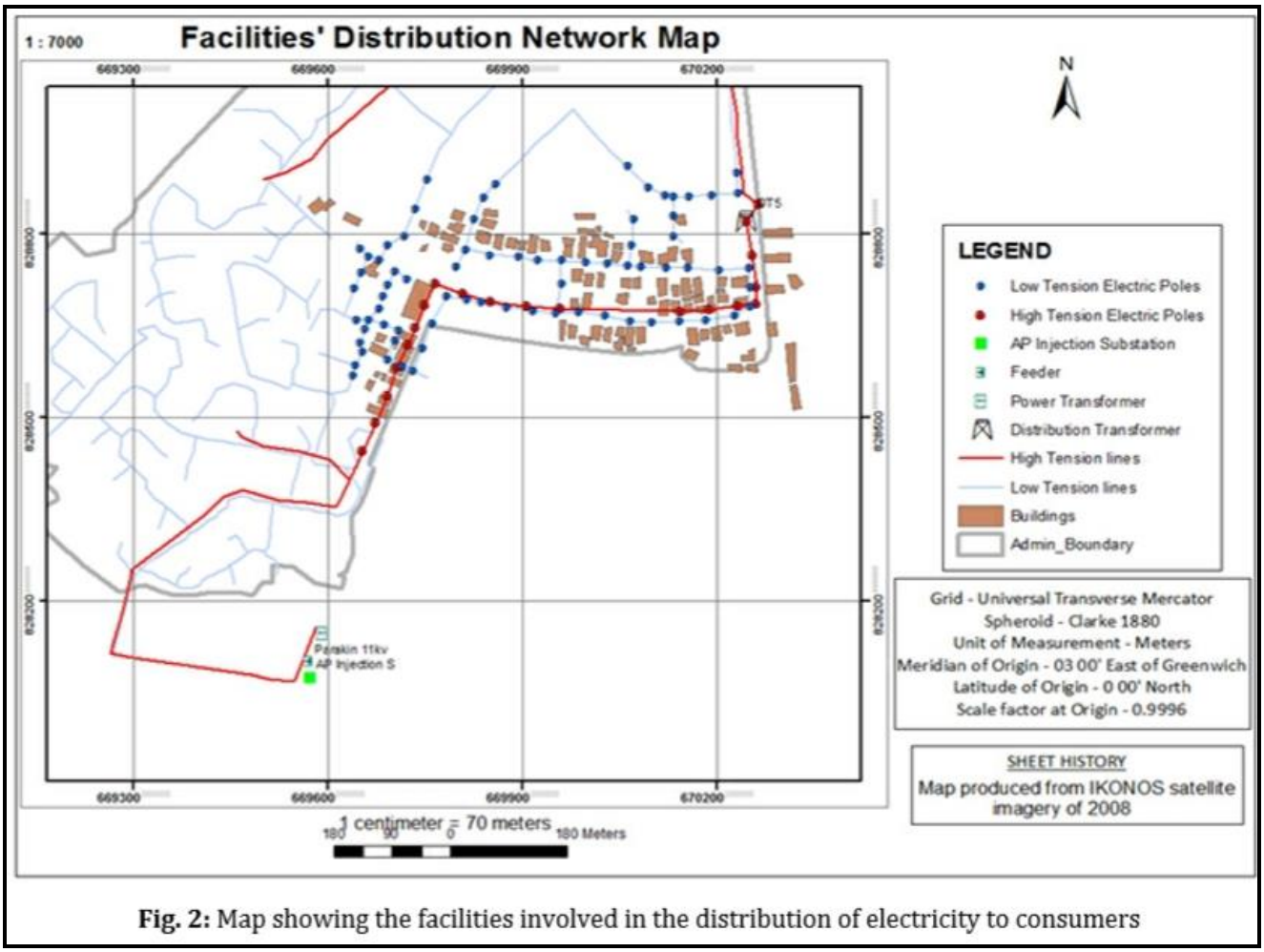


Fig. 2: Map showing the facilities involved in the distribution of electricity to consumers

Information showing all the facilities involved in the service of electricity to the consumers in a selected part of the area of interest is shown in Figure 2 above provides a better understanding of the electricity distribution network pattern.

3.2.Spatial Queries Generation

In GIS, a spatial operation tool is essential for processing or manipulating of data to suit user's need. Geographical Information Systems have a distinctive disparity from other Information Systems which is in the area of spatial analyses. In this project, the spatial search operation was carried out through query generation to retrieve the information stored in the database pertaining to certain systematically defined attributes within the database to answer spatially related questions. This operation involved the link between the database and the map of the selected area of interest. Spatial queries were generated to provide answers to the application use of GIS in developing a model (representation of reality) for the PHCN facilities.

Database query can be referred to as the selection of various combinations of various tables for examination; it involves the retrieval of information stored in the database using Structured Query Language (SQL). However, the queries generated are basically through database extraction in ArcMap 10.2. Some other queries can also be generated based on the user's need. In the course of the project, the results obtained have assisted in the several areas;

i. Monitoring the status of the facilities in the Feeder so as to know the status of the facilities.
ii. Retrieval of Geographical Information of the facilities (Transformer). The spatial information would include coordinates (X and Y), and symbol representation, capacity and location.

iii. Prioritization of facility (Transformer) replacement based on the available information. In this study, of all the existing facilities on ground, none was due for immediate replacement.
iv. Update the information system in the case whereby new facilities (Transformer) are to be installed. In the near future, when new buildings are connected to the existing facilities, the information system can be updated.

Figure 3 suggests proposed location for the installments of new Distribution Substation Transformers within the study area. This was generated based on the average consumption rate by the existing households and the amount of Power generated by each distribution substation transformers connected to them.

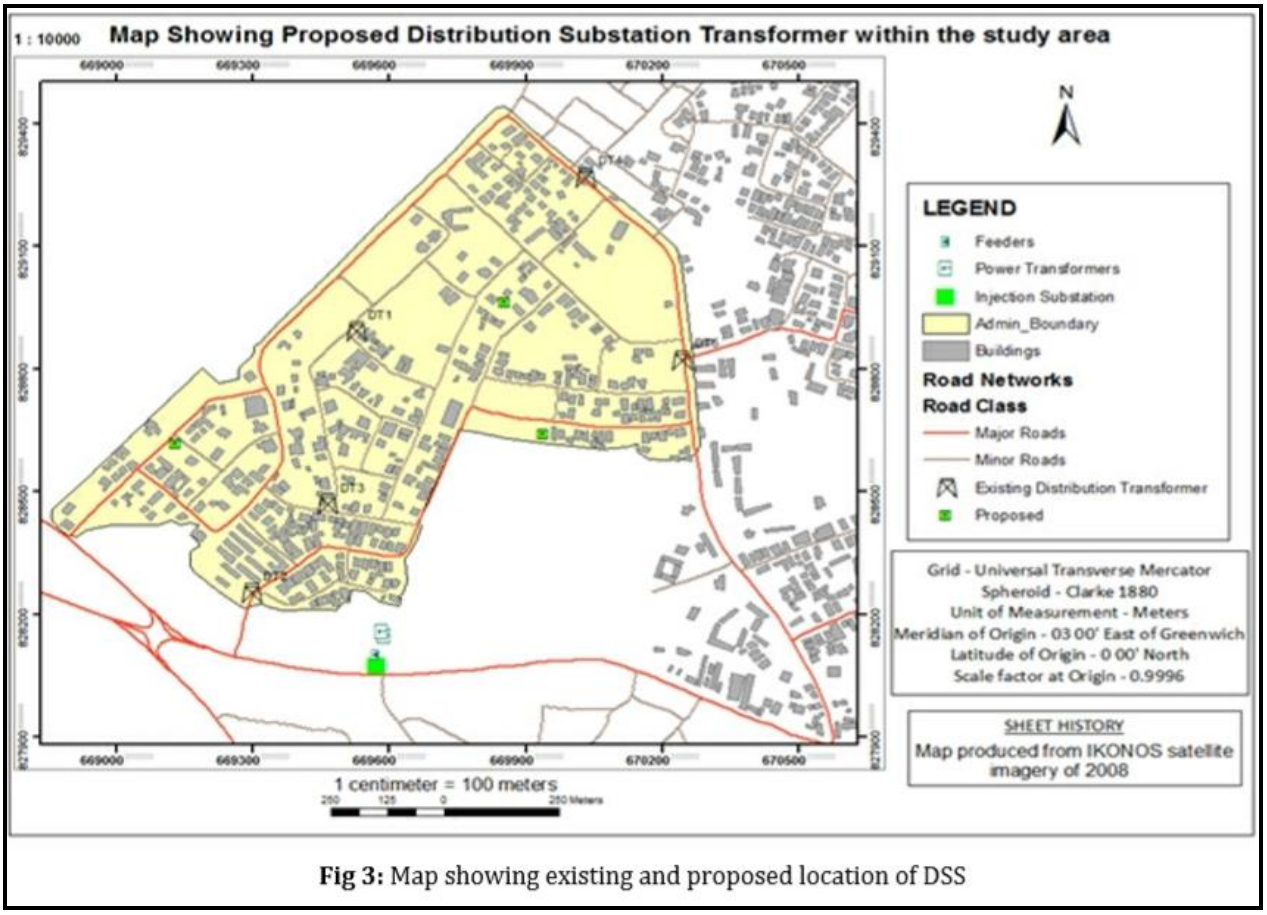


Fig 3: Map showing existing and proposed location of DSS

The above results are expected to equally provide good information for other professionals, which may include Urban Planners, Construction Managers, Engineers, etc. the knowledge of the locations (spatial locations) of the facilities will help to avoid or reduce damages done to these facilities during construction works especially in newly developed areas (such as, Parakin). Proximity Analysis and Estimation of Consumption Trends Other analyses were carried out during the course of this work. Proximity Analysis such as Buffering analysis was carried out to identify the areas for the requested availability and quality of energy supply. As shown in Figure 4 below, the buildings connected to a particular transformer (DT5 in this case) that fall outside a 300m radius experience a drop in voltage compared to those within the buffered area. The buffering analysis performed helped in identifying the pattern of the supply and distribution of power in terms of voltage within the study area.

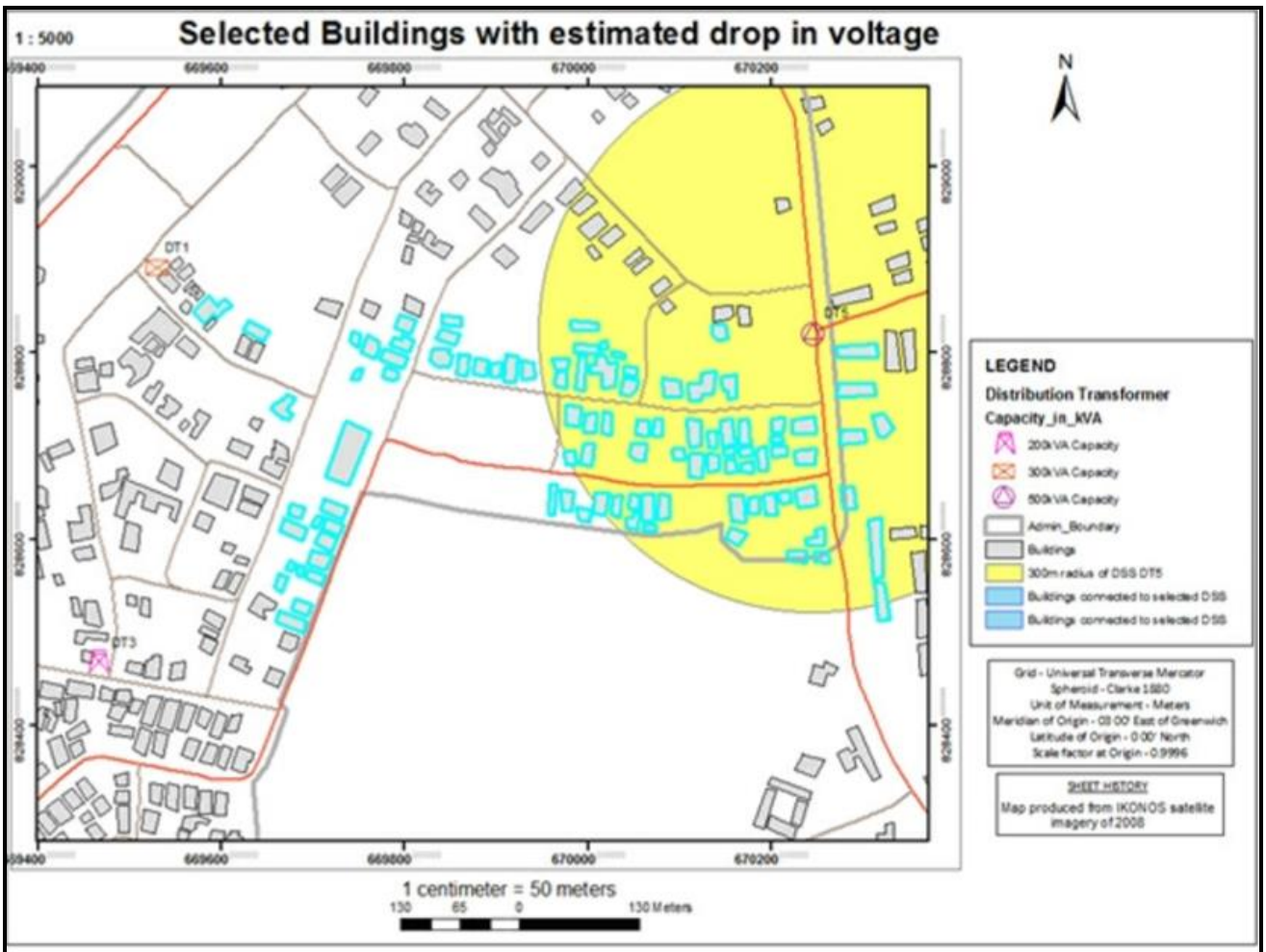


Fig 4: Selected Buildings with estimated drop in voltage

In the case of Figure 5, the result displayed was based on the average consumption rate of the households within the study area. Based on the level of consumption, it clearly shows that most of the households within the area of interest were classed as residential dwellings. The results displayed show the variation in the energy consumption rate within the area. This is because of the various types of electrical appliances used by different households within the area. As represented in the map displayed, the Buildings with darker colors consume higher consumption rate than those with brighter colors.

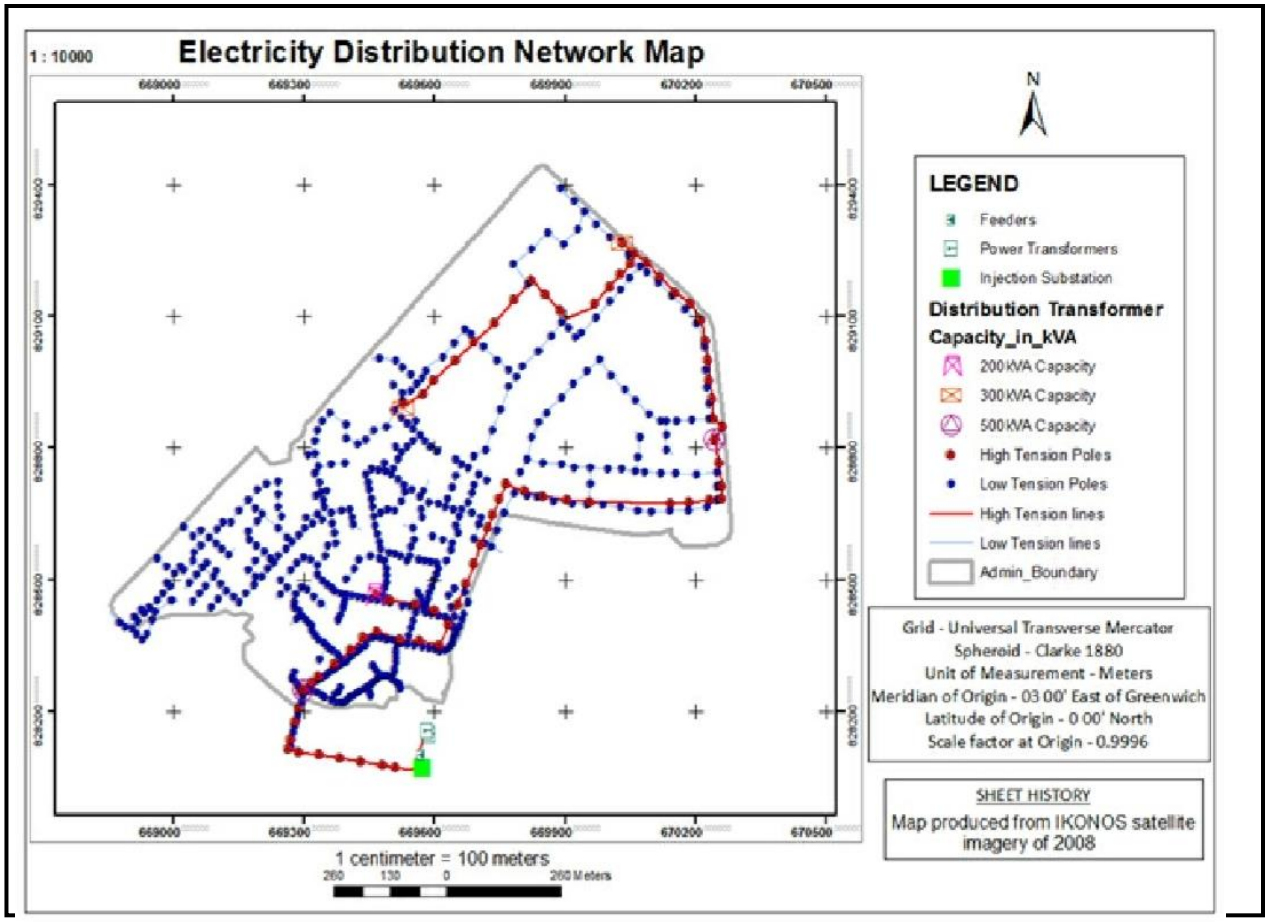


Figure 5: consumption rate of the households within the study area

In order to estimate the consumption trend during the course of this work, other analyses were performed. The relevance and importance of mapping utilities in GIS environment is to provide spatial solutions in a real time manner. The common problems that have been identified in the Power Utility Sector include the following;

1. Unequal distribution/transmission of load;

2. Bad conditions of the equipments due to improper monitoring and maintenance;

3. Frequent power surge due to natural incidences It was been observed that in the study area electrical facilities are either obsolete or too old for efficient performance.

Therefore, routine maintenance and replacement are necessary for efficient performance the electrical equipment in the study area. The threshold of each electrical facility should however be noted at installation and strictly adhered to during operational period. In advanced GIS, locating the facility and finding the exact fault wherein it has occurred could be substantiated by hand held palmtops. The notified location on the spatial network wherein the fault has occurred could be easily identified and trained engineers could be alerted to go to the exact area and rectify the problem immediately. This showed that the major advantage of GIS in Electricity distribution network is the location component.

The electricity distribution network map of the study area revealed the spatial distribution and locations of the electrical facilities within 11kV feeder in the area of interest. In the course of this work, it was observed that electricity consumers within a 300meter radius to a distribution substation transformer, experience higher voltage of power compared to the households outside the radius. It was however noted that about 55% - 60% of households connected to the distribution substation transformers experience average voltage of power supply while about 40% - 45% experience drop in voltage of power supply.

The gap analysis showed that, voltage drop in power supply is experienced the farther the households are connected to a distribution transformer.

As shown in figure 5, the system shows descriptive information about the households and their average rate of consumption. This clearly reveals the consumption rate pattern within the households connected to all the facilities within the area. Within the network, 37% of households consume high amount of electricity, 36% consume average and 27% consume less amount of electricity.

The spatial analysis and the resulting maps provided a better knowledge and understanding on spatial relationships among electricity facilities and consumer demands. Such knowledge is expected to help them know the sections of the service area that lack inadequate services. It is also expected to serve as a guide to them in order to know where they could extend or improve their services and areas that need new installation based on the distribution of power.

4.CONCLUSION AND RECOMMENDATIONS

The study clearly indicates the capability of spatially enabled information system in the management of electricity distribution network. Spatial and attribute data of power distribution network of any part of the selected areas of interest of this study, which are presently acquired, processed, managed, stored and presented in analogue form, can be digitalized. Ayeni et al., (2003) noted that Geospatial Information (GI) is very essential to economic planning and national development. This is buttressed further by Alamu and Ejiobih (2002), when they concluded that a well maintained utility information infrastructure gives up-to-date information on what is where, the state of it, the reaction other actions on it would cause, how it can be harnessed for optimum use of the people and economy. It has been shown that GIS has been employed as one of the technologies for better and improved delivery of networked services. This further enlightens that regardless of cost, GIS has prominently improved the manner of service delivery with respect to time period.

Conclusively, it has been observed that GIS applications have not reached the optimum exploration and there is still room for further exploration and extension in the field as far as utility service delivery like in the case of electricity distribution is concerned. The following recommendations are advanced for efficient electricity distribution networks;

1. More efforts must be made to bring in refined and scientific approaches such as Geographic Information Systems into the management of electricity distribution network.

2. Training programs should be organized for technical online staff of electricity distribution on the integration of GIS in the management/monitoring of electricity distribution.

3. The user requirement and survey analysis conducted before implementation of electricity projects should include spatial information system from the onset so as to forestall drop in voltage within the distribution network.

4. Establishment of more distribution substation transformers in low voltage area within the study area.

REFERENCES

Alamu, E. O. and Ejiobih, H. C. (2002). Utility Information Infrastructure Needs in Utility Organisations in Nigeria: A Case Study of Niger State Water Board. Proceedings of the Technical Session of the 37th Annual General Conference and Meeting of Nigerian Institution of Surveyors, Owerri, Imo State, Nigeria, pp. 85-88.

Ayeni, O. O., Kufoniyi, O. and Akinyede, J. O. (2003). Towards a National Geospatial Information Policy for Nigeria", Proceeding of the Technical Session of the 38th Annual General Conference and Meeting of Nigeria Institution of Surveyors, Lokoja, Kogi State, Nigeria, pp. 21.

Douglas, D. H., (1994), Least Cost Path in GIS Using an Accumulated Cost Surface and Slope Lines. Cartographica 31 (3): 37-51.

Emengini, E. J. (2004). Application of Geographic Information System (GIS) to Utility Information Management: A Case Study of Onitsha-North L.G.A., Anambra State, Nigeria. Unpublished M.Sc. Thesis Submitted to the Department of Surveying and Geoinformatics, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria.

Ezeigbo, C. U. (1998). Application of Geographic Information Systems (GIS) to Utility Mapping, in C. U. Ezeigbo (ed.), Principles and Applications of Geographic Information Systems, Lagos, Panef Press, pp. 124-131.

Korte, G (2001). The GIS book: how to implement, manage and assess the value of Geography

Kufoniyi, O. (1998). Database Design and Creation in C. U. Ezeigbo (ed.). Principles and Applications of Geographic Information Systems, Lagos, Panaf Press, pp. 1-15.

Sipes, J. L. (2007). Spatial Tech- GIS for the utilities Industry. Cadalyst Publications— www.usgs.com— www.gis.rgs.org

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

SELECTION OF HANDOFF METHOD FOR SERVING AIR TRAFFIC CONTROL COMMUNICATION IN LEO SATELLITE CONSTELLATION

DOI: 10.17261/Pressacademia.2017.614
PAP-WCTIE-V.5-2017(52)-p.379-387

Messaoud Garah[1] , Abdelali Achachi[2] , houcine oudira[1], Djamel Benatia[2]
[1] Department of Electrical Engineering, M'sila University, Algeria messa.gareh@gmail.com
[2] Department of Electrical Engineering, Batna University, Algeria

ABSTRACT

Due to the satellite motion with respect to the earth' s surface and the aircraft's mobility, it is required to handover frequently in the process of communicating , In This paper ,assuming that the aircraft's location and speed can be determined , we proposes a suitable study for managing handover in satellite constellations for serving air traffic communication . The main objective of This study is selecting the best scheme which reduce significantly contact failure Probability ,this techniques are: non priority, priority, and queuing handoff schemes ,in this paper we are interested only in schemes that prioritize handover request in particular handoff queuing schemes. Channel allocation strategy (FCA) is employed, and the simulation results obtained concern: call blocking probability, handover failure probability, and average waiting time in the queue. Also simulation results are compared to choose the best one which reduces the contacts failureownership ratios. That supports more firm specific information is incorporated into stock prices as the ratio of foreign ownership increases.

Keywords: Air traffic control, handoff, satellite constelltion, FCA

1. INTRODUCTION

Due to the motion of users (aircraft in our case), more requests of handover are created, a call of a fixed or mobile user can be transferred from a satellite to another since the contact in progress cannot profit a suitable channel of communication in the current satellite, handover [1-2] establishes the transfer of connection of communication of the current channel to another.

For ensure permanent relays between ground stations and aircraft throughout the entire globe The use of satellite systems is required. To handle the increasing aircraft number, the International Civil Aviation Organization (ICAO) proposes a system of air traffic navigation reliable presented in Figure 1[3], capacitive and global based on a concept called Communication Navigation Surveillance (CNS), a system of data link: Aeronautical Telecommunication Network (ATN). The current Air Traffic Management (ATM) procedures [4] are still based on VHF communications and claims for an improvement of ATM concepts. These challenges require the development of satellite communication, navigation and surveillance systems, aiming at providing high reliability and availability system. The increasing number of disasters [5], natural or man-made, occurred during the last years, in the first hours after the disaster, the existing solutions to overcome communication problems when terrestrial infrastructures are not available are the use of satellite communication systems.

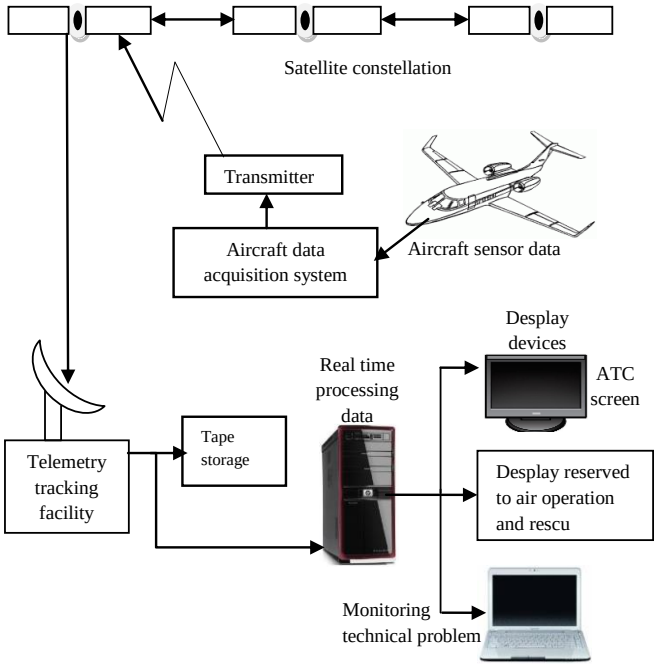
Many LEO constellations have been proposed in the literature (Iridium [6], Globalstar, etc.), while the operation of the Iridium system has given a very important experience for the study of the serious issues of these systems.

The visibility period of a satellite in LEO systems [7] can be about 5 min due to the high speed of satellites. This leads to a remarkable probability of communication interruption and the handover mechanism becomes important for the global

performance of the system. There are two types of handover events, as is the case in land mobile systems, the cell handover [8] and the satellite handover. The first one refers to the transfer of an ongoing contact from one cell to the next one in the same satellite footprint while the second one describes the transfer of an ongoing contact from a satellite to another one.

Few studies have been carried out on the issue of satellite handover, investigating channel allocation policies for new and handover contacts using mainly fixed channel allocation (FCA) techniques. In this paper different queuing policies for handover requests were investigated in order to enhance them in air navigation satellite communication [9]. This paper is organized as follows: Section 2 presents the data model and lists the assumptions and some preliminary notions. Section 3 presents the priority strategies, Section 4 presents the simulations and discussions and the last Section will be the conclusion.

Figure 1. Simplified Block Diagram of Satellite Air Communication System

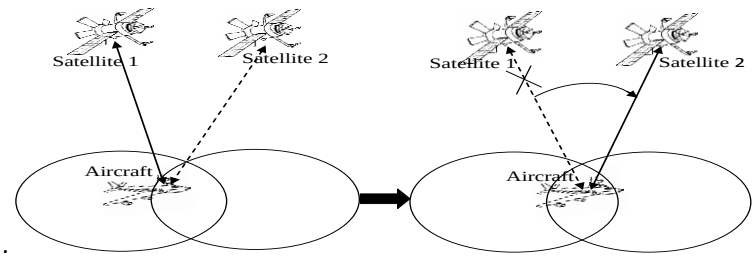


2. DATA MODEL AND PRELIMINARY

For the consistence with previous literatures [10] and its operational availability, the Iridium system is adopted as the basic system model. The satellite ground-track speed is approximately 26600 km/h, the Iridium satellite network is modeled as a one-dimensional environment in which mobile users move in straight lines and at constant speed.

We assume that the arrival of new contacts forms a Poisson process with an average λ and its intensity service is μ. The arrivals of handover requests, presented in Figure 2, form a Poisson process of average λh. If an aircraft is in the satellite cell, the contact duration (with mean 1/μ) is equal to the time during which the contact is in progress.

Figure 2: Hard Handoff between the Aircraft and Satellites



a. Before handoff **b. After handoff**

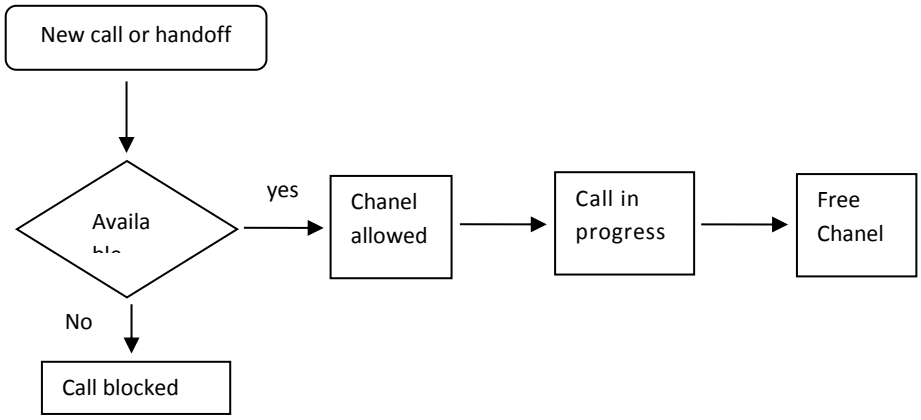
3. HANDOFF SCHEMES IN SINGLE TRAFFIC SYSTEMS

In the coming section, we introduce no priority, priority, and queuing handoff schemes for a single traffic system such as a voice or a data system. We assume that a system has many cells, and each has S channels. The channel holding time have an exponential distribution with mean rate μ. Both originating and handoff contacts are generated in a cell, respectively with mean rates λ_O and λ_H. We assume the system with a homogeneous cell. We concentrate our interest on a single cell (called the marked cell).

3.1. No Priority Strategy (NPS)

In this case, the Handover requests are handled exactly in the same way as an originating contact. So, theblocking probability of handover is equal to the probability of blocking new contacts. The NPS model is presented in Figure 3, where S is the number of channels present in the satellite cell **.**

Figure 3: No Priority Strategy Scheme



When the S channels are free, they will be used by new aircraft contact or Handover. If all channels are busy the new contact will be blocked.

Analytical computation

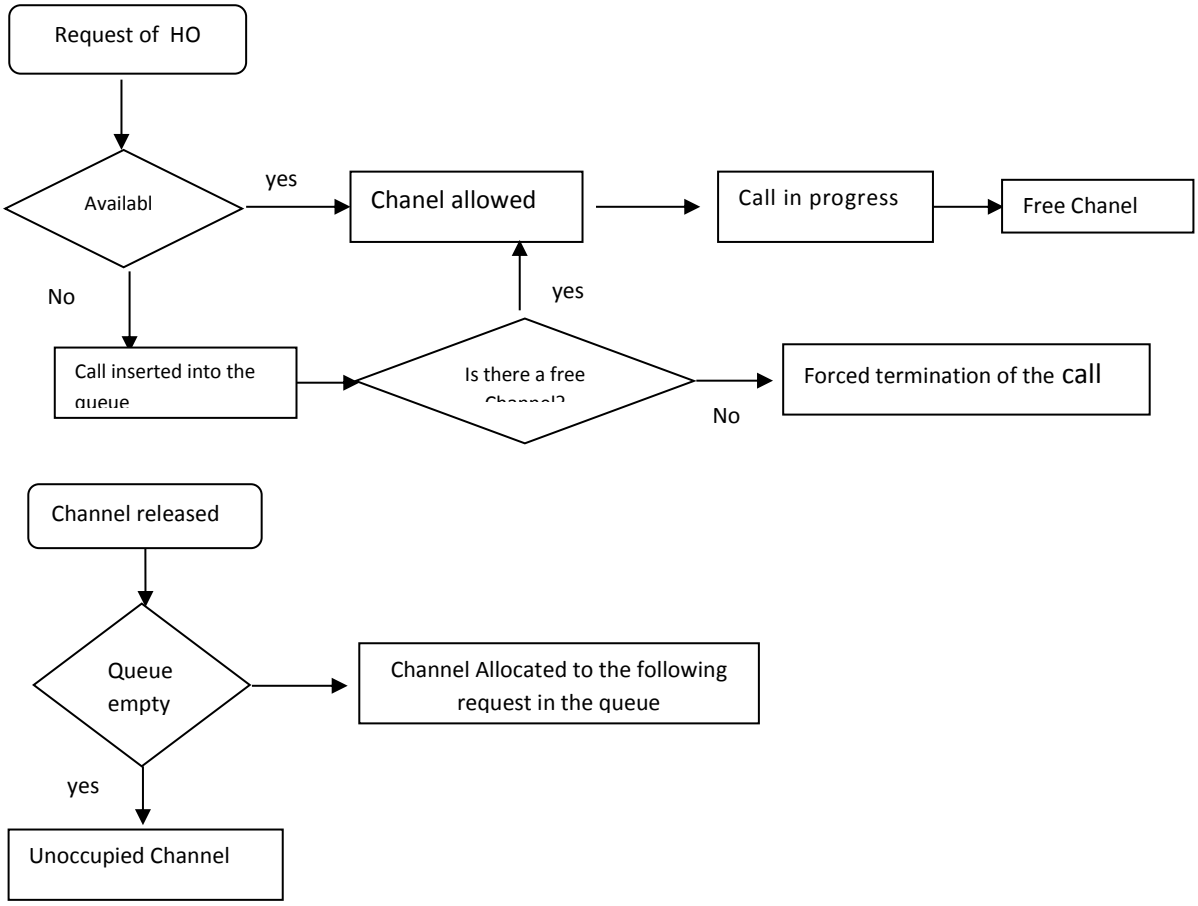
$$Pb = Ph = Ps = \frac{\frac{\left(\frac{\lambda + \lambda hi}{\mu}\right)^{s}}{s!}}{\sum_{k=0}^{s} \frac{\left(\frac{\lambda + \lambda hi}{\mu}\right)^{k}}{k!}} \quad (1)$$

3.2. Priority Strategy

3.2.1. Handoff Call Queuing Prioritizing Scheme (QPS)

in this scheme**,** If all channels in the destination cell are occupied, a handover request is put in the queue as it is shown in figure 4, if a channel is released when the queue of handover request is not empty, the channel is assigned following FIFO discipline, if the received signal strength from the current BS falls below the receiver threshold level before getting service in the target cell, the call is forced to termination.

Figure 4: System Model with Priority and Queue for Handoff Contact



Analytical computation

$$P_n \begin{cases} \dfrac{(\lambda+\lambda h)^n}{n!\,\mu^n} P_0 & 1 \le n \le s-1 \\ \dfrac{(\lambda+\lambda h)^s\, \lambda h^{n-s}}{s!\,\mu^s \prod_{j=1}^{n-s}(s\mu + j\mu_w)} P_0 & n \ge s \end{cases} \quad (2)$$

New contacts are blocked if all channels available in the satellite cell are occupied. We get :

$$Pb = \sum_{n=s}^{\infty} P_n \quad (3)$$

In the state n, the failure probability of handover is given by :

$$P_{b2|n} = 1 - \prod_{j=0}^{n-s}\left[1 - \frac{\mu_w}{(s\mu + \mu_w)^{2j}}\right] \quad (4)$$

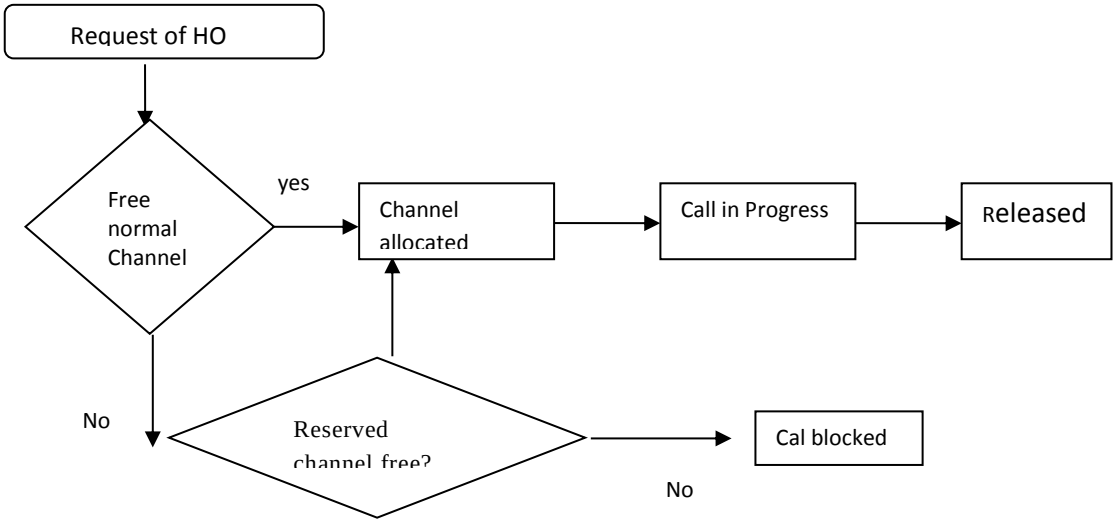
Therefore Ph is given by:

$$Ph = \sum_{n=s}^{\infty} P_{b2|n} P_n \quad (5)$$

3.2.2. Reservation Channels Strategy (RCS) :

Guard channels improve the probability of success of the handover by reserving a fixed number of guard channels reserved exclusively for Handover (Figure 5). The remaining channels are used for Handover and new contacts

Figure 5. System Model with Priority and Reservation Channel for Handoff Contact



Analytical computation

$$P_0 = \left[\sum_{j=0}^{n} \frac{\left(\frac{\lambda + \lambda h}{\mu}\right)^j}{j!} + \sum_{j=n+1}^{s} \frac{(\lambda h)^{j-n}(\lambda + \lambda h)^n}{j!\,\mu^j}\right]^{-1} \quad (6)$$

$$\text{Or } a = \left(\frac{\lambda + \lambda h}{\mu}\right), \quad r = \frac{\lambda}{a\mu}$$

The blocking probability of new contacts is equal to:

$$P = 1 - \sum_j \frac{(a)^j}{j!} P_0 \qquad (7)$$

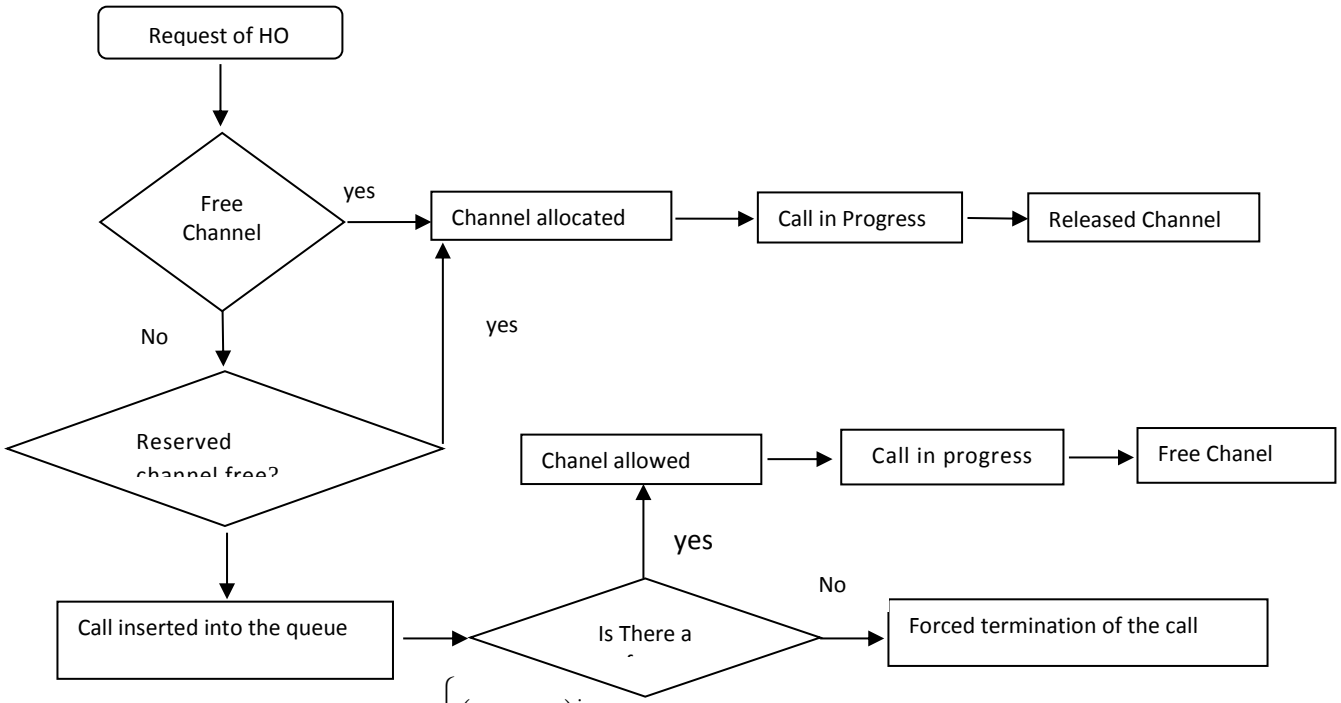
The blocking probability of handover is equal to:

$$Ph = P_0 \frac{a^s}{s!} (1-r)^n$$

3.2.3. Guard Channels with Queue for Handoff contact (QPS +RCS):

It is a combination of the two previous techniques queuing requests and guard channels strategy reserved exclusively for guards Handover .

Figure 6: System Model with Reservation Channel and Queue for Handoff Contact



$$P(i) = \begin{cases} \dfrac{(\lambda + \lambda h)^i}{i!\,\mu^i} P(0) & 0 \le i \le s_C \\ \dfrac{(\lambda + \lambda h)^{s_C} \lambda h^{i-s_C}}{i!\,\mu^i} P(0) & s_C < i \le s \\ \dfrac{(\lambda + \lambda h)^{s_C} \lambda_H^{i-s_C}}{s!\,\mu^s \prod_{j=1}^{i-s} [s\mu + j(\mu_w)]} P(0) & s < i < \infty \end{cases}$$

Where:

$$P(0)=\left\{\sum_{i=0}^{S_C}\frac{(\lambda+\lambda h)^i}{i!\,\mu^i}+\sum_{i=S_C+1}^{S}\frac{(\lambda+\lambda h)^{S_C}\,\lambda h^{i-S_C}}{i!\,\mu^i}+\sum_{i=S+1}^{\infty}\frac{(\lambda+\lambda h)^{S_C}}{S!\,\mu^S}\frac{\lambda h^{i-S_C}}{\prod_{j=1}^{i-S}[S\mu+j(\mu_w)]}\right\}^{-1} \quad (8)$$

So we obtain:

$$Pb=\sum_{i=S_C}^{S}P(i) \quad (9)$$

$$Ph=\sum_{k=0}^{\infty}P(S+k)P_{fh|k} \quad (10)$$

$$P_{fh|k}=1-\left(\frac{\mu_w}{\mu S+\mu_w}\right)\prod_{i=1}^{k}\left\{1-\left(\frac{\mu_w}{\mu S+\mu_w}\right)\frac{1}{(2)^i}\right\} \quad (11)$$

Where $P_{fh|k}$ *is* a probability that a handoff request fails after joining the queue in position *k*+1.

4.Result and Discusion

we suppose that

- Contact duration is exponentially distributed with a mean of 1 min.
- The trafic in the cells follows a poisson distrubtion
- The average waiting time in the queue is exponentially distributed with a mean of 5 min.
- Blocked contacts are lost and cleared.
- The system has a total of 10 available channels per cell.
- The queue length is infinite.
- The simulation results obtained are taken after 10 000s

For the RCS model, we studied three cases according to the guard channel rates used versus to all channels allocated for each cell.

1-Cr=10%

2-Cr=20%

3-Cr=3

Figure 7: Handover Probability Failure as Function of Traffic Intensity for RCS Model

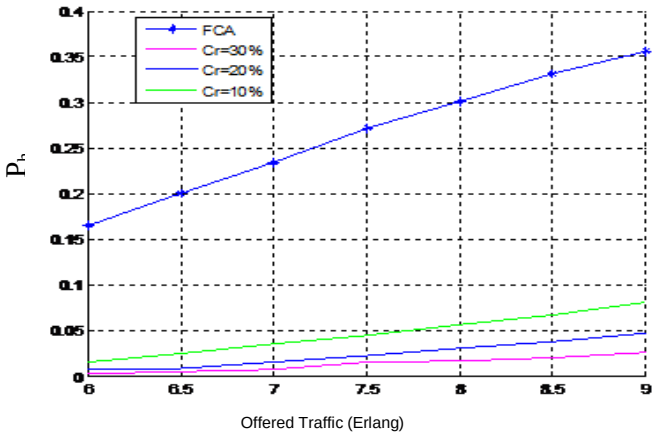
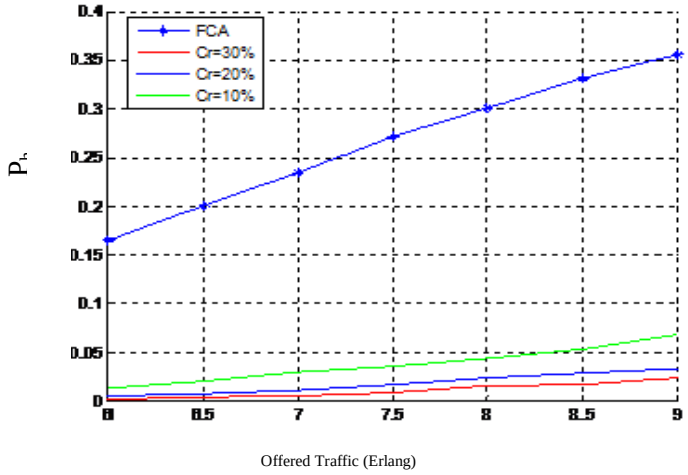


Figure 8: Handover Probability Failure as Function of Traffic Intensity for QPS +RCS t Models



Figures (7) and (8) show handover blocking probability for the RCS and RCS + QPS models respectively depending on traffic density, in the second case (RCS+QPS)the decrease in the probability failure of handover is improved compared to the RCS model, this is due to the presence of the queue.

5. CONCLUSION

The handover is one of the critical procedures of communication in LEO satellite networks, the management of this mechanism must be set appropriately in order to maintain communication between pilots and controllers, and thus ensure an acceptable level of quality of communication.

In this paper we have studied a handoff schemes for satellite constellation network to reduce the probability of forced termination of ongoing contact for serving air traffic control. The scheme employs a mobile aircraft priority based on the residual time of this one in overlap region. Fixed channel allocation (FCA) strategy is adopted for simulation results. In terms of CDP (Contact dropping probability), the results show that the value of this parameter is minimized for the RCS technique, this minimization is improved when using (RCS + QPS) technique.

REFERENCES

[1] B. Jabbari, " Teletraffic aspects of evolving and next-generation wireless communication networks," IEEE Pers. Commun. , vol. 3, Dec.1996, pp. 4–9.

[2] N. D. Tripathi, J. H. Reed and H. F. Vanlandingham, "Handoff in cellular systems" IEEE Pers. Commun., Dec. 1998, pp. 26-37.

[3] Abdelali Achachi, Djamel Benatia, Messaoud Gareh''Satellite Handover Techniques for LEO Networks Serving Air Traffic Control Communication''International Journal of Future Generation Communication and Networking'' vol 7(2014) pp 49-62

[4] G. Satapathy, J. Chen, and D. Tolani, "A traffic information service-broadcast model for mixed equipage aircraft simulationg", Integrated Communications Navigation and Surveillance (ICNS) Conference, Fairview Park, Ohio, **(2010)** May, pp. 1-15.

[5] M. Berioli, J. M. Chaves, N. Courville, P. Boutry, J. L. Fondere, H. Skinnemoen, H. Tork, M. Werner and M. Weinlich, "A rapidly deployable satellite backhauling system for emergency situationsg", Int. J. Satell. Commun. Network, vol. 29, **(2011)**. pp. 419–440

[6] E. D. Re, R. Fantacci, and G. Giambene, "Handover queuing strategies with dynamic and fixed channel allocation techniques in low earth orbit mobile satellite systems", IEEE Trans. Commun, vol. 47, issue 1, **(1999)** Jan, pp. 89–102.

[7] E. Papapetrou, S. Karapantazisny, G. Dimitriadis and F.N. Pavlidou, "Satellite handover techniques for LEO networks", Int. J. Satell. Commun. Network. vol. 22, **(2004)**, pp. 231–245.

[8] M. Gareh and D. Benatia, "Handover prioritizing scheme for reducing call failure probability in cellular wireless network", Wireless Communications and Mobile Computing, vol. 9, **(2009)**, pp. 1660-1667.

[9] Nour El Houda Hedjazi, Malika Ouacifi, Rachida Bouchouareb, Meriem Ourghi, Messaoud Gareh, Djamel Benatia''The Handover in the Constellations of Satellites in Low Orbit''International Journal of Advanced Science and Technology Vol. 41, April, 2012 pp.39-48

[10] L. Chen, Q. Guo, Z. Na, K. Jiang, "A Reservation Pooling Resource Management Scheme for LEO-MSSs with Multi-class Traffic", Journal of Computational Information Systems, vol. 24, issue 8, **(2012)**, pp. 10493–10500.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

APPLICABILITY OF RADIATION ABSORBING AND NON-ABSORBING NANOPARTICLES IN PHOTO-THERMAL ENERGY CONVERSION: A COMPARATIVE STUDY

DOI: 10.17261/Pressacademia.2017.615
PAP-WCTIE-V.5-2017(53)-V.4-p.388-396

Layth Al-Gebory
Department of Mechanical Engineering and CEEE, Ozyegin University, Istanbul-Turkey. Department of Materials Engineering, University of Technology, Baghdad-Iraq. layth.ismael@ozyegin.edu.tr

ABSTRACT

The radiative properties of nano-sized particles play a significant role in a widely range of industrial and engineering applications, such as chemical, electric power, industry, meteorology, biomedicine, biophysics, astronomy, combustion, fire and flame, as well as solar power plants. In the thermal applications field, nanoparticles are found to be an effective electromagnetic wave absorbers within UV–Visible wavelength range; where, short wavelength radiations include high energy photons. Because of their excellent and unique thermo-optical properties which are the basis of thermal phenomena and energy conversion, nanoparticles are used as working media in solar thermal collectors for photo-thermal energy conversion. The purpose of the present study is to investigate the applicability of absorbing (conducting) and non-absorbing (dielectric) nanoparticles in photo-thermal energy conversion. These particles are dispersed in a media to produce nanoparticle suspensions. The comparison between the radiative properties of the two nanoparticles with their suspensions is carried out. The radiative properties include scattering, absorption as well as extinction coefficients which are the most important parameters in the capture and utilization of solar energy. The contribution of particle scattering in the radiation attenuation is investigated by computing the single scattering albedo. Through the analysis of the radiative and thermal behaviour of the particulate medium, it is clear that photothermal conversion is important to not only the solar thermal systems, but also to the electricity generation and solar chemical technology, in addition to the other applications.

Keywords: Nanoparticles, radiative properties, photo-thermal conversion.

1. INTRODUCTION

The most common energy transport that we feel every day is the radiative energy. The energy travels from the Sun to the earth is in the form of electromagnetic radiations, this energy is scattered and absorbed as it goes through the atmosphere, then it is heating up everything around. Electromagnetic radiations allow us to see and feel everything in our surroundings as heat and light. In this operation, it is not only the strength and direction of radiative energy that are important, but also the interaction of the radiations with the objects and how it energizes our life.

Thermal radiation plays a signifcant role in many industrial and engineering applications, particularly at high temperatures. In the cases of convection and conduction, heat energy transfers between two objects depends on the temperature difference between these objects, the temperature difference may be slightly larger than one, and may reach two. In the thermal radiation transfer case which takes place between two distant objects depends on the difference between the fourth power of their absolute temperatures. If the materials properties are depend on the temperature and this included in the calculations, the radiative energy flux can be proportional to such a higher power of absolute temperatures, and as a result the importance of radiative energy transfer is dramatically enhanced at high temperatures. Because of this, radiation energy contributes significantly to energy transfer in combustion chambers, furnaces, rocket plumes, fires, high temperature heat exchangers, and during chemical explosions. therefore, the radiative energy transfer would improve the

operation and design of such applications and devices. To this end, radiative energy transfer calculations need to be made strictly, which requires the utilize the accurate radiative properties.

Radiative energy transfer governs the distribution of temperature of the sun and emitted objects. Understanding this directional and spectral nature of radiative energy propagation is the basis for solar energy utilization. Given that we need more sustainable and cleaner energy production for the daily purposes and growing population of Earth, implementation of optimum and efficient radiation energy transfer principles for solar energy utilization and capture which have an important impact on the environment.

When an electromagnetic radiation or a photon interacts with a participating (particulate), the intensity of the incident radiation may be changed by scattering and/or absorption. One of the common examples of this interaction is absorption of sunlight by a cloud of smoke (which is a multitude of suspension of fine particles in air), another example is the scattering of sunshine by the atmosphere (molecules in the atmosphere, in fact, these are tiny particles) resulting in red sunsets and blue skies, and rainbow colors. Astrophysicists were the first dealt with radiation scattering by particles, and they were interested in the scattering of starlight by the dust. Scientists from other fields are concerned with the scattering of electromagnetic radiations: In the thermal field, photo-thermal energy conversion have been proposed by many researchers as one of the main energy sources. Therefore, solar power systems take a significant part of the research area, where the sustainable energy is needed.

Through the analysis of the solar energy utilization technology, it is clear that photothermal energy conversion which is important to not only the thermal utilization, but also solar chemical technology and the electricity generation. In solar thermal systems, solar thermal collector is the most important part in the photothermal conversion. In regular solar thermal collectors, surfaces of the evacuated tube and flat plate are heated by the incident solar radiation and then the thermal energy is transfered away by the working media (fluids) through convective heat transfer. The thermal resistant in converting the incident solar energy into thermal energy of fluids is large especially when vapor or gas is applied as the working media. Thus, the surface that absorbs the solar radiation has high temperature. One of the most important drawbacks of surface absorber is the limit thermal efficiency. Besides the surface absorber, volumetric solar absorption is another type of solar absorber in which the incident solar radiation is absorbed by a volume of fluid rather than the surface. Compared with surface absorbers, volumetric solar absorbers have many prominent advantages. Specifically, when a suspension of smal particles is used. Because the suspended particles are in contact with the base fluid, the incident solar energy absorbed by suspended particles can be efficiently transferred to the continues phase (surrounding fluid). Particles dispersed in liquid or gas have large surface areas to volume ratio that can enhance radiation absorption as well as convective heat transfer. Therefore, volumetric radiation energy absorbtion can efficiently operate under a high solar flux.

The progress in the nanotechnology and nanoscience during the past two decades or so that makes the application of nanoparticles in the solar thermal utilization. Specifically, the application of nanoparticle suspensions in solar thermal collectors which is proposed in different researchs. Indeed, nanoparticles are found to be an effective electromegnatic radiation absorbers at low wavelength range where short wavelength radiations included photons that carry high energy.

2. LITERATURE REVIEW

Radiative properties of nanoparticles and nanosuspensions have been widely investigated in the recent years, the effects of different nanoparticles types in addition to their physical and chemical characteristics on the radiative properties and radiative transfer phenomena have been observed.

The study of Lorenz predates that of Mie, this work proposed the general theory that describing radiative scattering by absorbing spherical particle, the theory known as the “Mie theory”. Recently, in recognition of the contributions of Lorenz, the “Lorenz–Mie theory” becomes one of the most popular and reliable theory. The Lorenz-Mie theory cover the general scattering solution spherical particle (absorbing or non-absorbing) without limitation on particle size (Mie G. A. 1908).

An extensive reviews on thermal radiation phenomena in particulate suspensions have been introduced by Dombrovsky and Baillis (Dombrovsky et al. 2010), and also by Tien and Drolen (Tien et al. 1987).

Lord Rayleigh discussed radiation scattering and absorption by single spheres during the later part of the nineteenth century (Rayleigh 1964).

the impacts of dependent scattering were studied in a medium consisting of two spheres at an arbitrary distance between them and the index of refraction typical for that of soot particles (Ivezic et al. 1996).

Dependent and independent scattering regimes were clarified for packed systems contain 10 spheres. A comparison study was carried out between the Foldy's effective field approximation (EFA) and the quasi crystalline approximation (QCA) for nanoparticle suspension in the Rayleigh scattering regime (Prasher 2005).

A hybrid methodology was introduced based on elliptically polarized light scattering (EPLS) to characterize metal oxide colloidal suspension (Aslan et al. 2006).

A comparison was done on cases include both fluid base absorbing and non-absorbing matrix, also for metal and non-metal particles (Wei et al. 2012).

A Theoretical method was developed for obtaining the radiative properties of nanoparticle suspensions with particle agglomeration. The microstructure of particle agglomeration groups have been generated using the diffusion limited cluster aggregation (DLCA) technique (Du et al. 2015).

The thermo-optical properties were investigated for nanoparticle suspensions for low temperature direct absorption solar thermal collectors (temperature between 00 C and 1000 C) applications to study the potential of using suspended particles in solar power plants (Karami et al. 2016).

In the present study, the applicability of the absorbing (conducting) and non-absorbing (dielectric) nanoparticles in photo-thermal energy conversion is investigated. For that, two types of nanoparticles are proposed in the calculations, aluminum (Al) and alumina (Al2O3) nanoparticles, these nanoparticles are with particle size dp=50 nm and suspended in a medium (water) to produce a nanosuspensions. The radiative properties (scattering, absorption as well as extinction coefficients) of the two nanoparticles types with their suspensions are calculated using the single scattering approximation technique based on the Lorenz-Mie theory. In addition, a comparison between the radiative properties of the two nanoparticles with their suspensions is carried out.

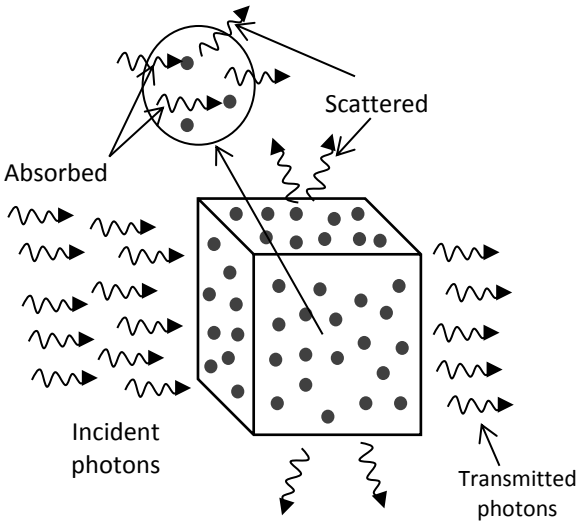
3. DATA AND METHODOLOGY

When incident radiations interact with participating (particulate) medium containing small particles, the intensity of the incident radiation may be changed by scattering and/or. How much it scattered or absorbed and into which direction an electromagnetic radiation passing through its vicinity depends on different parameters including (i) the shape and siz of the particle, (ii) the material of the particle (absorbing or nonabsorbing), (iii) the complex index of refraction, $\mathrm{m} = \mathrm{n} - \mathrm{ik}$, and (iv) the clearance (surface to surface distance) between particles (Modest 2012, Howell et al. 2016).

Different approaches have been considered to obtain the best fitted mathematical theory for describing the radiative properties of any particulate (participating) media. Light scattering by particles is characterized by the Rayleigh and the Lorenz-Mie theories. The Rayleigh scattering theory, after Lord Rayleigh, is applicable to dielectric (non-absorbing), spherical and small particle. On the other hand, the Lorenz-Mie theory, after Gustav Mie, is coverd the spherical particle (absorbing or non-absorbing) without limitation on particle size (Hahn 2009, Mishchenko 2014).

The conditions of the Rayleigh scattering theory are $\mathrm{x} \ll 1$ and $|\mathrm{m}|\mathrm{x} \ll 1$, where x is the dimensionless size parameter ($x = \pi d_p/\lambda$), m is the complex refractive index ($m = n_{particle}/n_{media}$), and n is the refractive index. Fig. (1) describes the radiative phenomenon in a particulate media. This figure shows an enclosure containing suspended particles in a medium. When the incident radiation travels through this media, it is attenuated by absorption and/or scattering, while other incident radiations are transmitted through this media to the other side.

Fig. (2) Incidence radiation on a particulate media



The Lorenz-Mie theory is proposed to present the interaction of incident radiation with a homogenous spherical particle. Based on such theory, the angular intensity functions (i_1 and i_2) are describe the differential scattering cross sections, which are given by (Bohren et al. 1983, Mishchenko 2000):

$$\sigma_{VV} = \frac{\lambda^2}{4\pi^2} i_1$$
$$\sigma_{HH} = \frac{\lambda^2}{4\pi^2} i_2 \qquad (1)$$

In these two equations, the subscripts VV and HH refer to vertically and horizontally polarized incident light. The radiative efficiencies Q_i for the interaction of radiation with a particle of radius r_p is defined as:

$$Q_i = \frac{C_i}{\pi r^2} \qquad (2)$$

Where, C_i is the cross sections normalized to the particle cross section πr_p^2, where i refers for extinction ($i = ext$), scattering ($i = sca$) and absorption ($i = abs$). Then, considering the conservation of radiative energy, the extinction efficiency and cross section is represented as (Howell et al. 2016):

$$or \begin{matrix} Q_{ext,p} = Q_{scat,p} + Q_{abs,p} \\ C_{ext,p} = C_{scat,p} + C_{abs,p} \end{matrix} \qquad (3)$$

In particulate suspensions with low particles volume fraction, the scattered intensity of a number of suspended particles is equal to the scattered intensity from a single particle multiple by the number of whole particles, which is known as the linear summation rule (SSA) (Howell et al. 2016). This rule is valid and suitable for most radiative transfer applications where independent scattering is considered. Then, the scattering and absorption coefficients of the suspended nanoparticles can be calculated as:

$$\sigma_{scat} = NQ_{scat} \text{ and } \sigma_{abs} = NQ_{abs} \qquad (4)$$

where, N refers to the number of particles per unit volume which can be replaced by the particles volume fraction as following:

$$\emptyset = NV_p = N\pi\frac{d_p^3}{6} \quad (5)$$

where, V_p is the volume of the particle.

Then, the scattering and absorption coefficients become:

$$\sigma_{scat} = \frac{3}{2}\frac{\emptyset\,(Q_{scat})}{d_p}$$
$$\sigma_{abs} = \frac{3}{2}\frac{\emptyset\,(Q_{abs})}{d_p} \quad (6)$$

The extinction coefficient (scattering+absorbtion) of the nanoparticles is as following:

$$\sigma_{ext,p} = \sigma_{scat} + \sigma_{abs} \quad (7)$$

Yet, only the effect of the suspended particles on the radiative phenomena is demonstrated. While, any absorption effect comes from the medium should be included in the calculations. Therefore, the extinction coefficient of the medium is obtained from (Otanicar et al. 2009):

$$\sigma_{ext,m} = \frac{4\pi k_{medium}}{\lambda} \quad (8)$$

where, k_{medium} is the absorption index for the medium (water) that varies as a function of wavelength.

Then, the total nanoparticle suspension extinction coefficient when combined is as given below:

$$\sigma_{ext,total} = \sigma_{ext,p} + \sigma_{ext,m} \quad (9)$$

To explain the importance of the particle scattering behaviour in the total radiation attenuation, the single scattering albedo (ω_λ) is calculated, which is an important parameter specifically for non-absorbing (dielectric) particles:

$$\omega_\lambda = \frac{\sigma_{\mathrm{scat,p}}}{(\sigma_{\mathrm{abs,p+m}} + \sigma_{\mathrm{scat,p}})} = \frac{\sigma_{\mathrm{scat,p}}}{\sigma_{\mathrm{ext,total}}} \quad (10)$$

4. FINDINGS AND DISCUSSIONS

Two types of nanoparticles are under investigation in this study, Aluminium (absorbing particle) and alumina (non-absorbing particle), the two nanoparticles at the same size $\mathrm{d_p} = 50\ \mathrm{nm}$ suspended in the base fluid (Water) with different particle volume fraction (0.001, 0.006, and 0.01 %). Therefore, the comparison in the radiative properties is carried out between two different particle types and with different particle concentration in the same medium. The radiative properties is the basis of the photothermal energy conversion, where the incident light intensity is converted to the thermal energy (the radiation is attenuated inside a media).

Fig. (2) shows the radiative efficiencies of the two nanoparticles types, the results obtained from the Lorenz-Mie theory. Zero absorption efficiency is obtained for the alumina ($\mathrm{Al_2O_3}$) nanoparticles, and aluminium (Al) particle shows a high scattering efficiency in comparison with that from the another particle. The radiative efficiency is inversely proportional with the wavelength, where nanoparticle is found to be an effective electromagnetic wave scattering and absorption at low wavelength (depending on te type of the particle). Aluminum particle shows high optical efficiencies in comparison with the second particle type ($\mathrm{Al_2O_3}$) particle at the same wavelength.

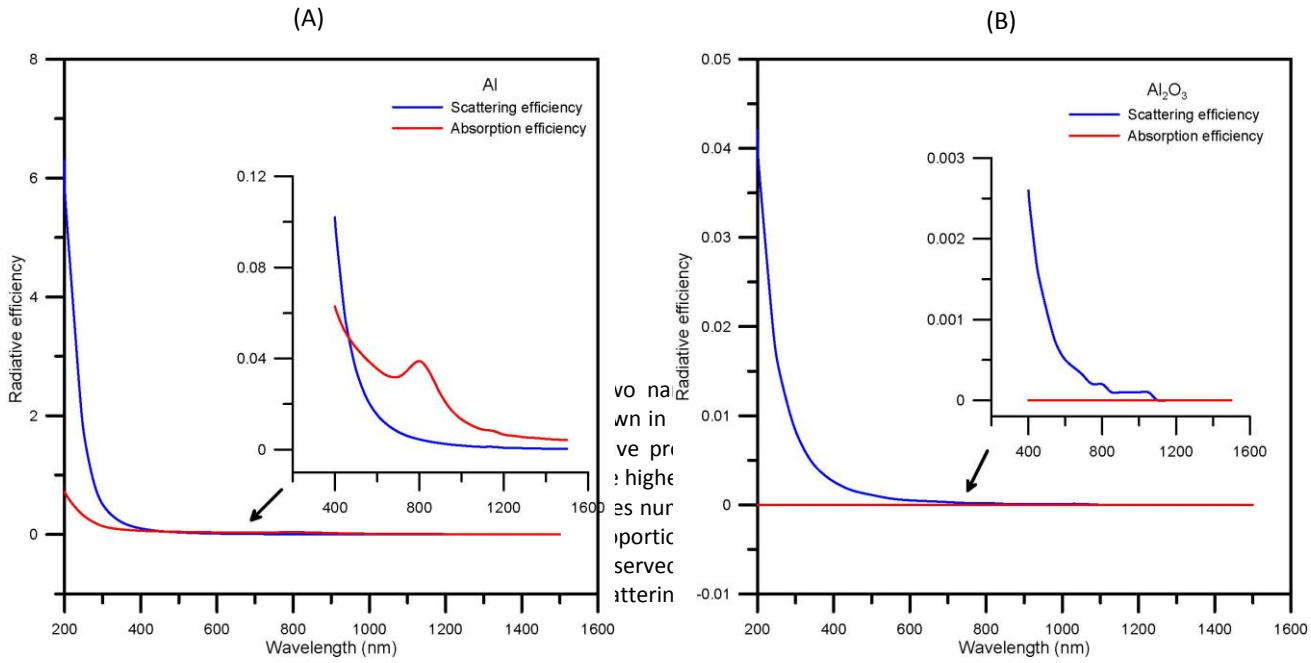


Fig. (2) Radiative efficiencies (scattering and absorption). (A) $\mathbf{Al}$ and (b) $\mathbf{Al_2O_3}$ nanoparticles

Scattering phenomena for particles in a participating media are affected by number of parameters including particle type, size, structure as well as the effect of the scattering from other particles in the same medium. The optical properties of a particle carries a significant impact on the radiative transfer phenomena.

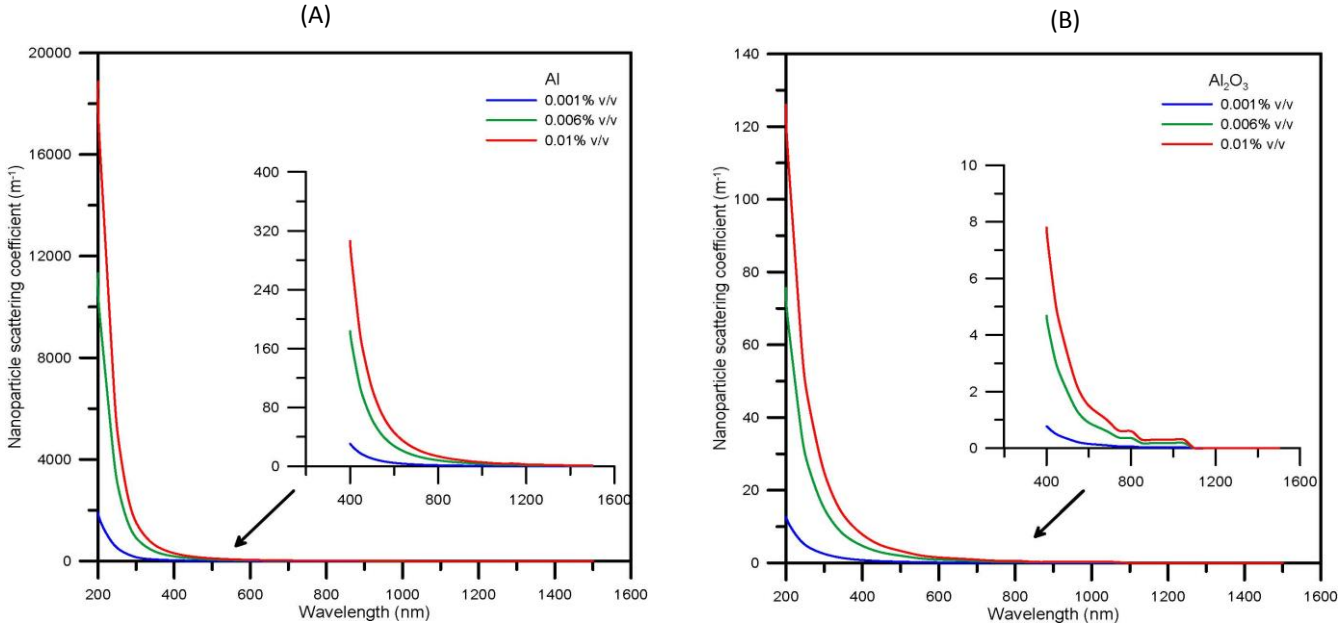


Fig. (3) Scattering coefficient for the nanoparticles. (A) $\mathbf{Al}$ and (b) $\mathbf{Al_2O_3}$ nanoparticles

The absorption coefficient of the Al nanoparticle is shown in Fig. (4). The same behaviour of the scattering coefficient, the absorption coefficient of the absorbing (Al) particle is inversely proportional to the wavelength. On the other hand, zero absorption coefficient is obtained for the non-absorbing (Al_2O_3) nanoparticle. These are a result of the optical properties for the both nanoparticles, where each type of nanoparticles has its own optical (refractive and absorbtion) indices. The effect of particle concentration is also demonstrated in this figure.

Absorption coefficient play a signifcant role in the radiative transfer phenomena, which can be observed from the radiative transfer equation. In general, scattering and absorption coefficients of particulate media carries important impact in different applications. In thermal applications, these coefficients represent how mutch light is attenuated inside a medium.

Fig. (4) Absorption coefficient for (Al) nanoparticle

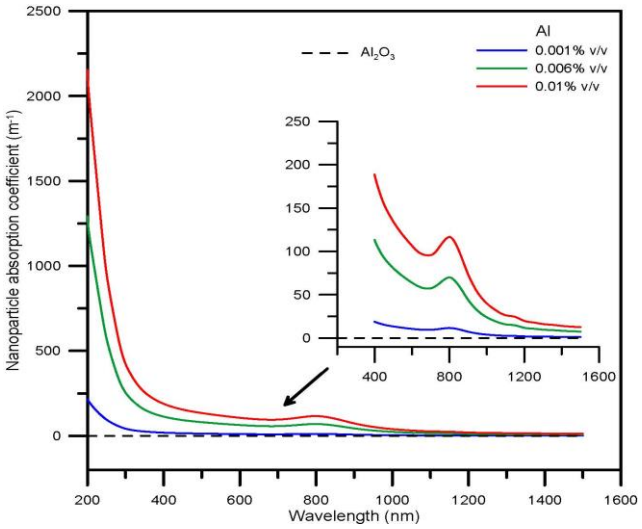
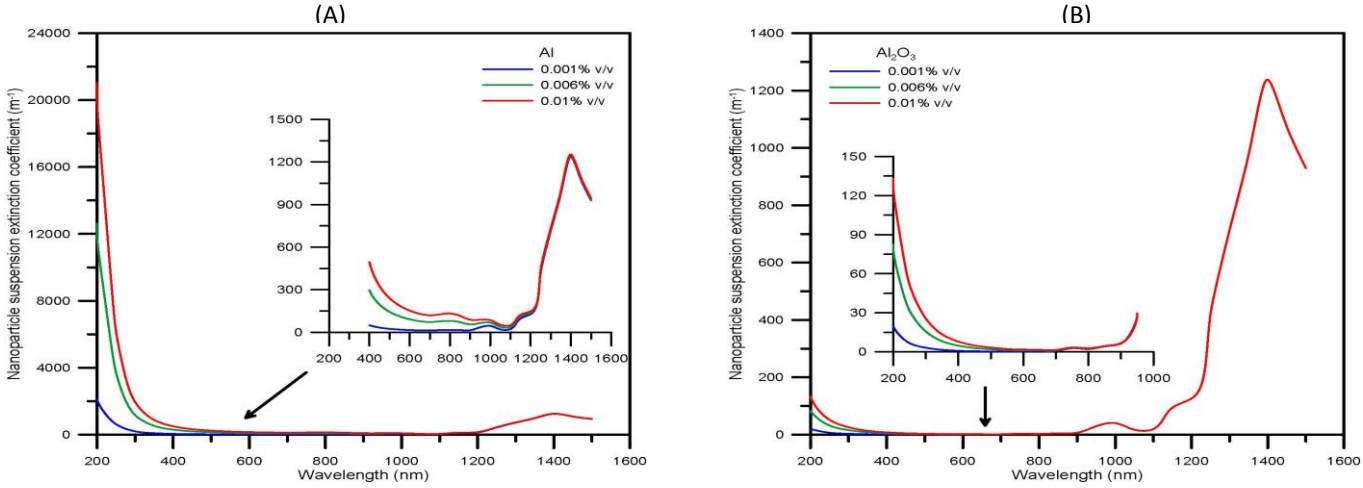


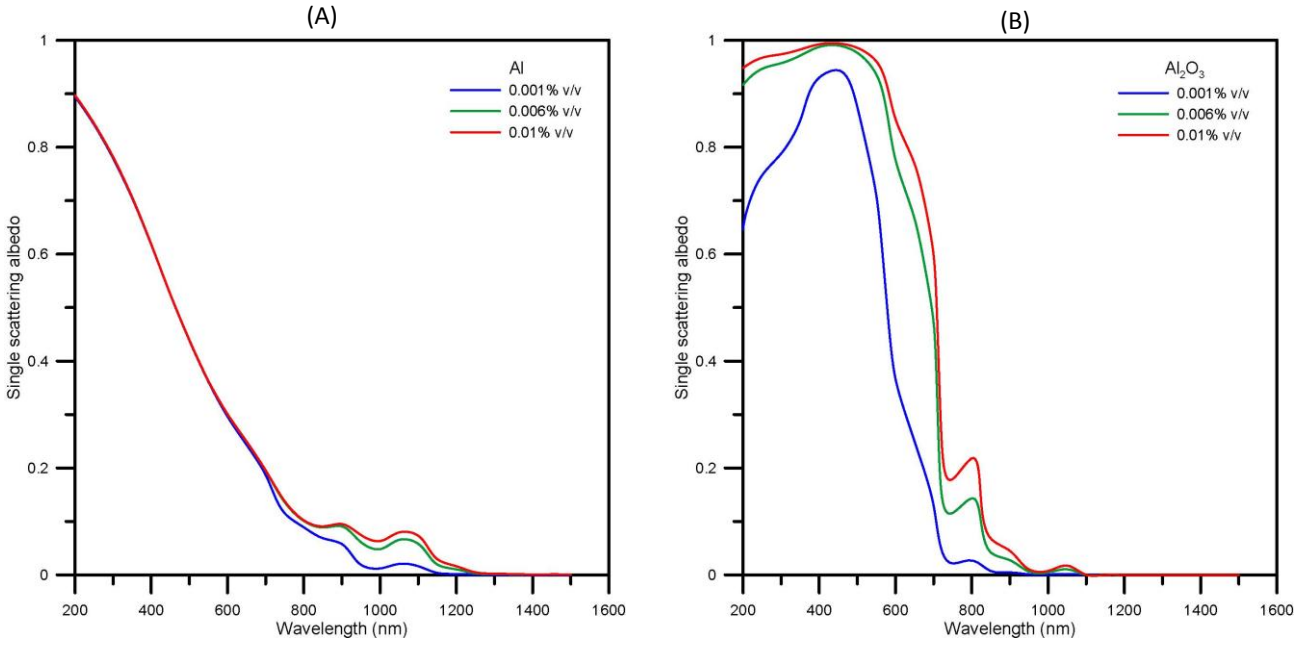
Fig. (5) shows the total extinction coefficient for the nanoparticle suspensions. The total extinction coefficient includes the effects of both nanoparticle and medium (water). The comparison between the two nanosuspensions types (water based Al and Al_2O_3) shows the importance of scattering and absorption coefficients of the nanoparticles in the radiation attenuation specifically at low wavelengths. While, the absorption coefficient of the medium play a significant role at longer wavelength (Infrared wavelength range). Radiation attenuation of a media indicates the photothermal energy conversion in such media.

Fig. (5) Extinction coefficient for the nanoparticle suspensions. Water based (A) Al and (b) Al_2O_3 nanosuspensions



To explain the effect of the particle scattering behaviour on the total extinction coefficient , Fig. (6) shows the single scattering albedo for the nanoparticle types. A signifcant value of the single scattering albedo is observed for the water based Al_2O_3 nanosuspensions, where this particle is a non-absorbing partilce and only scattering from the particle is contributed in the total extinction coefficient. While, for the other nanoparticle (Al) low value of the single scattering albedo is obtained because the effect of the particle absorption is included in the total attenuation coefficient. In general, the single scattering albedo decreases as the wavelength becomes longer as a result of decreasing the scattering values.

Fig. (6) Single scattering albedo for the nanoparticle suspensions. Water based (A) Al and (b) Al_2O_3 nanosuspensions



Nanoparticle are introduced for use in a wide range of industrial and engineering applications. Promising applications required suspended nanoparticles play a significant role in the thermal and optical properties at different wavelength ranges. Different nanoparticles have their own properties, and investigations are required to optimize the radiative properties for different nanoparticles (absorption and non-absorption nanoparticles). The present study, explored different radiative properties for these nanoparticles at different conditions (particle volume fraction) which is desired for a specific application including photothermal energy conversion.

Indeed, different radiative properties are obtained by changing particle concentration in the same medium, which is desired when a specified radiative properties is desired. In the field of photo-thermal energy conversion, optimizing the radiative properties of a participating medium is desired, where the amount of attenuated light represent the energy conversion in such media.

REFERENCES

Aslan M. M., Mengüç M. P., Manickavasagam S. and Saltiel C. (2006), "Size and shape prediction of colloidal metal oxide MgBaFeO particles from light scattering measurements", Journal of Nanoparticle Research.

Bohren C. F. and Huffman D. R. (1983), "Absorption and Scattering of Light by Small Particles", Wiley-Inter science Publication, Canada.

Du M. and Tang G. H. (2015), "Optical property of nanofluids with particle agglomeration", solar energy 122, 864-872.

Dombrovsky, L. A., and D. Baillis (2010), " Thermal Radiation in Disperse Systems: An Engineering Approach", Begell House, New York.

Fan L. and Zhu C. (1998), "Principles of Gas-Solid Flows", Cambridge University press.

Hahn D. W. (2009), "Light scattering theory", Department of Mechanical and Aerospace.

Howell J R, Mengüç M P, Siegel R. (2016), "Thermal Radiation Heat Transfer", 6th edition, CRC Press, New York.

Ivezic Z. and Mengüç M. P. (1996), "An investigation of dependent/independent scattering regimes using a discrete dipole Approximation", lnt. J. Heat Mass TransJer. Vol. 39, No. 4, pp. 811-822.

Karami M., Akhavan -Behabadi M.A., Dehkordi M. R. and Delfani S. (2016), "Thermo-optical properties of copper oxide nanofluids for direct absorption of solar radiation", Solar Energy Materials and Solar Cells 144, 136–142.

Mie, G. A. (1908), "Beitr¨age zur Optik tr ¨ uber Medien, speziell kolloidaler Metall¨osungen", Annalen der Physik, vol. 25, pp. 377–445.

Mishchenko M. I. (2014), "Electromagnetic Scattering by Particles and Particle Groups (An Introduction)", 1st Edition, Cambridge.

Modest M F. (2003), "Radiative Heat Transfer", Academic Press-Elsevier science", USA.

Otanicar T., Taylor R. A., Phelan P. E. and Prasher R. (2009), "Impact of size and scattering mode on the optimal solar absorbing nanofluid", Proceedings of the 3rd International Conference on Energy Sustainability, San Francisco, California, USA.

Prasher R. (2005), "Modification of Planck Blackbody Emissive Power and Intensity in Particulate Media Due to Multiple and Dependent Scattering", J. Heat Transfer 127(8), 903-910.

Rayleigh L. (1964), "On the light from the sky, its polarization and colour", Philos. Mag., vol. 41, pp. 107–120, 274–279, 1871, (reprinted in Scientific Papers by Lord Rayleigh, vol. I: 1869–1881, No. 8, Dover, New York.

Tien C. L., and Drolen B. L. (1987), "Thermal radiation in particulate media with dependent and independent scattering", in Annual Review of Numerical Fluid Mechanics and Heat Transfer, vol. 1, Hemisphere, New York, pp. 1–32.

Van de Hulst H. C. (1981), " Light Scattering by small particles", Dover publications, Inc., New York.

Wei W., Fedorov A. G., Luo Z. and Ni M. (2012), "Radiative properties of dense nanofluids", Applied Optics, Vol. 51, No. 25.

Mishchenko M. I. (2000), "Light Scattering by Nonspherical Particles (Theory, Measurements and Applications)", academic press, USA.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

HOMEWORK SYSTEM DEVELOPMENT WITH THE INTENTION OF SUPPORTING SAUDİ ARABIA'S VISION 2030

DOI: 10.17261/Pressacademia.2017.616
PAP-WCTIE-2017(54)-V.4-p.397-401

Atifa Elgimari[1], Shafya Alshahrani[1] and Amal Al-shehri[1]
[1] Department of Computer Science, University of Bisha, Bisha- Kingdom of Saudi Arabia, atfa000@gmail.com

ABSTRACT

This paper suggests a homework system based developing a website. By using the suggested homework system, hard copies of homeworks were replaced by soft copies. It is expected to participate in applying Saudi Arabia's Vision 2030, specially in the education sector, where it considers that the primary education is its foundation stone, as the success of the Vision depends in large assess on reforms in the education system generating a better basis for employment of young Saudis.

Keywords: Homework, teacher, parent, education, system, vision

1. INTRODUCTION

Homework is a long-standing education tradition that, until recently, has hardly ever been questioned. It has many types of definitions; it is defined as out-of-class tasks assigned to students by their teachers as amplification or an extension of classroom. Homework is a headache for most parents because it is a daily obligation [3]. The author [2] discussed how to advice parents to be powerful leaders in their kid's education. He answered the most common questions parents ask such as When do I reward and when do I punish?, How do I give advice about studying without getting backtalk?, How much should I review homework?, How do I help my stressed out child who takes school too seriously?

What do I say when my child says "I don't care" about school?, How do I help my disorganized child without feeling like I am his secretary?, and so on.

A popular book [1] introduced the term homework from the respect of recent changing in culture and in the different societies. However, Internet and bookstores are full of books offering parents advice on how to get children to do homework, but still homework has become problematic for more and more students, parents, and teachers. As the culture has changed, and as schools and families have also changed.

For sure homework places great importance in the educational process, moreover, [5] considered homework as a base for the learning process. The Ministry of Education in Saudi Arabia, general education sector, annually provides students with hard copies of homeworks note for free. Nevertheless that costs the Ministry of Education a huge amount of money. This paper suggests a homework system based on developing a website application. By using the suggested homework system, hard copies of homeworks are replaced by soft copies. From different viewpoint, parents were not involved in the educational process over the past few years. Recently, involving parents in children's education is being placed in the educational process. The suggested homework system engages parents and teachers on a day-to-day basis. Teachers can send homework to the system; parents can browse them, therefore they encourage their children to do their homework.

The objective of this paper is to provide a framework for a homework system that would help teachers and parents in collaborating goals that best fit students' achievement. The suggested homework system is expected to participate in

applying Saudi Arabia's Vision 2030, which it considers a cornerstone of education. The success of the Vision depends on large assess on reforms in the education system generating a better achievement.

2. SAUDI ARABIA's VISION 2030

Vision 2030 was approved by the Saudi cabinet on 25-April-2016, which confines the objectives of the government for the coming 15 years. According to reference [4], the vision carefully addresses three major themes (society, economy, nation), and assigns goals to each sub-theme, then concludes with commitments for each theme. Among the commitments that are applicable to education, research, and renewable energy. The vision gives young Saudi students extraordinary significance. It plans to build a centralized student database tracking students from early childhood through to K-12 and beyond into tertiary education (higher and vocational).

Vision 2030's view about the transformation process of Education in Saudi Arabia, its goal is to draw their attention to attract and retain the finest Saudi and foreign minds, and facilitate all needed requirements for them, in addition, to provide them with all they need.

To this end, and for getting excellent impacts, we will prepare a modern curriculum focused on rigorous standards in literacy, numeracy, skills and character development. The vision will track progress and publish a sophisticated range of education outcomes, showing year-on-year improvements. It will work closely with the private sector to ensure higher education outcomes are in line with the requirements of the job market [6]. Additionally, The vision will continue to improve and reform their regulations, paving the way for investors and the private sector to acquire and deliver services – such as…education – that are currently provided by the public sector. The vision will look to shift the government's role from providing services to focusing on regulating and monitoring the educational system. It will build the capability to monitor this transition.

Saudi Arabia's Vision 2030 sets out a motivated road-map for education development in the Kingdom of Saudi Arabia. This road-map reflects a collective vision for developing curricula, teaching methods, students' values and institution skills.

3. COLLECTING DATA

No doubt good data collection is the foundation, on which good research is built [7].

Our group was enabled to gather data information through interviews of respondents from the different schools. The respondents included teachers, psychologists, and socialists. This was used as the primary method in collecting data on the preferred use of the end-user on the supposed system.

All of this information was gathered in order to be of help in comparing between the current system and the suggested homework system.

We had a discussion and shared different ideas in sequence to decide the best way to develop the suggested homework system. The choices were, developing a website or developing a mobile application. After discussion and review of the data, the decision was to build a website. We also decided the programming language to be used and other applications and materials needed in turn to make the suggested system possible.

In addition to interviews, the data on which this study was based was obtained from a questionnaire. The questionnaire was distributed over different types of social media sites.

4. RESULTS

The questionnaire conducted in a nationwide of a randomly selected sample of 377 persons. 41.8 % out of them were teachers, and 58.2 % were parents. Regardless of Poll, (Fig.1) the majority members of the random selected sample were parents. Respondents completed an extensive questionnaire whose main issue of investigation was a comparison between the current system and our suggested system. In respect of using the current system, we examined the effect of handwriting notes of homeworks on lecture time. 85% of teachers answered that handwriting notes are time consuming (Fig.2). Specifically, we related these answers to the benefits of using the suggested system; it was found that 87.4 % out of 127 teachers' responses presented their willingness to use the suggested homework system (Fig.3). Comparing with 92 % out of 177 parents' responses presented their willingness to use the suggested homework system (Fig.5).

The data assembled asked parents many questions such as "has it happened that your son/daughter has forgotten his/her homework note, and that reflected negatively on his/her academic achievement?" Poll shows that about 39% of parents have agreed, 28.8% disagreed, 32.2 % said sometimes out of 177 respondents (Fig.4).

The suggested homework system will empower parents and teachers in collaborating together in order to improve students' achievement.

Figure 1:" Is the responder a teacher or a parent?"

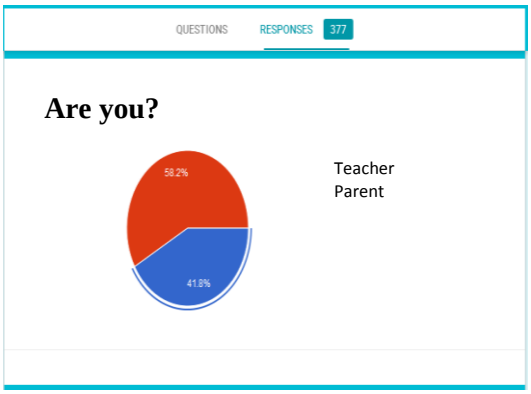


Figure 2: Questionnaire asked teacher" Do you think that handwriting notes of homework is time consuming for lecture time?"

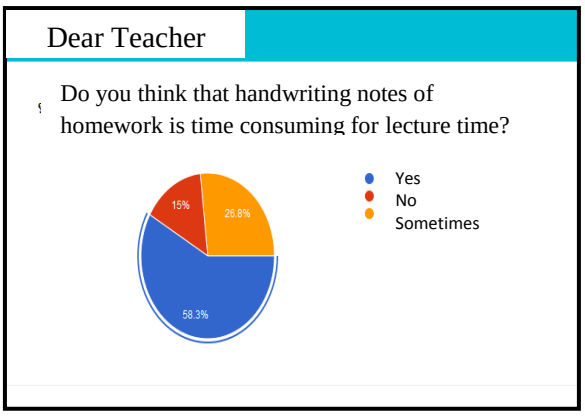


Figure 3: Questionnaire asked teacher "Are you going to use our suggested system?"

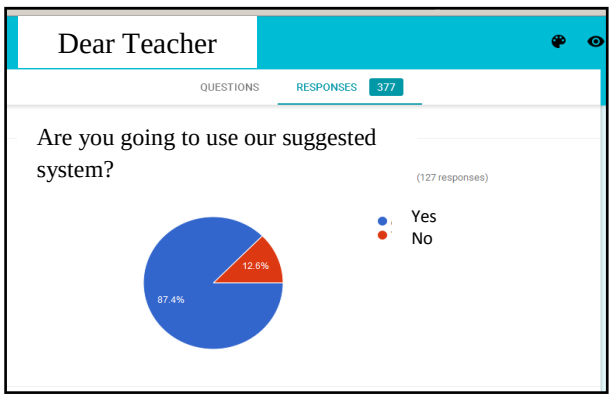


Figure 4: Questionnaire asked parent "Had it happened that your son/daughter forgotten his/her homework note, and that reflected negatively on his/her academic achievement?"

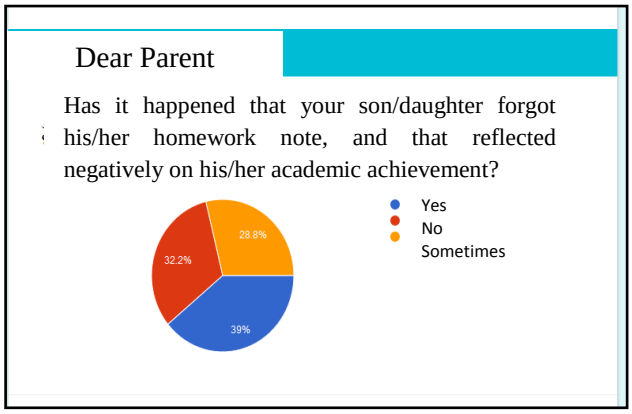
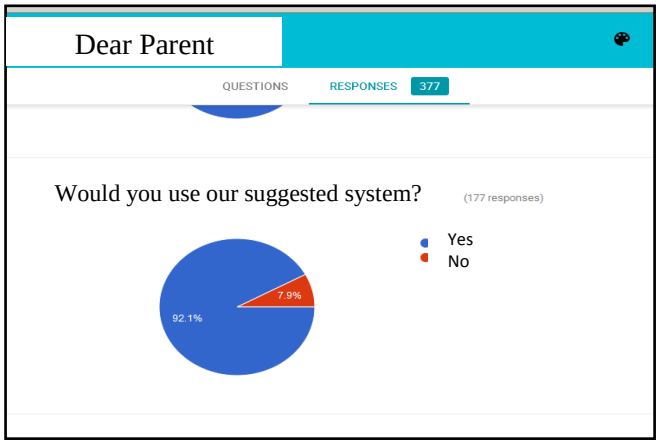


Figure 5: Questionnaire asked parent "Would you use our suggested system?"



5. CONCLUSION

In this paper, a homework system is suggested in order to support Saudi Arabia's Vision 2030 in transmitting from using traditional learning methods to contemporary learning methods. In sequence, to check the impact of our suggested homework system, a questionnaire was developed. In recent years, involving parents in children's education is being placed in the educational process; therefore, Respondents of our developed questionnaire included teachers and parents. Although both teachers and parents have been asked by our developed questionnaire, the questionnaire has been more attractive for parents (58.2 %) comparing with teachers (41.8 %). The end-users of the suggested homework system will be parents, teachers, and administrators. In summation, Polls show that both parents and teachers are willing to use our suggested homework system; moreover, teachers are enthusiastic to discontinue using the traditional methods in teaching. Our suggested homework system will support vision 2030 to be well applied on technological basis.

REFERENCES

[1] Greenfeld, Karl Taro, "My Daughter's Homework Is Killing Me," The Atlantic, September 18, 2013

[2] Neil McNerney, "Homework: A Parent's Guide To Helping Out Without Freaking Out!," integrated publishing 1984 Isaac, Newton square, suite 204, December 26, 2011

[3] Cathy Vatterott, "Rethinking HOMEWORK Best Practices That Support Diverse Needs," ; ASCD, Alexandria, Virginia USA,2009

[4] https://sacm.org.au/saudi-vision-2030-envisages-a-thriving-education-research-and-renewable-energy-sectors/?refreshed.

[5] Cathy Vatterott, "Making Homework Central to Learning," ASCD, November 2011

[6] http://www.tamimi.com/en/magazine/law-update/section-14/august-2/vision-2030-and-the-transformation-of-education-in-saudi-arabia.html

[7] Wendy Olsen, "Data Collection –Key Debates and Methods in social Resaerch," SAGE publications.Inc, 2012

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

OPTIMIZATION OF DISSOLUTION OF COLEMANITE ORE IN POTASSIUM DIHYDROGEN PHOSPHATE SOLUTION (KH2PO4)

DOI: 10.17261/Pressacademia.2017.617
PAP-WCTIE-V.5-2017(55)-p.402-409

Ozlem Karagoz, Soner Kuslu
Atatürk University, Engineering Faculty, Department of Chemical Engineering, 25240 Erzurum, Turkey.

ABSTRACT

Boron is one of the most important richnesses of Turkey which has approximately 72% of the known boron reserves globally. The production of boron compounds has essentially expanded recently due to increasing demands. Colemanite rich in boron is a common raw material to produce boron components and the first step of this process is the dissolution of colemanite by using different leaching solutions. The main aim of the study is to investigate the optimization of dissolution of colemanite in potassium dihydrogen phosphate **(**KH_2PO_4) solution. Taguchi method was used to determine the optimum conditions while effectiveness of the parameters were identified by variance analysis. Reaction temperature (T), KH_2PO_4 concentration (C), stirring speed (W), solid to liquid ratio (S/L), and particle size (D) of colemanite were selected as parameters affecting the rate of colemanite dissolution. The optimum conditions for these parameters were determined. As a result of the experiment made under optimum conditions, both 98% of B_2O_3 passed into the solution and potassium borate by-product were produced by crystallization.

Keywords: Colemanite, optimization, potassium dihydrogen phosphate, taguchi method.

1.INTRODUCTION

Boron is an important element which has the capacity to form a large number of complex chemical compounds. Boron is videly used in a diversity of applications including the nuclear, fuel, glass, electronic and computer, energy devices, medicine, cosmetics, paper, plastics, metallurgy, textile and agriculture industries [1]. Almost 72% of the world boron reserves are found in Turkey. The most important of these boron ores are colemanite, tincal and ulexite. Colemanite has a monoclinic crystal structure with a chemical formula of $2CaO.3B_2O3.5H_2O$. It is used to produce boric acid (H_3BO_3), which is an important boron chemical and based on the reaction of colemanite and sulfuric acid(H_2SO_4) at 88-92 ^oC [2]. The use of H_2SO_4 in boric acid production from colemanite ore has several problems cuse of decomposition of some minerals like calcium and magnesium in the colemanite ore by H_2SO_4. On the other hand, a form of gypsum that is a by-product of this process is released into the nature and causes environmental pollution[3, 4].

It was known that there are many studies in the literature about the investigation of the optimum condition for boric acid extraction from various minerals with using different leaching solutions. Some of them are showed in **Table 1.**

The aim of our study is to investigate the optimum condition for boric acid extraction from colemanite ore with using KH_2PO_4 solution in a mechanical stirrer system and declare an alternative process to produce the boric acid. The gypsum is not regarded as a by-product during the process. This prevents both environmental pollution and facilitates filtration. This is the most important advantage of our process when compared to other boric acid production processes. In addition, because it is a weak acid, KH_2PO_4 does not cause the dissolution of other minerals except boron. Thus, the boric acid purity is increased. For these reasons we believe that the investigation of the dissolution kinetics of colemanite in KH_2PO_4 solutions will be beneficial to the solution of these problems appeared during boric acid production. No study was found including the dissolution kinetics of colemanite in KH_2PO_4 solutions. Reaction temperature, KH_2PO_4 concentration, stirring speed, solid to liquid ratio, and particle size of colemanite were selected as parameters. Taguchi method was used to determine the optimum conditions for boric acid extraction from colemanite.

Table 1: Earlier Studies on the Reactions of Various Minerals with Using Different Teaching Solutions

Boron Minerals	Leaching Solutions	References
Colemanite	Supercritical carbon dioxide	[5]
	Propionic acid	[6]
	Nitric Acid	[7]
	Carbon Dioxide	[4]
Tincal	Phosphoric Acid	[8]
	Sulphur dioxide	[9]
	Water	[10]
Ulexite	Sulphuric acid	[11]
	Phosphate acid	[12]
	Sulphur dioxide	[13]

2. MATERIALS AND METHOD

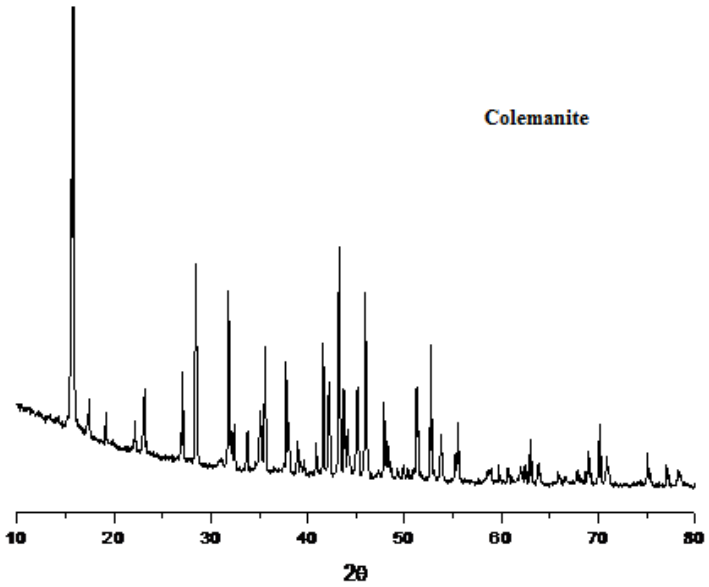
2.1. Materials

The colemanite ore used in the study was obtained from Emet-Mine (Kütahya-Türkiye). The samples was crushed, ground, and then sieved by using ASTM standard sieves to obtain 1550, 780, 390, 165 μm average size fractions. The chemical composition of the ore was determined by volumetric and gravimetric methods. The chemical analysis of these fractions is given in **Table 2.** There was no definitive trend between the B_2O_3 content and particle size ranges. XRD was obtained from the original sample and is shown in **Fig. 1**. Investigation of the diffractogram indicates that in the sample from Emet, apart from clay, the basic components are colemanite ($2CaO.3B_2O_3.5H_2O$) and Calcium oxide(CaO).

Table 2: Chemical Analysis of Colemanite Minerals

Component	B_2O_3	CaO	H_2O	SiO_2 and others
%	44.20	26.75	22.20	6.40

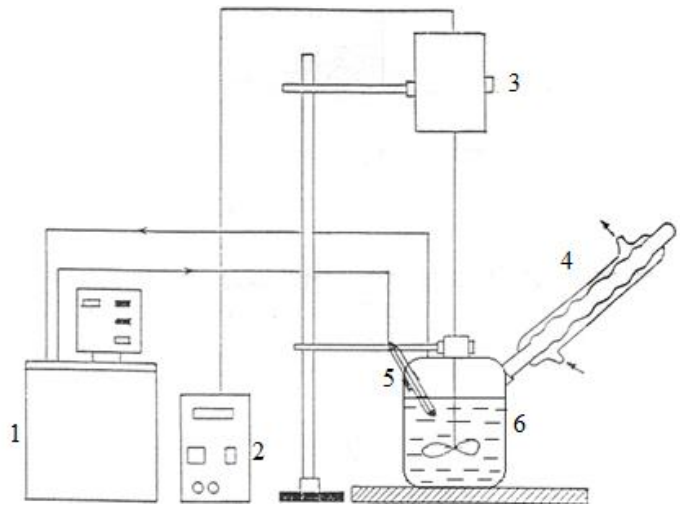
Fig. 1: X-ray Diffraction of Colemanite Ore



2.2. Experimental Procedure

A 500 mL jacketed glass reactor furnished with reflux condenser, a mechanical stirrer with tachometer for mixing, a fixed temperature circulator in order to keep the reaction temperature constant were used in the experiments carried out under atmospheric conditions. Solid material in accordance with the determined solid-liquid ratio was placed in the reactor and 200 mL KH_2PO_4 solution was added. After the reactor contents reached the determined temperature the experiment started. The experimental setup is shown in **Fig. 2.**

Fig. 2: Experimental Setup Used for the Dissolution Process



1- Fixed temperature circulator, 2- Speed control tachometer , 3- Mechanical stirrer, 4- Reflux condenser,
5- Thermometer, 6- Jacketted glass reactor

2.3 Statistical Method

The optimization of dissolution conditions of the ores has a great importance in the industrial fields. For this reason a lot of studies have been made by researchers using various methods such as Taguchi[12, 14, 15] the factorial experimental design [8, 16-18] and the orthogonal central composite design [19, 20].

In this study, the Taguchi method was used as the optimization method. Keeping experimental costs to a minimum, the advantage of the Taguchi method over other classic experimental design methods is that as the performance value is brought to the mean target,variability around the target is minimized and the optimum conditions obtained in the laboratory environment can be obtained in the true production environment [21].

During Taguchi method, the following stages were completed [22, 23].

- Determination of performance characteristic and choice of process parameters
- Determination of parameter levels for the process and possible internal interaction between parameters
- Choice of appropriate orthogonal pattern and insertion of parameters
- Completion of experiments in the orthogonal pattern
- Calculation of performance characteristic
- Analysis of experimental results using ANOVA and performance characteristic
- Determination of optimum levels of process parameters
- Confirmation of optimum conditions with confirmation experiments

During identification of optimum levels of parameters, the performance statistical formula for the three characteristics above are:

For situations where larger is better:

$$SN_L = -10Log\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_1^2}\right] \quad (1)$$

For situations where smaller is better:

$$SN_S = -10Log\left[\frac{1}{n}\sum_{i=1}^{n} Y_i^2\right] \quad (2)$$

For better target value:

$$SN_S = -10Log\left[\sum_{i=1}^{n} Y_i^2/S^2\right] \quad (3)$$

If the aim of the process is to reach maximum value, the parameter levels making the SN_L value maximum are optimum. If the aim is to reach a minimum, in this situation the parameter levels making SN_S maximum are optimum.

If the experiment with optimum study conditions in the Taguchi method is not in the experimental plan, in addition to performance value, (Eq.(4))may be used as an additional model to estimate the dissolution percentage.

$Y_i = \mu + X_i + e_i$ (4)

As the Y_i calculated from experimental results is a point estimate, to determine whether the additional model is sufficient, a confidence interval for estimate error should be created. The confidence interval is found using the equation below [24].

$$S_e = \pm 2\sqrt{\left[\frac{1}{n_0}\right] * \sigma_e^2 * \left[\frac{1}{n_r}\right]\sigma_e^2} \quad (5)$$

$$\sigma_e^2 = \frac{\text{sum of sequares due to error}}{\text{degrees of freedom for error}} \quad (6)$$

$$\frac{1}{n_o} = \frac{1}{n} + \left[\frac{1}{n_{Ai}} - \frac{1}{n}\right] + \left[\frac{1}{n_{Bi}} - \frac{1}{n}\right] + \left[\frac{1}{n_{Ci}} - \frac{1}{n}\right] \quad (7)$$

If the estimate error is outside these limits, the additional model may not be sufficient. If the opposite is true, the additional model is sufficient. If the experimental results are given as percentage, before using the Y_i equation, omega transformation of obtained percentage values is performed [23]. Later the omega value estimated for optimum conditions from obtained values is found using this equation:

$$\Omega(db) = -10Log\left(\frac{1}{P} - 1\right) \quad (8)$$

After calculations, the same equation is used for inverse transformation.

In this study parameter levels were determined in light of preliminary experiments and inserted into an orthogonal experimental design. To determine uncontrolled effects on the process (noise sources), experiments were repeated twice at separate times under the same conditions and larger, better performance characteristics were chosen as optimization criteria (Eq.(1)).

As the value calculated from (Eq.(4)) is a point estimate, to understand whether the additive model is sufficient or not, the confidence interval for the estimate error was calculated using (Eq.(5)).

3. RESULTS AND DISCUSSION

3.1. Dissolution Reactions

The reaction taking place in the solution can be written as follows [25]:

$$4\ KH_2PO_{4\ (aq)} \longrightarrow 4\ K^{+}_{(aq)} + 4H_2PO_4^{-1}{}_{(aq)} \quad (1)$$

$$4\ H_2PO_4^{-1}{}_{(aq)} + 4\ H_2O_{(aq)} \longleftrightarrow 4\ H_3O^{+}_{(aq)} + 4\ HPO_4^{-2}{}_{(aq)} \quad (2)$$

$$4\ K^{+}_{(aq)} + 2HPO_4^{-2}{}_{(aq)} \longrightarrow 2K_2HPO_{4\ (aq)} \quad (3)$$

When colemanite ore is added to the KH_2PO_4 solutions, the reaction taking place in the solution can be written as follows:

$$2CaO.3B_2O_3.5H_2O_{(s)} + 4H_3O^{+}_{(aq)} \longrightarrow 2Ca^{+2}_{(aq)} + 6\ H_3BO_{3\ (aq)} + 2\ H_2O_{(l)} \quad (4)$$

$$2Ca^{+2}_{(aq)} + 2\ HPO_4^{-2}{}_{(aq)} \longrightarrow 2\ [CaHPO_4.2H_2O]_{(s)} \quad (5)$$

The total reaction is as follows:

$$2CaO.3B_2O_3.5H_2O_{(s)} + 2\ KH_2PO_{4\ (aq)} + 6H_2O_{(aq)} \longrightarrow 2\ [CaHPO_4.2H_2O]_{(s)} + 4\ H_3BO_{3(aq)} + 2\ K_2HPO_{4(aq)} \quad (6)$$

3.2. Optimization Experiment Results

The number of parameters determined in optimization studies was 5, and for each parameter 4 different levels were investigated. The chosen parameter values and levels for the experiments are given in **Table 3.** In accordance with this situation, an $L_{16}(4^5)$ factorial fractional experimental design plan was chosen. According to this experimental plan, the heterogeneous mixture obtained at the end of experiments was filtered with a vacuum pump. The filtrate was analyzed for B_2O_3 and the results of % B_2O_3 amounts in filtrate are given in **Table 4.**

Table3: Parameters and Their Ranges

Parameters	Values
A: Reaction Temperature, K	303, 313, 323, 333
B: Particle size, µm	1550, 780, 390, 165
C: Solid/liquid ratio, g mL^{-1}	0.02, 0.04, 0.08, 0.17
D: KH_2PO_4 concentration, M	0.50, 1.00, 1.50, 2.00
E: Time, min	15, 30, 45, 60

Table 4: $L_{16}(4^5)$ Experimental Plan and Results of Experiments

Experiment No	Parameter Levels					1st Series Experiments	2nd Series Experiments
	A	B	C	D	E	B_2O_3(%)	B_2O_3(%)
1	1	1	1	1	1	15.83	22,17
2	1	2	2	2	2	31.95	32.61
3	1	3	3	3	3	35.33	38.48
4	1	4	4	4	4	54,45	55.60
5	2	1	2	3	4	50.14	49.32
6	2	2	1	4	3	90.87	87.20
7	2	3	4	1	2	22.03	21.03
8	2	4	3	2	1	51.76	49.32
9	3	1	3	4	2	48.81	44.06

10	3	2	4	3	1	28.40	28.63
11	3	3	1	2	4	99.8	99.90
12	3	4	2	1	3	67.00	60.83
13	4	1	4	2	3	28.15	30.42
14	4	2	3	1	4	39.93	41.93
15	4	3	2	4	1	99.85	99.95
16	4	4	1	3	2	93.96	98.65

3.3. Statistical Analysis

The data were analyzed using a statistical program. For effective parameters on the solution process and to find their confidence levels, variance analysis (ANOVA) was completed. ANOVA is used to understand whether process parameters are statistically significant or not. To determine the process parameters with most significant effect on the solution process the *F* test was used. For each parameter of the process, the *F* value is the ratio of the total square of the deviation error to the total square of the error. Generally parameters with larger *F* value have greater effect on the process.

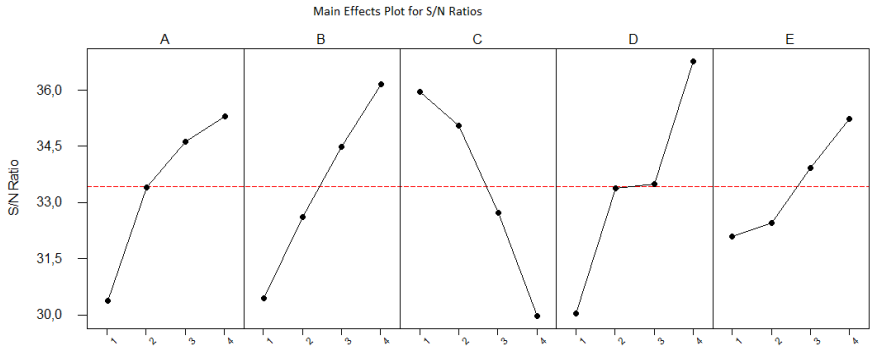
The ANOVA analysis results for B_2O_3 are given in **Table 5.** When **Table 5** is investigated, the highest *F* values were observed to be for solid/liquid ratio and KH_2PO_4 concentration.

Table 5: ANOVA Table for B_2O_3 in Optimization Experiment

Parameter	SS	Df	MS	F
A	4188.7	3	1396.2	271.07
B	5004.4	3	1668.1	323.86
C	5511.9	3	2837.3	550.85
D	5276.9	3	1759.0	341.49
E	790.1	3	263.4	51.13
Error	82.4	16	5.2	
Total	23854.4	31		

To obtain optimal dissolution performance, for situations where larger is better performance characteristic (Eq.(1)) has been taken for dissolution of B_2O_3. The order of graph in **Fig 3** is according to the degrees of he influences of parameters on the performance characteristics. The effect of each parameter on the optimization criteria for B_2O_3 is shown in **Fig. 3.**

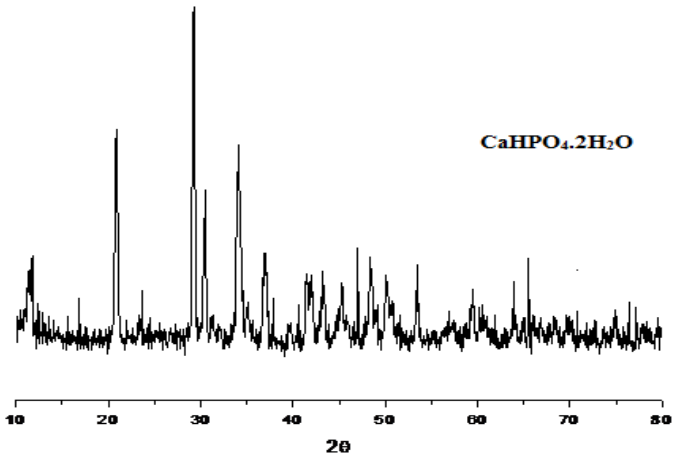
Fig. 3: The effect of each parameter on the optimization criteria for B_2O_3



The optimal level of process parameter is the level with the highest SN ratio value calculated by Eq.(1). The numerical value of a maximum point corresponds to the best value for that parameter in **Fig. 3**. These values are seem to be A4 (333 K), B4(165 μm), C1(0.02 g mL^{-1}), D4(2 M), E4(60 min).

X-ray diffractogram of the solid product obtained from the optimum condition is shown in **Fig. 4**. As seen from the diffractogram of the solid product in **Fig. 4**, the basic component is Calcium Hydrogen Phosphate Hydrate – $CaHPO_4$ 2.H_2O.

Fig. 4: X-ray Diffraction of Solid Product



4. CONCLUSION

- The effective parameters on the dissolution of colemanite ore in KH2PO4 solutions are solid/liquid ratio and KH_2PO_4 concentration
- The optimum conditions for the maximum B_2O_3 dissolution was determined as follows: temperature, 333 K; particle size, 165 μm; time, 60 min; solid-liquid ratio, 0.02 g/mL and KH_2PO_4 concentration, 2 M.
- Under these optimum conditions 98% B_2O_3 dissolved.
- As seen from the diffractogram of the solid product in **Fig. 4**, the basic component is Calcium Hydrogen Phosphate Hydrate – $CaHPO_4$ 2.H_2O.
- As a result, the additive model is appropriate to describe the dependence of the process on the studied parameters.

Nomenclature

N	Repetitive experiment number
y_i	Performance value
SN_L,SN_S,SN_{Tk}	Performance statistics
Y_i	Estimated performance value of experiment
μ	Total average of performance value
X_i	Total efficiency size of parameter levels used in the experiment
e_i	experimental error
S_e	confidence interval for effects of parameters
n_{A1}	experiment number belonging to A parameter
σ_e^2	Variance of error

REFERENCES

[1] Tombal TD, Özkan ŞG, Ünver İK, Osmanlıoğlu AE. Bor bileşiklerinin özellikleri, üretimi, kullanımı ve nükleer reaktör teknolojisinde önemi/Properties, production, uses of boron compounds and their importance in nuclear reactor technology. Journal of Boron. 2016;1:86-95.

[2] Gönen M, Nyankson E, Gupta RB. Boric Acid Production from Colemanite Together with ex Situ CO2 Sequestration. Industrial & Engineering Chemistry Research. 2016;55:5116-24.

[3] Davies T, Colak S, Hooper R. Boric acid production by the calcination and leaching of powdered colemanite. Powder Technology. 1991;65:433-40.

[4] Ata ON, Colak S, Çopur M, Çelik C. Determination of the optimum conditions for boric acid extraction with carbon dioxide gas in aqueous media from colemanite containing arsenic. Industrial & engineering chemistry research. 2000;39:488-93.

[5] Budak A, Gönen M. Extraction of boric acid from colemanite mineral by supercritical carbon dioxide. The Journal of Supercritical Fluids. 2014;92:183-9.

[6] Kuskay B, Bulutcu A. Design parameters of boric acid production process from colemanite ore in the presence of propionic acid. Chemical Engineering and Processing: Process Intensification. 2011;50:377-83.

[7] Yeşilyurt M. Determination of the optimum conditions for the boric acid extraction from colemanite ore in HNO 3 solutions. Chemical Engineering and Processing: Process Intensification. 2004;43:1189-94.

[8] Durak H, Genel Y, Çalban T, Kuşlu S, Çolak S. Optimization of the Dissolution of Tincal Ore in Phosphoric Acid Solutions at High Temperatures. Chemical Engineering Communications. 2015;202:245-51.

[9] Karagöz Ö, Çopur M, Kocakerim MM. Tinkalden Borik Asit Üretiminin Optimizasyonu. 10 Ulusal Kimya Mühendisliği Kongresi2012. p. 299-300.

[10] Abalı Y, Bayca SU, Edgunlu G. Optimization of pure borax pentahdrate extraction from calcined tincal. Physicochemical Problems of Mineral Processing. 2015.

[11] Tunc M, Celik C, Colak S, Kocakerim M. Determination of optimum conditions for dissolution of ulexite in sulphuric acid solutions. Transactions-Institution of mining and metallurgy Section C Mineral processing & extractive metallurgy. 1999;108.

[12] Doğan TH, Yartaşı A. Optimization of Dissolution of Ulexite in Phosphate Acid Solutions. Journal of the Chemical Society of Pakistan. 2014;36.

[13] Küçük Ö, Kocakerim MM. Optimization of dissolution of ulexite in water saturated with sulphur dioxide. Chemical Engineering and Processing: Process Intensification. 2005;44:1005-11.

[14] Beşe AV, Borulu N, Çopur M, Çolak S, Ata ON. Optimization of dissolution of metals from Waelz sintering waste (WSW) by hydrochloric acid solutions. Chemical Engineering Journal. 2010;162:718-22.

[15] Yeşilyurt M, Çolak S, Çalban T, Genel Y. Determination of the optimum conditions for the dissolution of colemanite in H3PO4 solutions. Industrial & engineering chemistry research. 2005;44:3761-5.

[16] Tokkan D, Çalban T, Kuşlu S, Çolak S, Dönmez Bn. Optimization of microwave-assisted removal of lead from anode slime in triethanolamine solutions. Industrial & Engineering Chemistry Research. 2012;51:3903-9.

[17] Çalban T, Kuşlu S, Çolak S. Precipitation conditions of Chevreul's salt from synthetic aqueous CuSO4 solutions. Chemical Engineering Communications. 2009;196:1018-29.

[18] Yalvac Can M, Yildiz E. Phosphate removal from water by fly ash: Factorial experimental design. Journal of hazardous materials. 2006;135:165-70.

[19] Çavuş Dişli F, Kuşlu S, Çalban T, Çolak S. Optimization of Dissolution of Ulexite Mineral at High Temperature in Aqueous and Borax Pentahydrate Solutions Saturated with Carbon Dioxide. Asian Journal of Chemistry. 2014;26.

[20] Çalban T, Çolak S. Optimization of Pure Copper Powders Production from Leach Solutions Containing Copper. Asian Journal of Chemistry. 2014;26.

[21] Çopur M. An Optimization Study of Dissolution of Zn and Cu in ZnS Concentrate with HNO~ 3 Solutions. Chemical and biochemical engineering quarterly. 2002;16:191-8.

[22] Nian C, Yang W, Tarng Y. Optimization of turning operations with multiple performance characteristics. Journal of Materials Processing Technology. 1999;95:90-6.

[23] Phadke MS. Quality engineering using robust design: Prentice Hall PTR; 1995.

[24] Roy R. A primer on the Taguchi method, 1990. New York. 1990.

[25] Nemodruk AA, Karalova ZK. Analytical chemistry of boron 5 B 10.811. 1965.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

INVESTIGATION OF THE EFFECTS OF RENEWABLE ENERGY SOURCES ON INTERCONNECTION NETWORKS

DOI: 10.17261/Pressacademia.2017.618
PAP-WCTIE-V.5-2017(56)-p.410-419

Mustafa Tekin[1], O. Fatih Kececioglu[2], Ceyhun Yildiz[3], Ahmet Gani[4], Hakan Acikgoz[5], Mustafa Kesenek[6] Mustafa Sekkeli[7]
[1]Department of Electrical and Electronics Eng., Faculty of Eng. and Architecture, K.Maraş Sütçü İmam University. mustafatekin@ksu.edu.tr
[2]Department of Electrical and Electronics Eng., Faculty of Eng. and Architecture, K.Maraş Sütçü İmam University. fkececioglu@ksu.edu.tr
[3]Graduate School of Natural and Applied Sciences, K.Maraş Sütçü İmam University. cyildiz@ksu.edu.tr
[4]Department of Electrical and Electronics Eng., Faculty of Eng. and Architecture, K.Maraş Sütçü İmam University. agani@ksu.edu.tr
[5]Graduate School of Natural and Applied Sciences, Kilis 7 Aralık University. hakanacikgoz@ksu.edu.tr
[6]TEDAŞ General Directorate, K.Maraş. kesenekksu@gmail.com
[7]Department of Electrical and Electronics Eng., Faculty of Eng. and Architecture, K.Maraş Sütçü İmam University. msekkeli@ksu.edu.tr

ABSTRACT

Demand for electricity is increasing due to population growth, the development of technology and the level of prosperity of communities. Coal, natural gas, oil, etc., which are called traditional sources can no longer meet this demand. In addition, these resources are harmful to the environment and have high costs. For all these reasons, the trend towards alternative energy sources is increasing. These sources are called renewable energy sources. Solar, wind, water, geothermal etc. as examples of renewable energy sources. In addition to producing enough energy, it is also necessary to use the produced energy efficiently. This is defined as the energy quality (power quality). The power quality parameters are defined to keep the obtained electric energy within certain limits in terms of quality. These parameters include; harmonic, voltage flicker, voltage deviation etc. Through this study; it has been analyzed, with the increasing number of solar power plants connecting to the network, what kind of changes occurred in the electricity network and how the solar power plants affect the network in terms of energy quality. The analysis was carried out by evaluating the one month data from a 5 MW solar power plant.

Keywords: Renewable energy, solar power plant (SPP), power quality, harmonic, voltage deviation.

1. INTRODUCTION

With each passing day the development of technology and the increase of the human population increases the energy demand (electricity energy). Figure-1 shows this situation clearly in Turkey scale. When Figure-1 is examined; while the energy demand of Turkey is 40,000 MW in 2015, it is expected that this demand will be about 60,000 MW in 2024.(Tekin et al., 2016a)

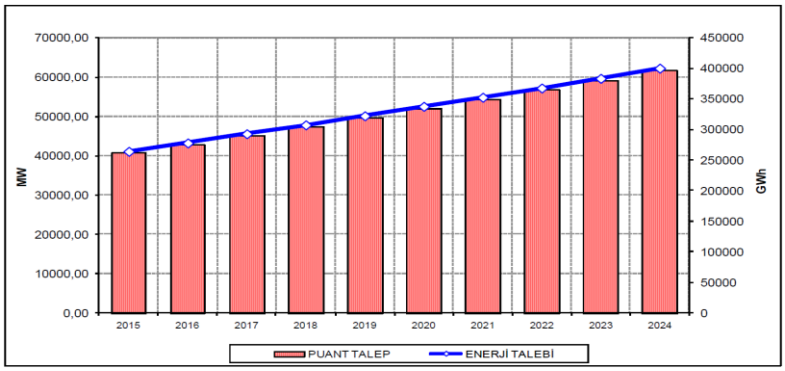
Energy circles in the world are turning to alternative energy sources because traditional resources (petroleum, coal, natural gas, etc.) are both harmful to nature and depletion of reserves. Hydro, solar, wind, biomass and geothermal energy can be given as examples to these alternative sources and clean energy production sources.(Tekin et al., 2016a)

In addition to the production of electricity, it is also important to use the produced energy efficiently and in good quality; because the lack of quality of electricity affects both producers and consumers as well as the economy in the negative. The concept of power quality has been introduced to reduce these negatives the least and to determine the sources that affect the quality of electricity in a negative way more quickly.

Power quality is a set of constraints that allow devices to behave in the desired way without performance and lifetime losses. The poor quality of the energy can cause equipment such as microprocessors which are sensitive to current or voltage imbalance used in the industry and the equipment used by the consumer to be damaged.(Kocaman and Özdemir, 2009) Energy quality distortions classified as; current and voltage harmonics, voltage sags & swells, flicker, neutral-to-ground potential difference, current and voltage imbalance between phases, frequency change, etc.(Şahin et al., 2014)

Figure-1: Electricity Energy Demand Forecast of Turkey Between 2015-2024

(Türkiye Elektrik Enerjisi 5 Yıllık Üretim Kapasite Projeksiyonu, 2015)



In order to avoid the above mentioned inconveniences, power quality parameters have been controlled all over the world by introducing restrictions by regulations. "Regulation on Procurement Duration, Commercial and Technical Qualities of Electricity Power Supplied in Electricity Market Distribution System" covers these limitations for Turkey.

Based on the importance of the above-mentioned issues, the 1-month electrical data from a 5 MW Solar Power Plant (SPP) was taken with the help of a power analyzer and from this data it was evaluated how this SPP influenced the network in terms of harmonics and voltage deviations.

2. LITERATURE REVIEW

Gabriel Tanasescu and colleagues (2016) investigated how a 5 MW solar power plant affected the interconnection network in terms of flicker. In another study, Semih Hürmeydan et al. (2016) investigated how small-scale SPPs affect the network's in terms of power quality (harmonic and voltage imbalance) at low voltage levels. Pertti Pakonen et al. (2016) they evaluated the voltage flicker, total harmonic distortion and total demand distortion values by connecting a 6 kW SPP to the network. And they observed that flicker increased in cloudy weather. They also emphasized the need for good design of the inverter so that the SPP does not adversely affect the power quality. Amirullah et al. (2015) investigate how total harmonic distortion and total demand distortion are affected by the coupling of SPPs to the networks to which residential and industrial loads are separately connected. As a result, both values are understood to comply with IEEE standards. Carlos Gonzalez et al. (2012) investigate how overvoltage, undervoltage and voltage imbalance phenomena are affected by the connection of a feeder where residential loads are connected together with SPP at low voltage level. Minas Patsalides and colleagues (2012) have modeled the system to analyze harmonic variations by connecting the increasing number of SPP to the distribution network. And as a result of the study, they arrived at the conclusion that the SPPs did not pose a threat to the network and that the harmonics were within the limit values. Pedro A. B. Block et al. (2014), in their study; they evaluated how to change the power quality parameters by connecting a 1 MVA SPP to the network. And at the end of the study, increases in current harmonics were observed at sunrise and sunset.

In our study, actual measured values were taken from the point (medium voltage level=36,000), which is connected to a 5 MW large-scale SPP interconnection network. In the direction of datas; daily, weekly and monthly changes in power quality (harmonics and voltage deviations) caused by SPP during the winter months in the network were investigated.

3. DATA AND METHODOLOGY

3.1. Power Quality Parameters

3.1.1. Harmonics

Harmonics can be defined as all electrical signals originating from nonlinear loads or from unbalanced generators, except for network frequency (50 Hz). Harmonics (3, 5, 7, ...), which are single times the main frequency, are called odd harmonics, and which are double times the main frequency (2,4,6 ...) are called even harmonics. Harmonics which smaller than 50 Hz are called subharmonics, and which is not a complete rigid of 50 Hz are called interharmonics (for example, 170 Hz).(Tekin et al., 2016b)

Harmonics are one of the most important of energy quality parameters. Due to the economic effects of harmonics and the ability to adversely affect network operation, power systems must be monitored in detail.(Şahin et al., 2014)

Harmonics do not only distort the waveforms, but also they can lead to overheating of transformers and conductors; deterioration of di-electric materials such as cables and capacitors, heavy damage, reduced efficiency, reduced economic

periods of use; increased losses in power lines; communication systems near the power lines may be adversely affected, overheating in electric machines, noisy operation and mechanical oscillations.(Keçecioğlu et al., 2015) Harmonic contamination can often be carried over the electricity distribution system and affect neighboring facilities in the same system.(Şahin et al., 2014)

The most important parameter defined for harmonic distortion is Total Harmonic Distortion (THD). THD is an indication of an impairment in an electrical signal. The boundary values of this variable are specified in IEEE 519-2014 standards. THD formula for voltage is given in Equation-1.(Tekin et al., 2016b)

$$THB_V = \frac{\sqrt{\sum_{i=2}^{40} V_i}}{V_1} \times 100 \quad \text{(EPDK, 2016)} \qquad (1)$$

The boundary values of voltage harmonics in the IEEE 519-2014 standard are shown in Table-1.

Table 1: Boundary Values for Voltage Harmonics (IEEE 519-2014 Standard, 2014)

Barrier Voltage at Common Port (V)	Odd Harmonic (%)	Total Harmonic Distortion THD (%)
V≤1 kV	5	8
1 kV<V≤69 kV	3	5
69 kV<V≤161 kV	1.5	2.5
161 kV<V	1	1.5*
* High-voltage systems can have up to 2.0% THD where the cause is an HVDC terminal whose effects will have attenuated at points in the network where future users may be connected.		

3.1.2. Voltage Deviation (ΔV)

The voltage deviation is that the network voltage goes below or exceeds the rated voltage. The limit value defined for voltage deviation in Turkey is ± 10%. 95% for a week of values taken from the system should remain in this limit. It is the responsibility of the distribution company to keep the mains voltage at these values. There can be many reasons for voltage deviation; but this study focuses on the effects of grid-connected SPP on this parameter.

In Turkey, the limit values for the voltage deviation allowed to be formed in the network by connecting a SPP to the network are given in table-2. According to the table-2, the threshold allowed when connecting and disconnecting SPP is ± 3%.(Şimsek and Bizkevelci, 2015)

Table 2: Voltage Deviation Boundary Values Applied in Turkey (Şimsek and Bizkevelci, 2015)

	Limit Values
SPP Connecting	±3.3
SPP Disconnecting	±3.3

The SPP raises the mains voltage when connecting to the network, and reduces the mains voltage when disconnecting from the network. Therefore, the point at which SPP connects to the network is important, and this connection point needs to be chosen according to the worst cases. For this reason, the analysis of the feeders to which SPP is connected should be done with great care.

4. FINDINGS AND DISCUSSIONS

4.1. Harmonic Measurements

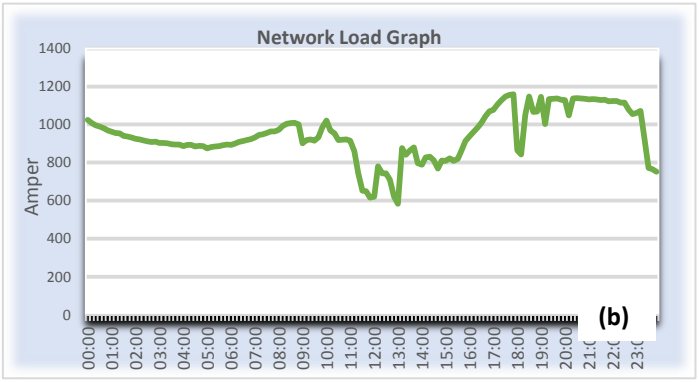
In this part of the study, the data which received from the 5 MW SPP that was connected to the network at medium voltage level, between the 01.01.2017 and 31.01.2017 were evaluated. Evaluation was made in terms of voltage harmonics and voltage deviations. One-day, one-week and one-month graphs of harmonics and voltage deviations were created. Thus how the SPP that was connected to the interconnected network, influences the network in terms of voltage deviations and harmonics has been examined. In the study, the harmonics values except for 3, 5, 7th were not considered because they were very small.

4.1.1. Daily Harmonic Measurements

Figure-3-(a, b, c, d). a) 3rd Harmonic, b) 5th Harmonic, c) 7th Harmonic, d) THD Figures

The figures in this section of the study were obtained from the SPP's datas that were taken at 10 minute intervals on 16.01.2017. Figure 2, which shows the load flow in the network, is provided to help evaluate the harmonic distortions. When Figure-2 is examined, it is observed that the system current falls to the lowest value (600 A) when time is between 12:00-13:00. It is also clear that consumption increased after 17:00.

Figure 2: Network Load Graph

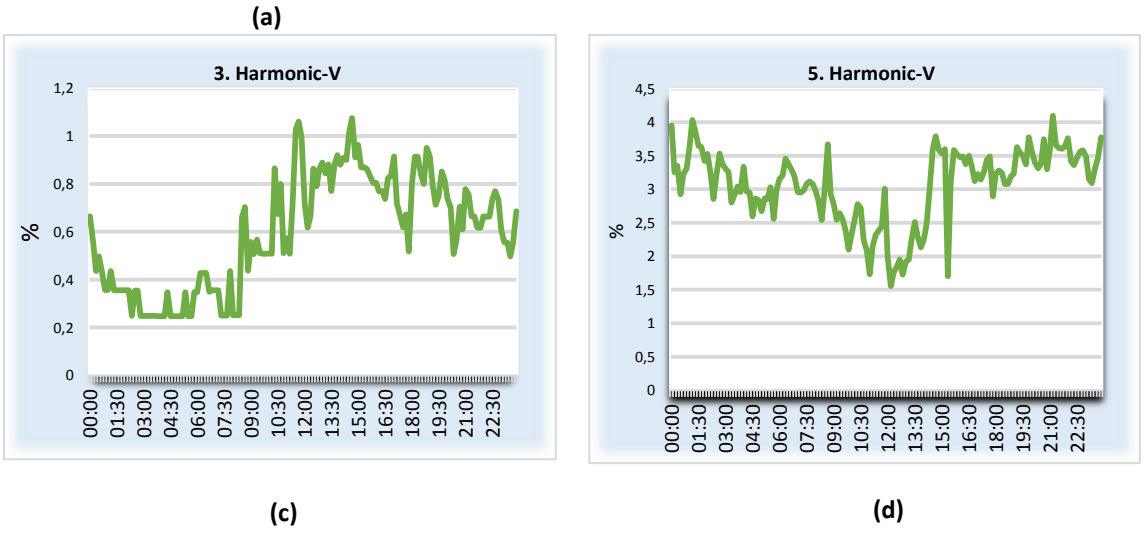


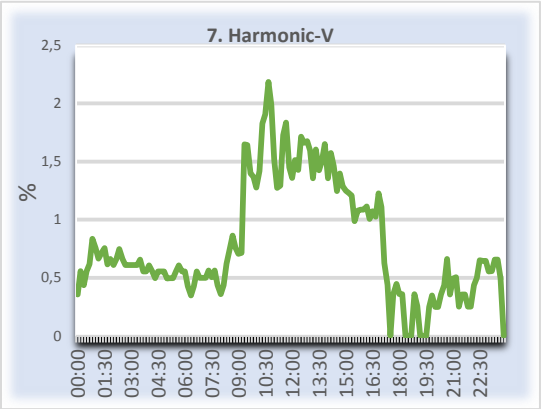
3, 5, 7th harmonics and THB values are seen in Figure-3-(a, b, c, d) respectively. According to these graphs; it is seen that the 3rd harmonic during the day does not exceed the limit value of 3% for medium voltage level. It is understood that the 3rd harmonic reaches its maximum value at noon hours while it is at the lowest values at night hours. It is thought that SPP causes this harmonic to reach its maximum value at noon. Because SPP is producing the most intensive energy at that time.

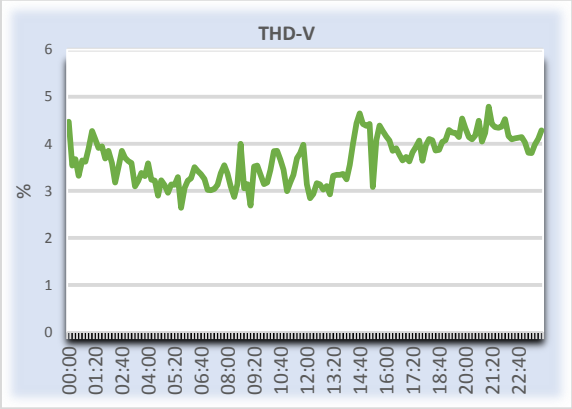
It is seen that from the figüre-3-b, 5th harmonic generally exceed the limit value (3%) throughout the day. It is observed that 5th harmonic drops below the limit (3%) between 08:00 and 13:30. The reason of this situation is evident from figure-2 (network load graph). That is, the hours that the 5th harmonic falls below the limit value are the hours when the load is minimum. So that the load current reduction causes the decrease in the 5th harmonic.

It is seen that the 7th harmonic is observed at small values as 0,5% at the times when SPP is not in operation (18:00-07:00) and reaching 2% at the times when SPP is active during the day, but the limit value is not passed.

When the figure of THD is examined, it is observed that it changes between 3% and 4% in general during the day. Since the 5th harmonic is dominant in the system, there is a decrease in THD between 08:00-13:30 as in the 5th harmonic. However, THD has not exceeded the limit of 5% for medium voltage.







4.1.2. Weekly Harmonic Measurements

In this part of the article, the graphs, obtained from the weekly data which were received from the network on 01.01.2017 -

Figure-4-(a, b, c, d). a) 3rd Harmonic, b) 5th Harmonic, c) 7th Harmonic, d) THD Figures

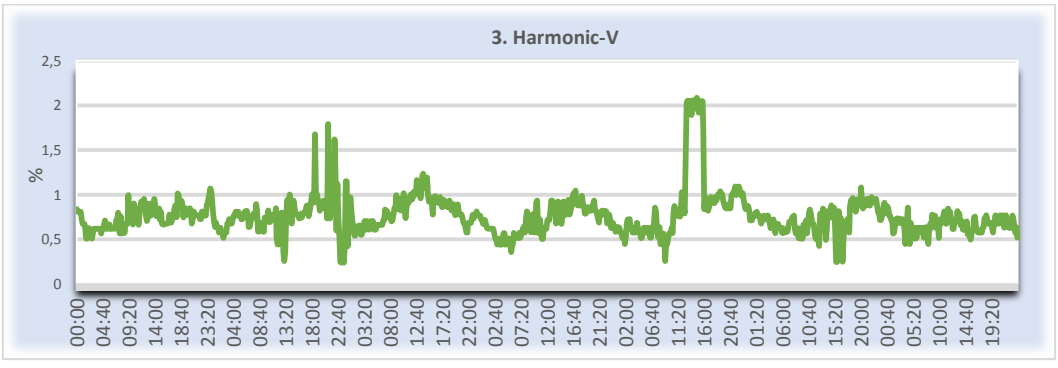
07.01.2017, were evaluated. 3, 5, 7th harmonics and THD values are seen in figure-4-(a, b, c, d) respectively.

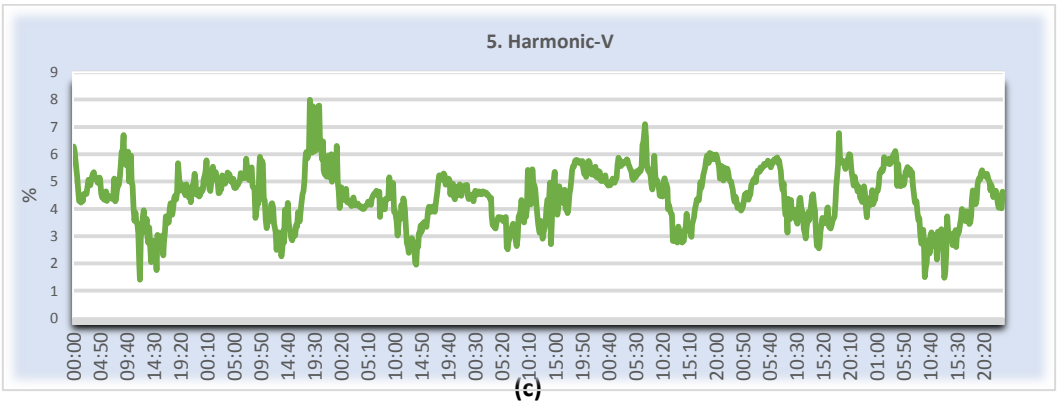
When looking at figüre-4-a; 3. harmonic varies from 0,5% to 1% in general, except for some sudden ups and downs. 3rd harmonic therefore provides standards for weekly measurements. As in the case of one-day measurements, it can be seen from the graph that the 3rd harmonic tends to increase at the times when SPP is in operation (07:30 - 18:00) in a weekly measurements.

When the 5th harmonic graph is examined for one week measurements; the 5th harmonic has exceeded limit values for medium voltage (3%), many times. As in the case of one-day measurements, it is understood that the 5th harmonic tends to decrease at times when the network current decreases. It is understood that the 5th harmonic does not comply with the standard that is 95% of the weekly measurements should not exceed the limit value. On some days it seems that the 5th harmonic reached as 8%, 7%.

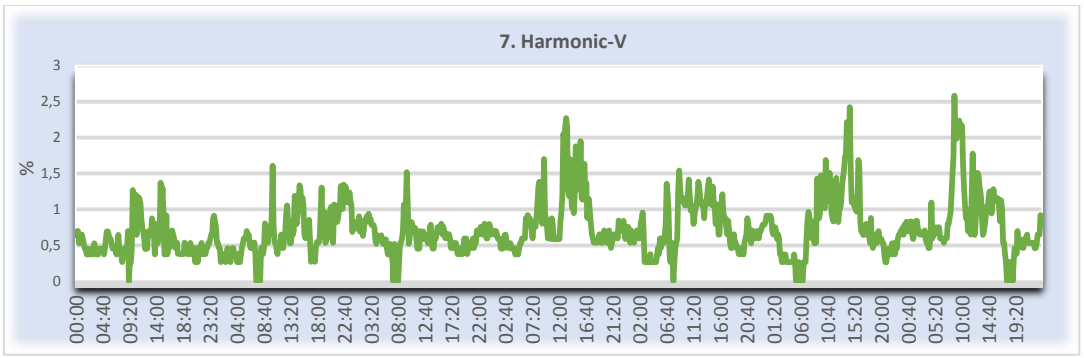
When one week's 7th harmonic graph is examined; it is understood that, as in the 3rd harmonic, the 7th harmonic is also increasing at the times when SPP is active. However, it appears that the 7th harmonic has never exceeded the 3% limit value that is defined in the standards. It is observed that the 7th harmonic reaches a maximum value of 2.5% within one week period.

When a weekly figure of THD is considered; it is seemed that, in parallel with the 5th harmonic, the THD has also exceeded the limit value (5%).

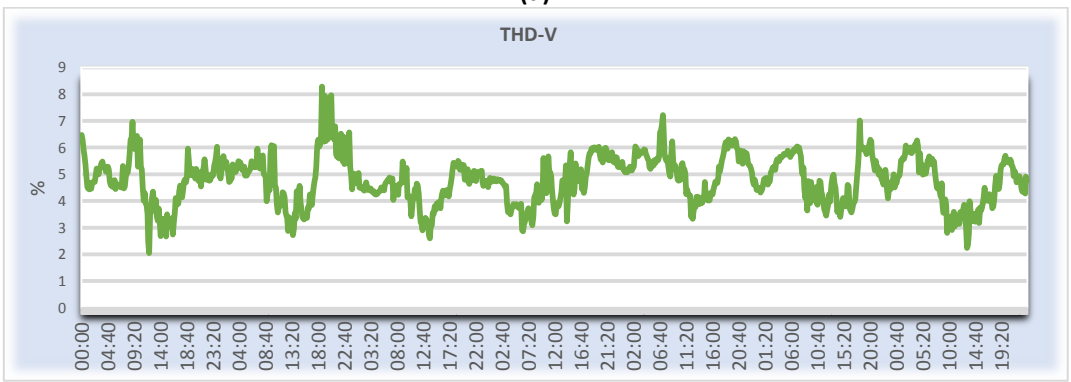




(c)



(d)



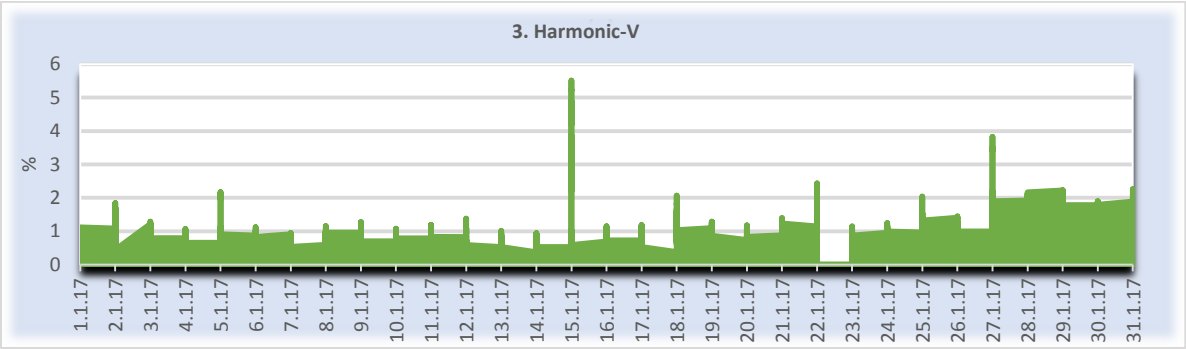
4.1.3. Monthly Harmonic Measurements

In this section of the article, the graphs, obtained from the monthly data which were received from the network on 01.01.2017-31.01.2017, were evaluated. 3, 5, 7th harmonics and THD values are seen in figure-5-(a, b, c, d) respectively.

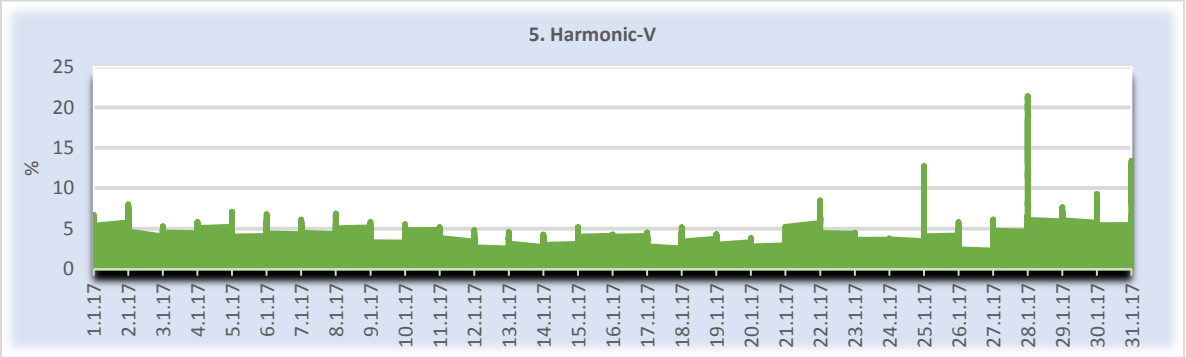
Figure-5-(a, b, c, d). a) 3rd Harmonic, b) 5th Harmonic, c) 7th Harmonic, d) THD Figures

From the area graph for the 3rd harmonic; it is understood that the 3rd harmonic has passed the border value only 2 times for 30 days. It is observed that the 3rd harmonic is generally around 1% for a month but sometimes it is about 2%.

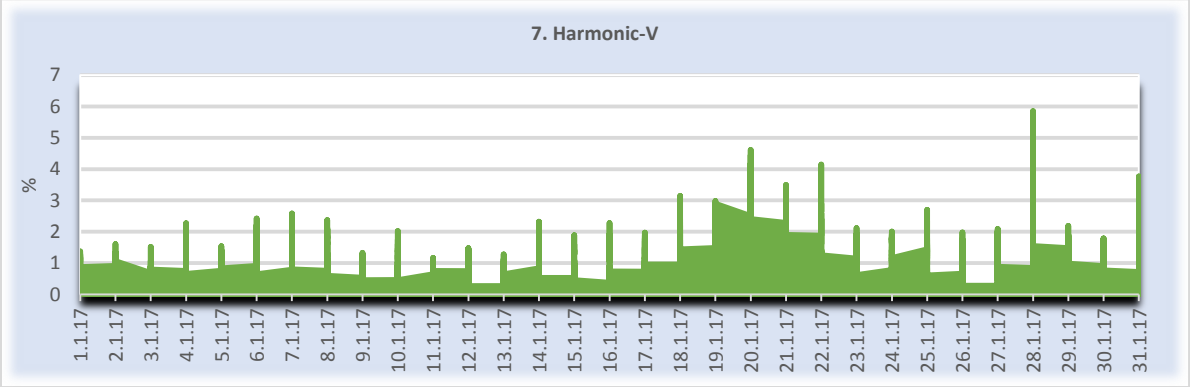
It is seemed that the 5th harmonic is generally observed at 5% level. In this case, it is observed that the standard value of the 5th harmonics (3%) has passed. From the 7th harmonic graph; it is understood that this harmonic is generally below 1%, but in some days it has increased to 2% and very rarely exceeded the limit value of 3%. But since these limitations crossings are too short, it does not affect the compliance of the 7th harmonic with the standards. Looking at the THD's one month graph; it is observed that this parameter is generally above 5%, even on some days it has risen up to a value of 13%, 23% for a very short period of time. This shows that THD does not comply with the standards.


3. Harmonic-V
%
0
1
2
3
4
5
6
1.1.17
2.1.17
3.1.17
4.1.17
5.1.17
6.1.17
7.1.17
8.1.17
9.1.17
10.1.17
11.1.17
12.1.17
13.1.17
14.1.17
15.1.17
16.1.17
17.1.17
18.1.17
19.1.17
20.1.17
21.1.17
22.1.17
23.1.17
24.1.17
25.1.17
26.1.17
27.1.17
28.1.17
29.1.17
30.1.17
31.1.17

(b)


5. Harmonic-V
%
0
5
10
15
20
25
1.1.17
2.1.17
3.1.17
4.1.17
5.1.17
6.1.17
7.1.17
8.1.17
9.1.17
10.1.17
11.1.17
12.1.17
13.1.17
14.1.17
15.1.17
16.1.17
17.1.17
18.1.17
19.1.17
20.1.17
21.1.17
22.1.17
23.1.17
24.1.17
25.1.17
26.1.17
27.1.17
28.1.17
29.1.17
30.1.17
31.1.17

(c)


7. Harmonic-V
%
0
1
2
3
4
5
6
7
1.1.17
2.1.17
3.1.17
4.1.17
5.1.17
6.1.17
7.1.17
8.1.17
9.1.17
10.1.17
11.1.17
12.1.17
13.1.17
14.1.17
15.1.17
16.1.17
17.1.17
18.1.17
19.1.17
20.1.17
21.1.17
22.1.17
23.1.17
24.1.17
25.1.17
26.1.17
27.1.17
28.1.17
29.1.17
30.1.17
31.1.17

(d)

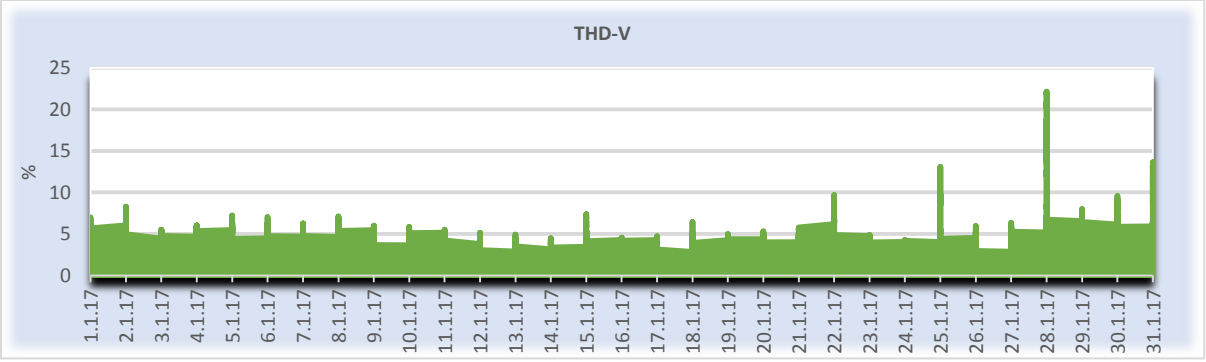

THD-V
%
0
5
10
15
20
25
1.1.17
2.1.17
3.1.17
4.1.17
5.1.17
6.1.17
7.1.17
8.1.17
9.1.17
10.1.17
11.1.17
12.1.17
13.1.17
14.1.17
15.1.17
16.1.17
17.1.17
18.1.17
19.1.17
20.1.17
21.1.17
22.1.17
23.1.17
24.1.17
25.1.17
26.1.17
27.1.17
28.1.17
29.1.17
30.1.17
31.1.17

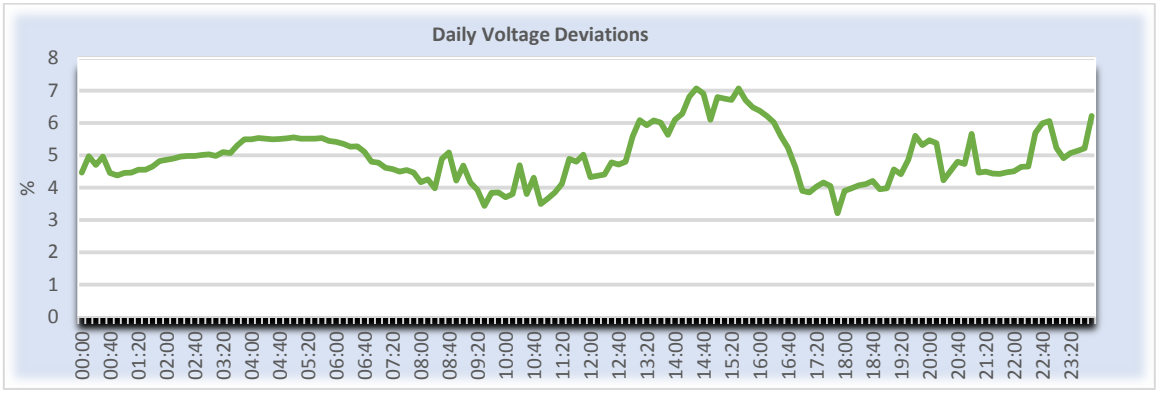
Figure-6-(a, b, c). a) Daily Voltage Deviation, b) Weekly Voltage Deviation, c) Monthly Voltage Deviation

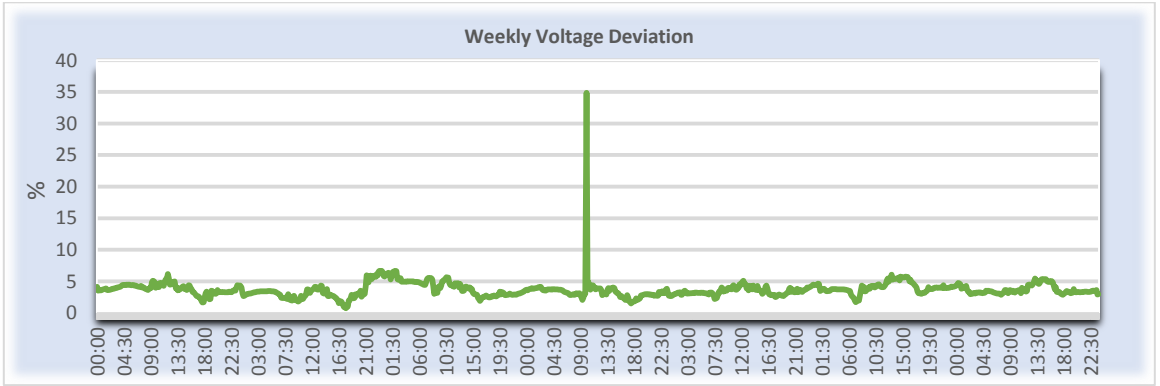
4.2. Voltage Deviation Measurements

In this part of article figure-6-(a, b, c); which is three separate graphs obtained from one day, one week and one month measurements, is examined to see the deviations over the grid voltage.

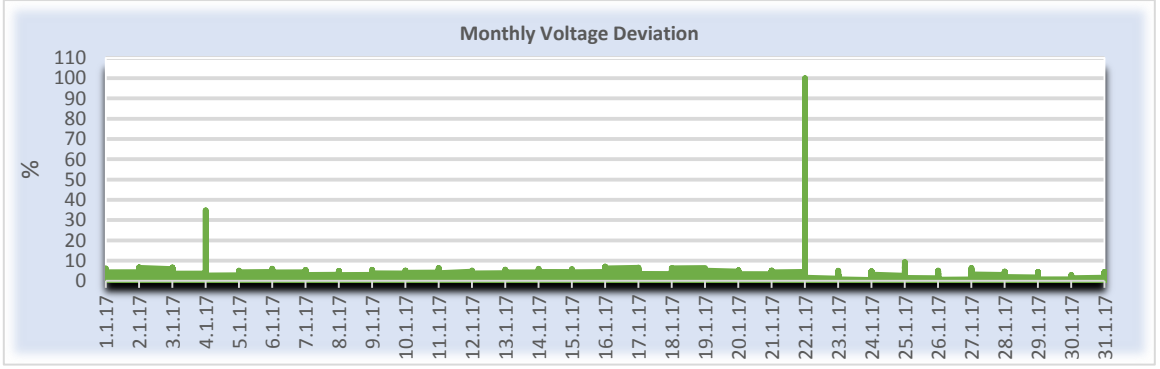
From a daily chart; it is understood that a 1% change in system voltage is caused when the SPP enters and exits the circuit. Therefore, it is seen that the SPP has not exceeded the allowable 3% change in the system and stayed at the limit values, and it is understood that the network does not exceed the 10% deviation which is the limit value for the network. Looking at the weekly and monthly graphs; it is observed that the system voltage is below the limit value (10%) except for short-term voltage drop.

(a)





(c)



5. CONCLUSIONS

In this study, it was investigated how the 5 MW SPP which was connected to the interconnected system from the medium voltage level, affects network's harmonics and voltage deviations. Within the scope of the research, electrical data were obtained from the place, where SPP was connected to the network, between 01.01.2017- 31.01.2017. One day, one week and one month of harmonics and voltage deviations graphs were generated from these data. With interpretation of these graphs yielded the following results:

- As a result of examining the daily graphs; it has been understood that the 3rd and 7th harmonics are in accordance with the limit values in the standard throughout the day. It is seen that the 3rd and 7th harmonics show their maximum values during the daytime hours (when SPP is active) while at the lowest values during night time. The 5th harmonic generally exceed the threshold value; but it was found to be the lowest value between 08:00 and 13:30, when the load on the system decreased. It has been observed that THD is below the limit of 5% during the day.
- As a result of reviewing the weekly figures; it is understood that 3rd and 7th harmonics are appropriate for the weekly limit values defined in the standards. 3rd and 7th harmonics showed the highest values during the daytime hours of weekday (when SPP was active). It has been observed that the 5th harmonic generally exceed the limit values in the week and do not comply with the standards. THD's, in parallel with the 5th harmonic, limit values generally exceeded during the week.
- It is understood that monthly graphs show parallel results with weekly graphs.
- It is observed that the network voltage deviation is within the limits defined in the standard. SPP deviates the network voltage by 1% when entering and exiting the circuit; but remains within the standards (3%).
- In the end of the work; SPP is effective on 3rd and 7th harmonics; but it is understood that these harmonics do not affect enough to cause the limit values to be exceeded. The analyzed network does not comply with the standards in terms of 5th harmonic and THD so active filtration method should be applied this network. In terms of voltage deviation, it was the result of providing the network and SPP were coply with the standards.

REFERENCES

A. Hill, C., Such, M. C., Chen, D., Gonzalez, J., Grady, W. M., "Battery Energy Storage For Enabling Integration Of Distributed Solar Power Generation", IEEE Transactions On Smart Grid, vol.3, no.2, 2012.

Akkoyunlu, A. (2006), "Türkiye'de Enerji Kaynakları ve Çevreye Etkileri".

Amirullah, Penangsang, O., Soeprijanto, A. (2015), "Effect of Installation of Photovoltaic (PV) Generation to Power Quality in Industrial and Residential Customers Distribution Network", International Seminar on Intelligent Technology and Its Applications, p.p. 193-200.

Baggini, A. (2008), "Handbook of Power Quality", Wiley, p.p. 151-155.

Bhattacharyya, S., Myrzik, J. M., Kling, W. L. (2007), "Consequences Of Poor Power Quality- An Overview", UPEC, p.p. 651-656.

Block, P. A. B., Salamanca, H. L. L., Teixeira, M. D., Dahlke, D. D., Shiono, O. M., Donadon, A. R., Camargo, J. C. (2014), "Power Quality Analyses of a Large Scale Photovoltaic System", 5th International Renewable Energy Congress (IREC), p.p. 1-6.

"Elektrik Piyasasında Dağıtım Sisteminde Sunulan Elektrik Enerjisi Tedarik Sürekliliği, Ticari ve Teknik Kalitesi Hakkındaki Yönetmelik" (2016), Enerji Piyasası Düzenleme Kurumu (EPDK).

Enslin, J. H. R., Heskes, P. J. M. (2004), "Harmonic Interaction Between A Large Number Of Distributed Power Inverters And The Distribution Network", IEEE Transactions On Power Electronics, Vol. 19, No. 6.

Gonzalez, C., Geuns, J., Weckx, S., Wijnhoven, T., Vingerhoets, P., Rybel, T. D., Driesen, J. (2012), "LV Distribution Network Feeders in Belgium and Power Quality Issues due to Increasing PV Penetration Levels", 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe), p.p. 1-8.

Hürmeydan, S., Rosin, A., Vinnal, T. (2016), "Effects of PV Microgeneration on Rural LV Network Voltage Quality – Harmonics and Unbalance", Electric Power Quality and Supply Reliability (PQ), p.p. 97-100.

IEEE 519-2014 Standard. (2014), "IEEE Recommended Practice and Requirements for Harmonic Control in Electric Power Systems", IEEE Power and Energy Society.

Kandemir, Ç., Bayrak, M. (2015), "Fotovoltaik Sistemler Şebekeye Bağlı Olduğunda Oluşan Sorunlar", VI. Enerji Verimliliği, Kalitesi Sempozyumu Ve Sergisi Bildirileri (EVK).

Keçecioğlu, Ö. F., Tekin, M., Gani, A., Sarı, M., Şekkeli, M. (2015), "Endüstriyel Isı Santrallerinde Enerji Kalitesi Ölçümü Ve Değerlendirilmesi, Kahramanmaraş Sütçü İmam Üniversitesi Örneği", VI. Enerji Verimliliği, Kalitesi Sempozyumu Ve Sergisi Bildirileri.

Kocaman, Ç., Özdemir, M. (2009), "Güç Sistemlerinde Oluşan Güç Kalitesi Bozunumlarının Çarpıklık Ve Basıklık Katsayıları İle Belirlenmesi", 17th Signal Processing and Communications Applications Conference, IEEE, p.p. 341-344.

Pakonen, P., Hilden, A., Suntio, T., Verho, P. (2016), "Grid-Connected PV Power Plant Induced Power Quality Problems –Experimental Evidence", 18th European Conference on Power Electronics and Applications (EPE'16 ECCE Europe), p.p. 1-10.

Patsalides, M., Georghiou, G. E., Stavrou, A., Efthimiou, V. (2012), "Assessing the Power Quality Behaviour of High Photovoltaic (PV) Penetration Levels inside the Distribution Network", 3rd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), p.p. 709-716.

Şahin, M., Oğuz, Y., Tuğcu, H. Z. (2014) "Güç Sistemlerinde Enerji Kalitesini Etkileyen Harmoniklerin İncelenmesi", EÜFBED - Fen Bilimleri Enstitüsü Dergisi Cilt-Sayı: 7-2, 199-218.

Şimşek, B., Bizkevelci, E. (2015), "Türkiye Elektrik Dağıtım Şebekesinde Fotovoltaik Sistemlerin Güç Kalitesine Etkisi", IV. Elektrik Tesisat Ulusal Kongre Ve Sergisi Bildirileri.

Tanasescu, G., Busoi, S., Gorgan, B., Notingher, P. V. (2016), "Photovoltaic Power Plants Contribution To The Flicker Level In The Medium Voltage Energy Network. Case Study", Conference On Diagnostics In Electrical Engineering (Diagnostika), IEEE, p.p. 1-4.

Tekin, M., Sarı, M., Keçecioğlu, Ö.F., Gani, A., Açıkgöz, H., Yıldız, C., Şekkeli, M. (2016), "Power Quality Analysis Of A Campus Field: Case Study For Avşar Campus Of Kahramanmaraş Sütçü İmam University", 4th International Symposium on Innovative Technologies in Engineering and Science, (ISITES2016) Alanya/Antalya, Turkey, 3-5 November.

Tekin, M., Keçecioğlu, Ö.F., Gani, A., Açıkgöz, H., Yıldız, C., Şekkeli, M. (2016), "The Role Of Kahramanmaraş In Turkey From The Standpoint Of Based On Traditional And Renewable Electricity Generation", 10th International Clean Energy Symposium (UTES), 24-26 October, Istanbul, Turkey.

"Türkiye Elektrik Enerjisi 5 Yıllık Üretim Kapasite Projeksiyonu (2015 – 2019)", 2015.

Yanıktepe, B., Özalp, C., Savrun, M. M., Köroğlu, T., Cebeci, Ç. (2011), "Rüzgar-Güneş Hibrid Güç Sistemi Osmaniye Korkut Ata Üniversitesi Uygulama Örneği", 6th International Advanced Technologies Symposium (IATS'11), Elazığ, Turkey.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

CONSIDERING AIR DENSITY EFFECT ON MODELLING WIND FARM POWER CURVE USING SITE MEASUREMENTS

DOI: 10.17261/Pressacademia.2017.619
PAP-WCTIE-V.5-2017(57)-p.420-430

Ceyhun Yildiz[1], Mustafa Tekin[2], Ahmet Gani[3], O.Fatih Kececioglu[4], Hakan Acikgoz[5], Mustafa Sekkeli[6]
[1]Graduate School of Natural and Applied Sciences, K.Maraş Sütçü İmam University. cyildiz@ksu.edu.tr
2Department of Electrical and Electronics Eng., Faculty of Eng. and Architecture, K.Maraş Sütçü İmam University. mustafatekin@ksu.edu.tr
3Department of Electrical and Electronics Eng., Faculty of Eng. and Architecture, K.Maraş Sütçü İmam University. agani@ksu.edu.tr
4Department of Electrical and Electronics Eng., Faculty of Eng. and Architecture, K.Maraş Sütçü İmam University. fkececioglu@ksu.edu.tr
5Graduate School of Natural and Applied Sciences, K.Maraş Sütçü İmam University. hakanacikgoz@ksu.edu.tr
6Department of Electrical and Electronics Eng., Faculty of Eng. and Architecture, K.Maraş Sütçü İmam University. msekkeli@ksu.edu.tr

ABSTRACT

Manufacturers develop power curves for their wind turbines. Customers use these wind turbine power curves for wind farm planning and estimating nearly total production of planned plant. When wind farm is installed and connected to the grid, these power curves are not useful. In literacy, researchers proposed wind turbine power curve measurement methods to obtain an accurate power curve for turbine on site. But it is not easy to develop power curves for clusters of wind turbines. Developing a single power curve for a wind farm slightly simplifies this problem. Accurate wind farm power curve is a very useful tool for converting wind speed forecasts to power. Also plant owner can use this tool to detect anomalous operations. In this study we developed and tested wind farm power curves by using real site measurements. Two different methods are used to develop power curves. They are polynomial curve fitting and mean bins method. Wind speed and power output relation is investigated. A method is proposed to add effect of air density on power curve. Developed power curve has two inputs. They are hourly mean wind speed and air density values. This approach uses variable air density in calculation of wind farm power output. Results of this study showed that performance of mean bins method is better than polynomial curve fitting. Also proposed air density effect adding method improves performances of obtained power curves.

Keywords: Wind farm, power curve, air density effect, mean power bins, wind turbine power curve.

1. INTRODUCTION

Over the last decades wind energy has been the fastest growing renewable energy[1]. Promoting policies of countries and decrease in installation costs increased the total installed capacity of wind power all around the world. It is well known that wind energy has variable, uncertain nature and it is not easy to balance power system with high wind power penetration [2]. Transmission system operators must ensure and maintain the balance between supply and demand in their electrical grids. Within this scope numerous studies on wind power forecast have been carried out. Forecasts help TSOs to create generation plan.

Wind energy forecast systems use wind farm power curves to transform forecasted wind speeds to electrical power. So accuracy of power curve has an impact on performance of the forecast. Developing a power curve for a wind farm is not an easy process. Because wind farms will have clusters of wind turbines spread over a large area and site conditions differs cluster to cluster. Also air density, topographical obstacles, wake effects, wind direction etc. effects wind speed-power output relation.

It is necessary to have accurate site measurements to develop a wind farm power curve and to evaluate the performance of power curve. Hard weather conditions like icing, very high wind speeds etc. make difficult get accurate measurement data.

These measurements must also include wind farm power output, wind speed, wind direction, pressure, temperature and humidity. Other meteorological data air density, is function of pressure, temperature and humidity [3].

A lot of study on wind farm power curve development have focused wind and power relation [4-10]. In literacy small number of studies investigated effect of air density on power curves. These studies have different conclusions about considering the effect of air density. Farkas [11] suggests considering air density in wind power calculations.On the other hand, according to Wan and at all. [12] using air density as an additional input in calculation didn't improve the performance. Also [14]standard on wind turbine performance test proposed normalizing wind power data with constant measured mean air density.

In this study we developed a wind farm power curve. Real site measurements are used during the development and test stages. Wind speed, power output and air density data are used in calculations. Polynomial curve fitting and mean bins methods are used to develop power curves. A new method is used to assess variable air density effects.

The rest of this paper is structured as follows: Section 2 describes measurement stations; Section 3 analyzes measured data; Section 4 describes application of methods; Section 5 analyzes power curve performance; Section 6 includes conclusions of the study.

Nomenclature	
c: wind power conversion constant	M_a: molar mass of dry air (g/mol)
h: humidity (0<h<1)	M_v: molar mass of dry water (g/mol)
p: pressure (Pa)	**R**: molar gas constant ($JK^{-1}mol^{-1}$)
P: power (MW)	**T**: thermodynamic temperature
Pc: power curve	X_v: mole fraction of water vapour
P^v: air density rejected power (MW)	**Z**: compressibility factor
t: temperature (^{0}C)	ρ :air density (kg/m^2)
v: wind velocity (m/s)	ρ_m :measured air density (kg/m^2)

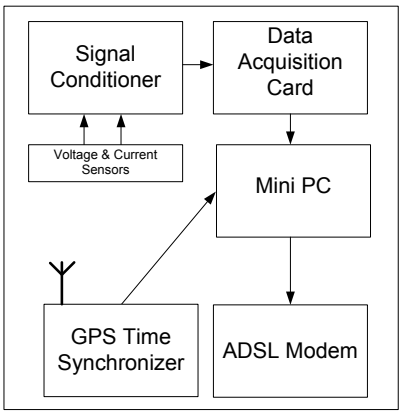
2. MEASUREMENT STATIONS

In this study two kinds of measurement stations (electrical power and a meteorological data measurement station) are used as wind farm power curve is a result of relation between meteorological data and electrical power output of plant. Rest of this section describes measurement stations.

2.1. Electrical Power Measurement Station

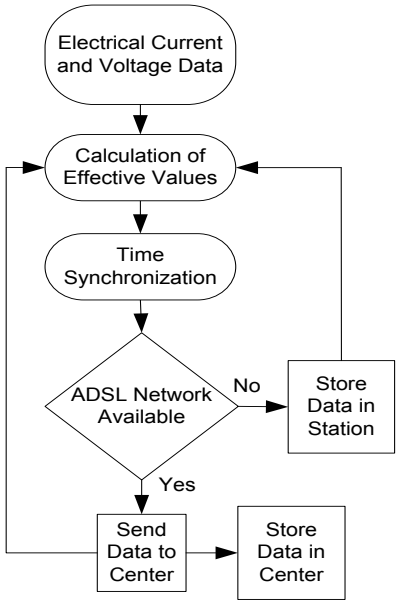
This station is installed in the power house of the plant. Power output and it's quality are measured. Hardware of the station is composed of signal conditioner and data acquisition cards, mini pc, modem and GPS time synchronizer equipment. Power plant's sensors are used to get voltage and current values. Signal conditioner card converts raw sensor signals to suitable form for data acquisition card. After signal conditioning process data acquisition card converts these signals to numerical values. Mini PC performs calculation, communication and time synchronization algorithms. GPS time synchronizer gets accurate time data to create time label of calculated values. Simple hardware configuration of the system is given in Figure 1.

Fig. 1: Hardware Configuration of Power Measurement Station



After transforming raw analog (voltage, current) measurements into numerical values, mini PC performs calculations in accordance with IEC standard numbered 61000-4-30 [13]. Simple flow chart of electrical power measurement algorithm is as follows.

Fig. 2: Flow Chart of Electrical Measurement Algorithm

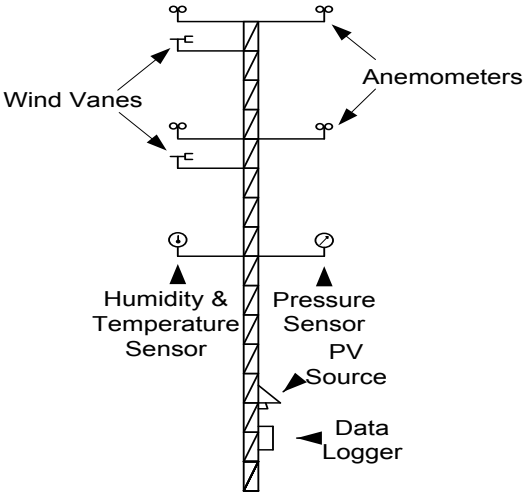


At first step of the algorithm effective values are calculated. Second step is time synchronization. After them algorithm checks network. If network is available system sends data to data center. If network is not available system stores data in measurement station.

2.2. Meteorological Measurement Station

Wind turbines are spread over a large area. So it is not easy to measure all turbine site's meteorology. For this reason a specific site that can represent whole plant area meteorology is selected. The hardware of the system is composed of wind speed, wind direction, pressure, humidity, temperature sensors; data logger, photo voltaic energy source system and tower. Installation is carried out in accordance with IEC standard numbered 61400-12-1 [14]. Simple hardware scheme is given in Figure 3.

Fig. 3: Schematic Expression of Meteorological Measurement Station



Four anemometers and two wind vanes are used to measure wind speed and direction values. Object of duplicate using of anemometer and vane is to detect malfunction operation and to get values at different altitudes. The system will check sensor measurement by comparing another sensor's value. GPS time synchronization is also used in this station and GPRS modem is used for communication. Software of this system has simple algorithm. Algorithm gets sensor data, synchronizes it with GPS time, if network is available sends data to data center if network is not available stores data in station. Data logger performs these algorithms.

3. DEVELOPING WIND FARM POWER CURVE

Wind farm power output and meteorological data relation gives power curve of the plant. First step of the investigation is to obtain accurate data set. Data set comprises wind speed, air density and power output of wind farm. Air density is function of three types of measurements (air pressure, temperature and relative humidity). Detailed calculation of air density is given in this section. Second step is applying methods to obtain relation between meteorological data and power output. In this study polynomial curve fitting, mean bin methods are used. Third step, performance evaluation is given in the next Section 4.

3.1. Meteorological and Power Data Set

Detailed explanation of measurement stations are given in previous sections. In this section short overview of measured data and air density calculation are given. Meteorological measurements involve wind speed, temperature, air pressure and relative humidity. Fig. 4 is histogram of wind speeds and fitted Weibull curve with k and c parameters (k=1.95, c=8.5).

Fig. 4: Wind Speed Histogram and Fitted Weibull Curve

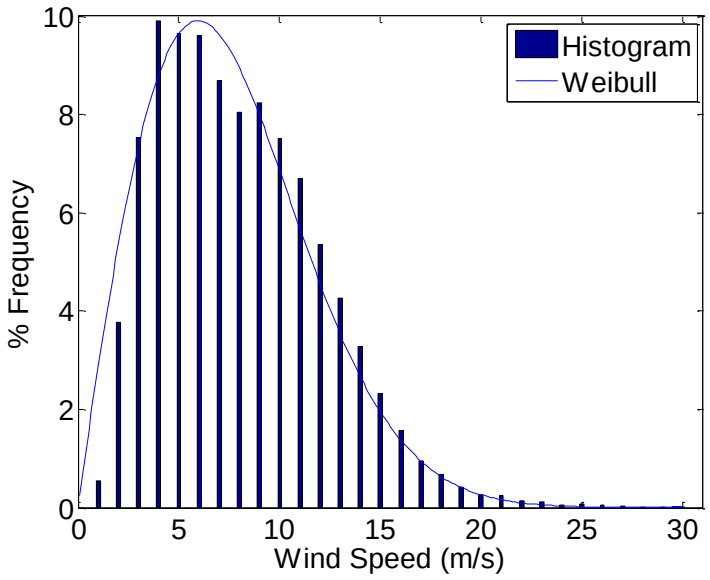


Fig. 5-7 represent available % humidity, air pressure in (mb) and temperature in (C^0) data respectively. Because of the hard weather conditions and PV source shortage some data are not available. Measurement period involves 30553 hours as 21123 hours are available. There are five kinds of measurement sensors so it is not easy to operate them simultaneously. It is needed to have three kinds of measurements (pressure, humidity, temperature) to calculate air density also wind vane and anemometer are needed to measure wind data. So number of available data reduces.

Fig. 5: Humidity Data

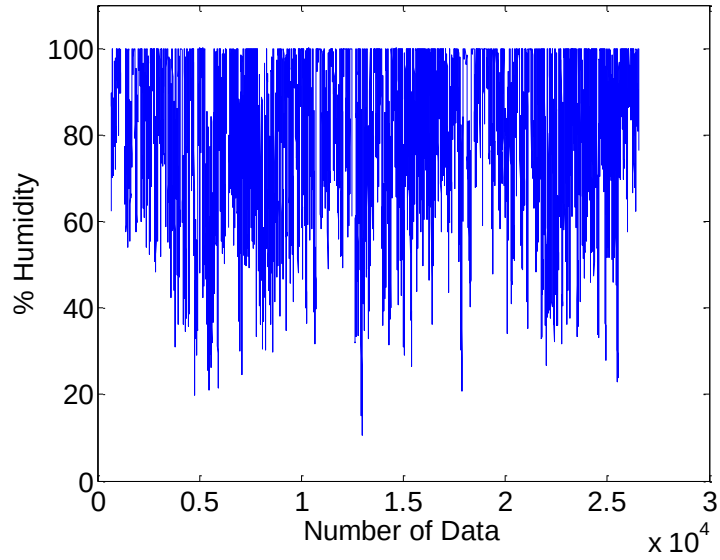


Fig. 6: Pressure Data

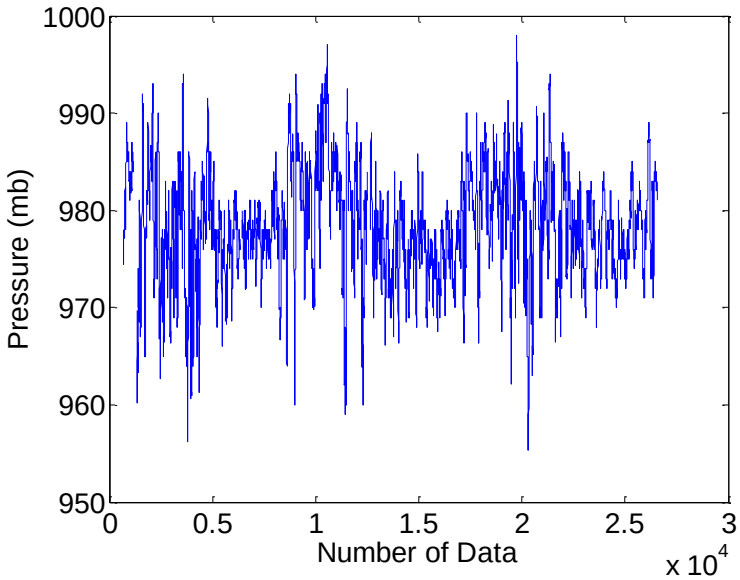
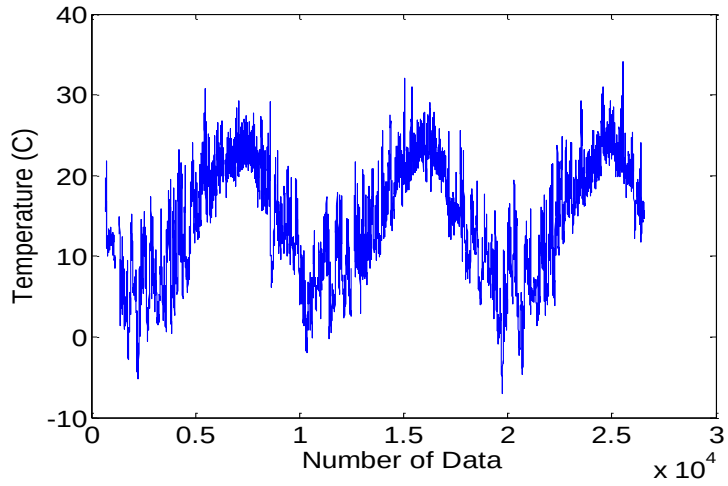


Fig. 7: Temperature Data



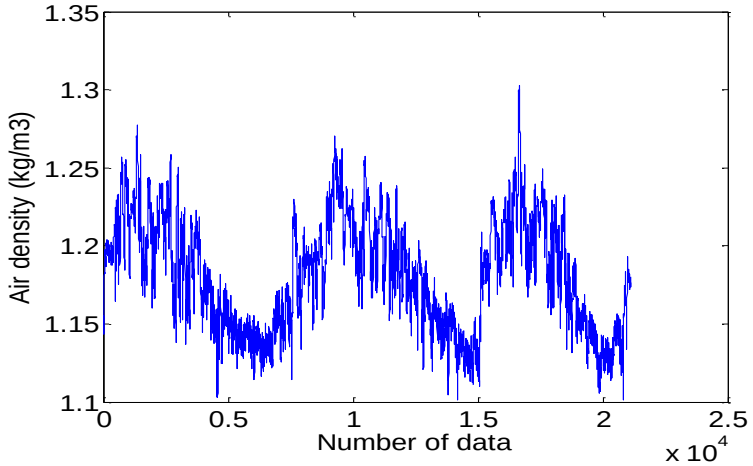
Formula of moist air density is used to calculate density of air. Formula has three variables. *t* is temperature (^{0}C), *p* is air pressure [Pa], *h* is relative humidity (0≤h≤1). Other parameters are constants, given in [3]. Calculated air density is as in Fig. 8.

$$\rho(t,p,h)=\frac{pM_a}{Z(t,p,h)RT(t)}(1-x_v(\mathrm{t,p,h})[1-\frac{M_v}{M_a}]) \quad (1)$$

$$x_v(t,p,h)=h[\alpha+\beta p+\gamma t^2]\frac{e^{[\mathrm{AT(t)^2+BT(t)+C+}\frac{D}{T(\mathrm{t})}]}Pa}{p} \quad (2)$$

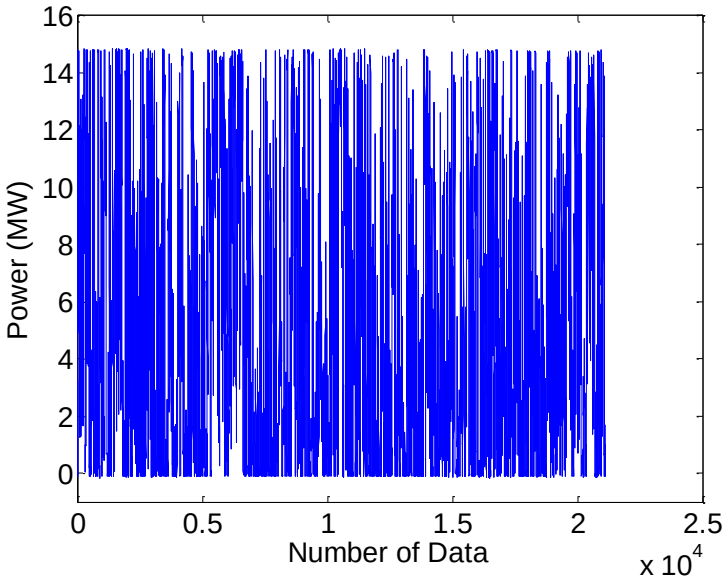
$$Z(t,p,h)=1-\frac{p}{T(t)}\left(a_0+a_1t+a_2t^2+[\mathrm{b_0+b_1t}]\mathrm{x}_v+[\mathrm{c_0+}c_1\,\mathrm{t}]\mathrm{x}_v^2\right)+\frac{p^2}{T(\mathrm{t})^2}[\mathrm{d+ex}_v^2] \quad (3)$$

Fig. 8: Calculated Air Density



Last data set involves output powers of the plant. Wind farm has five 3 MW wind turbines. Installed power is 15 MW. Electrical power output of the system is measured simultaneously (time synchronization between meteorological data is carried out). Total power output of the farm is as in Fig. 9.

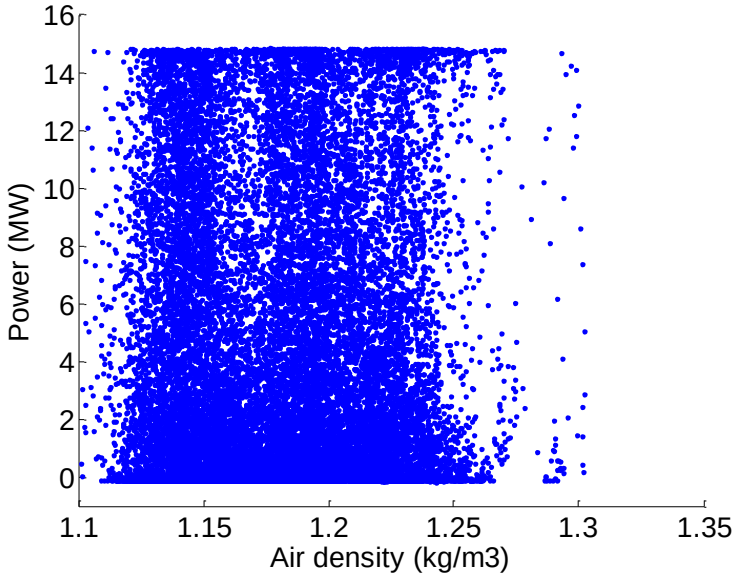
Fig. 9: Wind Power Plant Output



4. APPLIED METHODS

Before applying methods, scatter plots of wind speed-power and air density-power are created. As expected, these plots showed that relation between wind speed and power output is stronger than air density-power relation. It is known that wind speed has cubic, air density has proportional relation between power[15]. Stronger relation of wind speed is result of this situation. Scatter plots of air density-power and wind velocity-power are as in Fig. 9 and Fig. 10.

Fig. 9. Scatter Plot of Air Density-Power



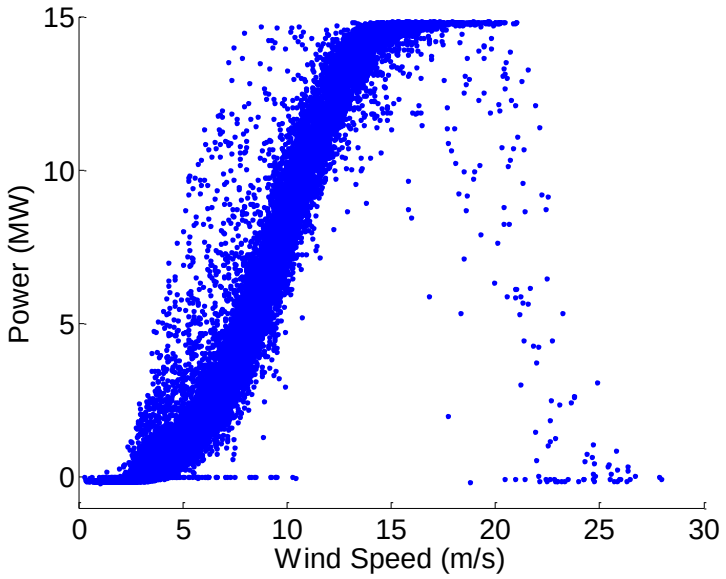


Fig. 10. Scatter Plot of Wind Velocity-Power

A polynomial curve fitting method is applied to create power curve. One year period of available data is used to perform method applications. 7. order polynomial fit is selected as a fitting method. Matlab Curve Fitting toolbox is used to fit curve. Other method is mean bins, this method is specified in [14]. In this method averages of powers and wind speeds are calculated for each 0.5 (m/s) wind speed bins. These values formed power curve. Comparison of these power curves is given in Fig. 11.

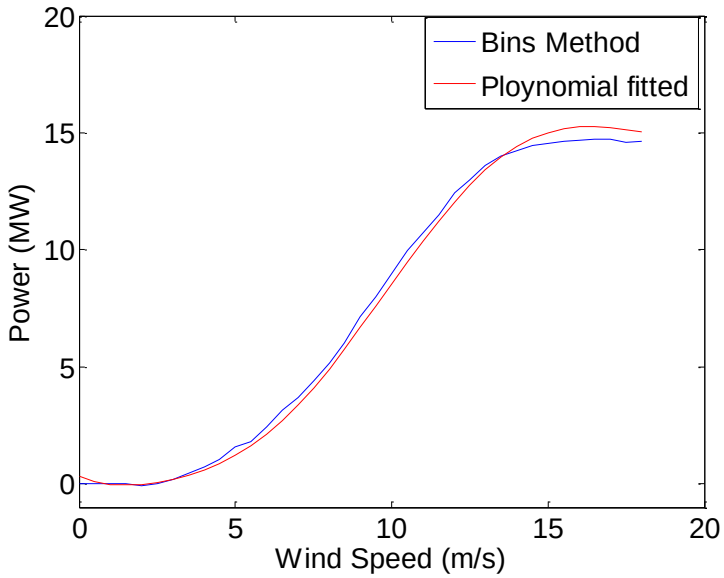


Fig. 11: Polynomial Fitted and Mean Bins Method Power Curves

Mean bins method power curve has smaller values near installed power output. It is the result of low power outputs of wind farm at rated wind speeds. These two methods have single input and single output. Effect of air density is not included.

A new method is proposed to investigate effect of air density on power curve. It is well known that wind power has proportional relation between air density as given in (4). To reject the effect of measured air density on measured power, hourly measured power is divided to measured air density (5). Air density rejected values are used to calculate air density rejected power curve. After this process, calculated power curve is tested with measured data by multiplying it's output with hourly measured air density values as in equation (6). By applying this method air density effect is rejected during curve development phase and it is injected during test phase.

$$P = \frac{1}{2} c \rho v^3 \tag{4}$$

$$P^{v}(v) = \frac{1}{2} c v^3 = \frac{P_m(\rho, v)}{\rho_m} \tag{5}$$

$$P = Pc(\mathrm{v})\rho_m \tag{6}$$

At the second phase of the method applications. Effect of air density on power curves is argued. Comparison of air density rejected power curves and others are given in Fig. 12 and Fig. 13.

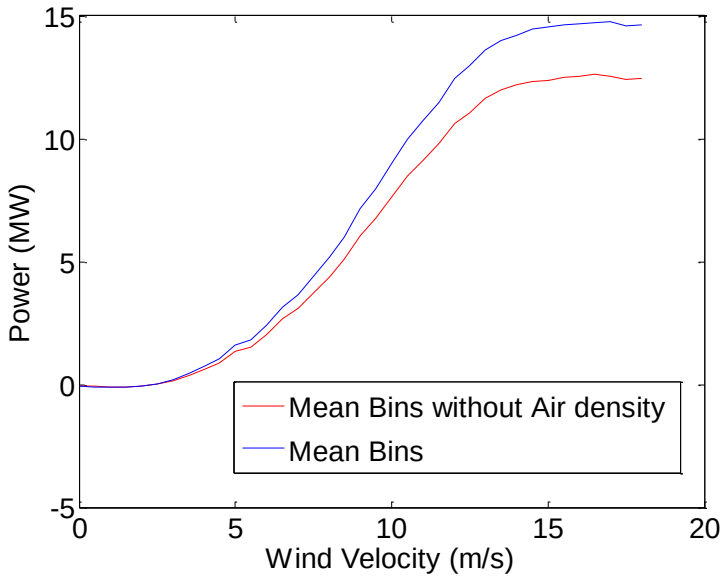


Fig. 12. Mean Bins and Mean Bins Without Air Density Method Power Curves

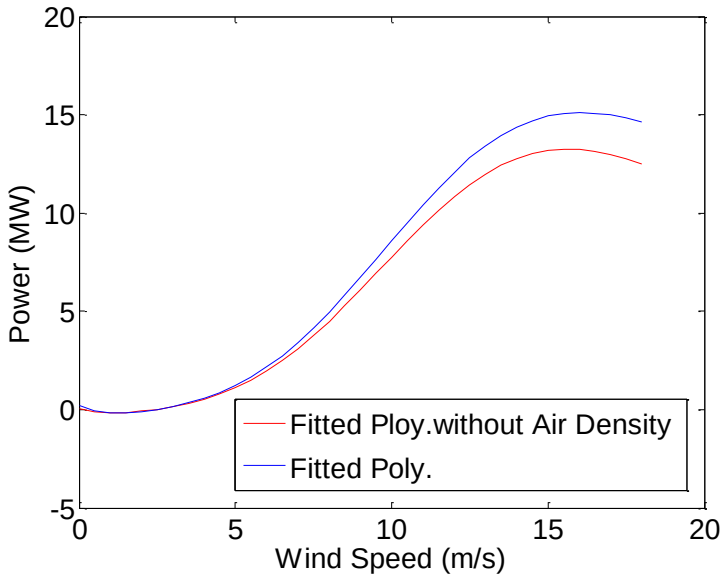


Fig. 13: Polynomial Fitting and Polynomial Fitting Without Air Density Method Power Curves

5. PERFORMANCE TEST RESULTS

In this section performances of methods are analyzed. One year period of available data is used to carry out performance tests. One year period is necessary to see seasonal effects on performances. Number of all available data is 21113 hours, it is nearly equal to 2.4 years. One year period was used to perform methods applications. Other year is used for performance tests. MAE of power curve outputs are calculated to evaluate performances. Table 1 summarizes MAE values. Test resuts showed that proposed air density rejection method improved performances.

Table 1: MAE Values of Methods

Method	% MAE
Polynomial Fit.	5.46
Polynomial Fit. Without Air density	4.80
Mean Bins	4.36
Mean Bins Without Air Density	4.32

6. CONCLUSIONS

This paper presents development of power curve using two different methods and air density effect on these methods. Power curves are developed and tested using real site measurements.

Power curves are very useful to detect anomalous operation of plants and to convert meteorological forecasts into power. So accuracy of power curve increases accuracies of these important calculations. Great number of researches on this topic focused on wind velocity-power relation. There is restricted number of studies in literacy that argues effect of air density on power curve. So results of this study are significant. On the other hand a new method is proposed to reject air density effect on power curve. This method is applied and performance improvements are assessed.

REFERENCES

[1] Bitar, Eilyan Y., et al. "Bringing wind energy to market." Power Systems, IEEE Transactions on 27.3 (2012): 1225-1235.

[2] Evans, Annette, Vladimir Strezov, and Tim J. Evans. "Assessment of utility energy storage options for increased renewable energy penetration." Renewable and Sustainable Energy Reviews 16.6 (2012): 4141-4147.

[3] A. Picard, R. Davis, M. Gl¨aser, and K. Fujii, "Revised formula for thedensity of moist air (cipm-2007)," Metrologia, vol. 45, pp 149–155, 2008

[4] Marvuglia, Antonino, and Antonio Messineo. "Monitoring of wind farms' power curves using machine learning techniques." Applied Energy 98 (2012): 574-583.

[5] Chang, Tian-Pau, et al. "Comparative analysis on power curve models of wind turbine generator in estimating capacity factor." Energy 73 (2014): 88-95.

[6] Lydia, M., Sahoo Subhendu Kumar, and G. E. P. Kumar. "Advanced algorithms for wind turbine power curve modeling." Sustainable Energy, IEEE Transactions on 4.3 (2013): 827-835.

[7] Olaofe, Zaccheus O., and Komla A. Folly. "Wind energy analysis based on turbine and developed site power curves: A case-study of Darling City." Renewable Energy 53 (2013): 306-318.

[8] Carrillo, C., et al. "Review of power curve modelling for wind turbines." Renewable and Sustainable Energy Reviews 21 (2013): 572-581.

[9] Lydia, M., et al. "A comprehensive review on wind turbine power curve modeling techniques." Renewable and Sustainable Energy Reviews 30 (2014): 452-460.

[10] Milan, Patrick, Matthias Wächter, and Joachim Peinke. "Stochastic modeling and performance monitoring of wind farm power production." Journal of Renewable and Sustainable Energy 6.3 (2014): 033119.

[11] Farkas, Zeno. "Considering air density in wind power production." arXiv preprint arXiv:1103.2198 (2011).

[12] WAN, Y.; ELA, Erik; ORWIG, Kirsten. Development of an equivalent wind plant power curve. In: Proc. WindPower. 2010. p. 1-20.

[13] Compatibility, Electromagnetic. Part 4: 30: Testing and measurement techniques–Power quality measurement methods. IEC 61000-4-30 Std, 2003.

[14] Turbines, Wind. Part 12-1: Power performance measurements of electricity producing wind turbines; IEC TC/SC 88. IEC 61400-12-1, 2005.

[15] Ackermann, Thomas, ed. Wind power in power systems. John Wiley & Sons, 2005.

PressAcademia Procedia - (PAP), ISSN: 2459-0762

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

DETERMINATION OF ENHANCEMENT EFFECTS AT THE COMPTON TO COHERENT SCATTERING INTENSITIES X-RAYS USING FE-CU AND FE-ZN BINARY COMPOUNDS WITH WDXRF

DOI: 10.17261/Pressacademia.2017.620
PAP-WCTIE-V.5-2017(58)-p.431-433

Elif Boydas[1] Demet Yilmaz[2] Mustafa Gokalp Boydas[3]
[1]Atatürk University .ebaydas@atauni.edu.tr
[2]Atatürk University. ddemir@atauni.edu.tr
[3]Atatürk University .mboydas@atauni.edu.tr

ABSTRACT

Matrix enhancement effects have been examined with scattered X-rays by using WDXRF system for Fe-Cu and Fe-Zn binary compounds. Fe as analyte and elements which some of them the K absorption edge just below the characteristic K X-ray energies of the analit (Fe) as matrix have been chosen. Pure and binary with Fe ratio ranging from 10% to 90% samples were prepared with these elements. These samples have been analyzed in WDXRF system. At the end of measurement, the coherent and Compton scattering X-rays data of Rh Kα X-rays from these samples have been obtained. Matrix enhancement effects have been examined by calculating coherent / Compton scattered X-rays intensity ratio.

Keywords: Analyte, enhancement, compton scattering, coherent scattering , x-rays

1. INTRODUCTION

Compton scattering is one of the methods giving information about the electronic structure, electronic momentum distribution and the wave functions of atoms, molecules and solids. Compton scattering is the scattering of a photon of high energy from an electron considered

being free and stationary or from an atomic electron whose binding energy is small compared with the incident photon energy. X-ray scattering techniques are nondestructive and often have a great advantage over the traditional chemical techniques used for determination of the effective atomic number of samples of environmental, biological, agricultural, industrial and medical interest. The 3d transition metals and their alloys have played an important role in the development of modern technology, and knowledge of their physical parameters such as mass attenuation coefficients, total atomic and electronic cross-sections, and effective atomic and electron numbers is very important for understanding their physical properties .

X-ray scattering techniques are nondestructive and often have a great advantage over the traditional chemical techniques used for determination.of the effective atomic number of samples of environmental, biological, agricultural, industrial and medical interest. Han et al (2009) shows that the 3d transition metals and their alloys have played an important role in the development of modern technology, and knowledge of their physical parameters such as mass attenuation coefficients, total atomic and electronic cross-sections, and effective atomic and electron numbers is very important for understanding their physical properties. Speller et al. (1991) have introduced applications of the coherent and Compton scattering in the fields such as medicine and biology. Webster et al (1985) have measured the coherent to Compton scattering intensity ratio in order to assess trabecular bone mineral content. Cooper et al (1982) examined sample composition change in alloys by using scattered intensity ratio. Gigante et al. (1985)have analyzed the metal alloys by Rayleigh to Compton ratios and X-ray fluorescence peaks in the 50–122 keV energy range. Shakeshaft et al (1997) have introduced applications of this technique in the fields of medicine and biology .

In this study, We measured coherent and Compton peaks of binary systems. We examined enhancement effects of matrix on analyt by using coherent/Compton scattering area ratios.

2. EXPERIMENTAL

Fe was used as analyt. As a matrix were used elements such as Cu, Zn. While concentration of analyt gradually increased in steps of 10% from 10% to %90, that of matrix gradually decreased in steps of 10% from 90% to %10.Analyt and one of matrice were mixed these proportions. For example, this proportion can be such as Fe of 30% and Ca of 70% in binary systems. Compton and coherent peaks of this samples were measured by using wavelength-dispersive X-ray fluorescence (WDXRF) spectrometry.

3. FINDINGS AND DISCUSSIONS

R (R= Icoh/IC) versus concentration of Fe in binary systems of Fe-Cu, Fe-Zn are given in Fig. 1-2. Thus this enhancement effects is observed in the Zn element with respect to the Fig.2. These dispersed X-rays of matrix elements might also excite the examining element (analyte) as well, be called the Matrix enhancement effects effect. Concluded that ratio of Icoh/Ic ascended in the position of decrement for all systems concentration of Fe, while increment of other Matrix concentration. Increment in this ratio of these elements evident that could be an intensify impulse on Fe K X-ray. In as much as K shell absorption edge of Co, Ni, Cu and Zn reside inside of short wavelength of Fe K edge. Accordingly these elements' primary X-ray absorption is very narrow so that characteristic X-rays of these elements will be intensifying the Fe K X-rays and can be considered as intensifying elements of previous stated elements.

Figure 1: R (R= Icoh/IC) versus Concentration of Fe in Binary Systems of Fe-Cu

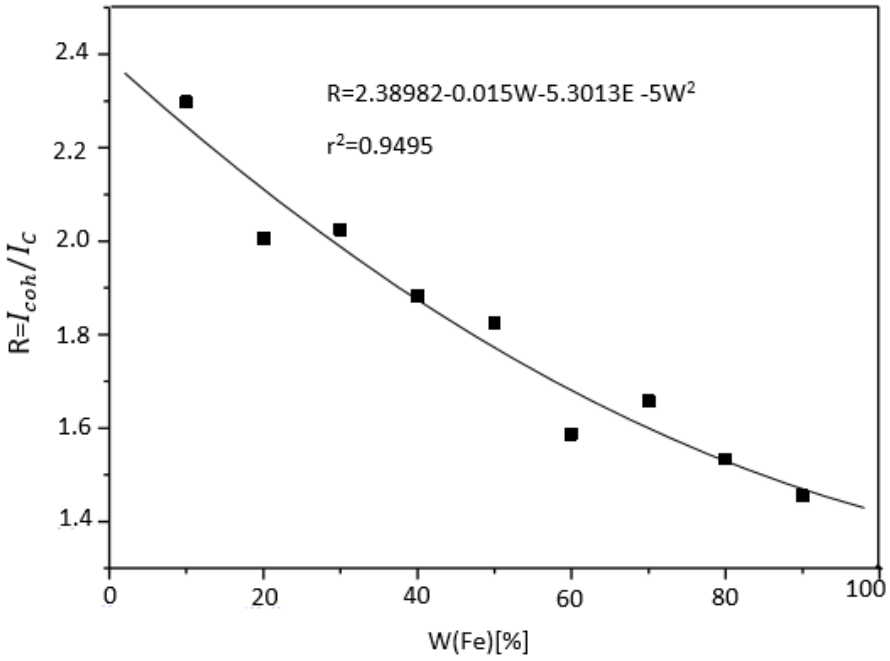
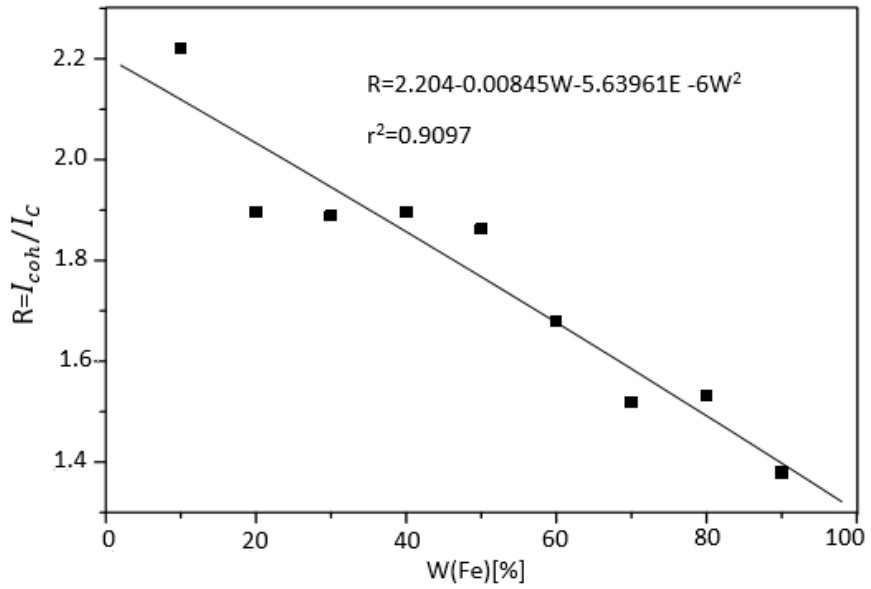


Figure 2: R (R= Icoh/IC) versus Concentration of Fe in Binary Systems of Fe-Zn



4. CONCLUSION

The aim of the present study is to determine enhancement effects using coherent / Compton scattered X-rays intensity ratio. Matrix enhancement effects have been examined with scattered X-rays by using WDXRF system for Fe-Cu and Fe-Zn binary compounds. the coherent and Compton scattering X-rays data of Rh Kα X-rays from these samples have been obtained. Matrix enhancement effects have been examined by calculating coherent / Compton scattered X-rays intensity ratio.

REFERENCES

Han, I., Demir, L. 2009. "Mass attenuation coefficients, effective atomic and electron numbers of Ti and Ni alloys",Radiation. Measurement, vol.44 no. 1 ,pp. 289-294.

Speller, R. D., Horrocks, J. A. 1991," Photon scattering--a "new' source of information in medicine and biology?" Physics in Medicine and Biology .Vol.36, no.,pp.1.

Webster, D. J., Lillicrap, S. C.1985,"Coherent-Compton scattering for the assessment of bone mineral content using heavily filtered X-ray beams" ,Physics in Medicine and Biology, volume 30, no. 6, pp. 531-539.

Cooper,M.J., Rollason, A.J., Tuxworth, R.W. 1982," Gamma-ray scattering studies of composition variations in alloys" J. Phys. E: Sci. Instrum,vol.15, pp. 568-572

Gigante,G.E., . Pedraza, L.J, Sciuti, S. 1985," Analysis of metal alloys by Rayleigh to Compton ratios and x-ray fluorescence peaks in the 50 to 122 keV energy range" Nucl. Inst. Meth, vol.12, pp. 229–234

Shakeshaft,J., Morgan H.M., Lillicrap, S.C.1997, "Gamma-ray scattering for fat fraction measurement", Physics in Medicine and Biology, vol. 42, no.7,pp. 1403-1413.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

MODELING A CONTEXT-AWARE FRIEND-OF-A-FRIEND(FOAF) APPLICATIONS FOR MOBILE PLATFORMS

DOI: 10.17261/Pressacademia.2017.621
PAP-WCTIE-V.5-2017(59)-p.-434-442

Bujar Raufi[1], Florije Ismaili[1], Jaumin Ajdari[1], Xhemal Zenuni[1]
[1]South East European University, b.raufi, f.ismaili, j.ajdari, xh.zenuni@seeu.edu.mk

ABSTRACT

We live in an era where the evolution of social networks and developments in the field of information technology witnesses and exponential growth. On a daily basis, we have an ever growing, an ever expanding social network users and as a result we affirm an increase of data that is distributed across different platforms. In this paper, we aim to treat the semantic web from the prism of creating a suitable infrastructure for data integration on the web. One of the most promising applications of the semantic web is the presentation of profiles using an RDF (Resource Description Framework) schema called Friend-of-Friend (FOAF) which represents a machine-processable ontology for describing persons, their activities and relations to other people and objects. In order to demonstrate the concept on how to integrate, relate and share information using FOAF, we model and develop a mobile application called "Find professional" which consists in creating, finding and interacting in a context aware fashion with other users via FOAF. The "Find Professional" mobile app can create a FOAF profile either by asking user's direct input or it can utilize other social network profiles such as Facebook for importing user public profiles into FOAF. User profiles represent RDF files that can be read and queried by semantic query languages like SPARQL, which returns data to their user profiles as well as finds professionals in the FOAF app based on user's geographic location. The application is also evaluated with a field usability testing performed against ten randomly picked user from various profiles and backgrounds. Usability testing process did not yield severe problems during application evaluation and comments are more related to user interface rather than difficulty in user task completion.

Keywords: Semantic web, foaf dictionary, semantic social network, context-awareness.

1. INTRODUCTION

FOAF (Friend-Of-A-Friend) consists on the connection between people and objects in one side and semantic information on the other using a web platform. FOAF ontology is machine redable and processable which serves to describe activities and their relationship with other persons and objects as well as can be used for entity self-description. The role of FOAf is threefold, first it integrates *social networks* for human collaboration, friendship and association; second it integrates the so called *representational networks* which gives a simplified view of a universe based on factual terms, and third it integrates *information networks* that utilize the various types of linkings in the Semantic Web to share independently published descriptions of resources on this inter-connected world (Brickley, 2016). Management from the perspective of the social network integration that are distributed in different profiles is seen as a very profitable union of all these relationships in a single set of data. In an integrated social networks, persons who have several user profiles will be presented as a single entity. Information about users which are distributed accross various user profiles will also be interconnected and joined. This paper presents the importance of FOAF vocabulary components as well as conduct an analysis on how these components enable data integration in a single mobile platform.

As a result, all theoretical analysis based on earlier scientific research and W3C specifications about semantic web and FOAF technologies as pillars of supporting this platform will be conceptualized and developed through an android application. The application consists in creation of social network linkings as well as their localization and matching of similar profiles. This application uses simple forms where users can create their FOAF profile as well as integrate an approach that aims to

join Facebook application data that can be added to a FOAF profile as RDF. The application generates the RDF file that can be read and queried by specialized languages such as SPARQL (SPARQL, 2016), which returns data to their user profiles. The paper focuses also on examination of all necessary classes and FOAF vocabulary items (such as components and the persons he or she knows), expressed in RDF format. All this is achieved through a mobile application that converges all the aspects of FOAF usage in one context-aware environment.

The rest of this paper is organized as follows: section 1 introduces related work done with relation to FOAF based mobile applications, section 3 introduces the FOAF mobile application from the sense of its architecture and functionality, section 4 evaluates the application with findings from usability testing and section 5 concludes the paper with some outlines for future improvements.

2. RELATED WORK

The proliferation of the web made it possible for users to visit different sites, read the information published there and interact with the web in various ways. In recent years, new types of applications emerged which allow visitors not only to read the information passively, but also to act and influence upon user's activities, such as the online reservations, ticketing services and recommender systems among others.

Web applications are being exponentially populated and as a result information is stored in huge databases that utilize various technologies for that purpose such as Datacubes (Edoh-Alove, 2016), (Höffner, 2016) or noSQL databases (Swaminathan, 2016), (Gantz, 2016). But the information stored in these databases are either semi-relational, un-relational (in case of noSQL) or hidden from user activities and existing search engines. Even though that search engines are used to find and collect information, most of the time they fail in finding a direct and accurate information to whom users are interested (Dan & Davidson, 2016). The information retrieval process that lies out of the scope of semantic web is facing serious challenges that influence the search process in many ways. Problems that IR is facing are ranging from spams, poor content quality, lack of quality evaluation, absence of web conventions up to presence of duplicate sources and vagely structured data and they all represent obtrusive factors for a good search (Henzinger, 2002). Based on the above mentioned, the approach for the future would be to structure the data as much as possible, together with their association with the additional information that may be descriptive or helpful to end users of the web (Heath, 2011). This is the core issue which semantic web faces. Semantic search helps in finding information by dealing with complex queries, presenting precise and self-explanatory results and quick response (Lei, 2006). It also provides a common carrier structure, which allows the distribution and reuse of data between applications, enterprises and community boundaries beyond (Berners-Lee, 2016), (Torres, 2017). The goal is to allow communication between the automated processes, which make the data readable by machines in a detailed way and helps in finding information on millions of subjects in their web pages. In order to achieve the goal, it is encouraged to develop web applications, especially in the mobile application area that will allow users to enter semantically structured data that can be later reused by semantic agents.

In many approaches, challenges in creating an appropriate infrastructure for the process of integration of information on the web is risen but not holistically addressed (Allemang, 2011). In Meng et. al, an analysis of the problem of knowledge base semantic integration using crowd intelligence is elaborated. As a result, a novel hybrid framework for knowledgebase semantic integration considering the semantic heterogeneity of such KBs class structures is introduced. The proposal also addresses the problem of vastnes of such knowledgebases as well as problems that cowdsourcing introduces. The approach, faces the challenges of crowdsourcing as well as does not provide a unified format of data representation of such integration. Another approach proposes the use of Bluetooth technologies to connect multiple mobile devices in an ad hoc network in a user-centered location-based service that enables stable and efficient data communication (Jiang, 2007). However the lack of the approach is in the underlying infastructure of the Bluetooth which does not support large scale infrasturture that would allow social network data integration.

The objective of this paper is to build a semantic social network based on FOAF ontology. There are various approaches that deal with integration of semantic web into different web and mobile platforms. Many of the attempts are made in the field of Linked Open Data. Some of the ideas involve creating a user friendly interfaces with relation to navigation of the resources in Linked Open Data Repositories (Bizer et. al, 2009), (Becker, 2008). In Becker et. al, a mobile application is proposed which presents an overview of the Geospatial Semantic Web and introduces DBpedia Mobile, a location-aware Semantic Web client which is an app that can be used in many mobile platforms (Becker, 2009). Context awareness is an important aspect when it comes to mobile application and search of semantic repositories. It enables computer systems to anticipate users' needs and to act in advance in one hand and improve the user interactions with computing system on the other (Chen, 2000).

To the best of our knowledge, there are many applications and architectures that tend to bring social networking paradigm to mobile platform in one holistic design (Rana, 2009). We can mention few of them as: (Sheth, 2009), (Schaffert, 2009),

(Jung, 2010), (Torres, 2017) as well as some industry approaches such as (Serena, 2017) . However they lack proper data integration from social networks as well as their applicability in the domain of context-awareness in such cases.

In this paper we integrate all of the above mentioned by developing a mobile social network application called "Find professional" which consists in utlizing other social network profiles for importing and finding professionals in FOAF based on their geographic location.

3. FRIEND-OF-A-FRIEND (FOAF) APPLICATION FOR MOBILE PLATFORMS

In order to demonstrate the functionality of FOAF and its application to social networking, we have introduced a small app that builds upon the above mentioned ideas and it tends the integration of profiles from other social network accounts in order to create a FOAF user profile for finding various professionals.

The FOAF ontology is the most widely used domain ontology on the semantic web. Many people are using it in an open and extensible fashion by defining new classes and properties to use with FOAF (Ding et. al, 2004). The FOAF vocabulary in itself includes classes and properties found useful to describe people online. It is interoperable with other semantic web standards as well as support auto discovery and cross reference with other FOAF documents. It can be serialized in various support formats for semantic web such as: RDF / XML, RDFa, Microdata, N3, N-Triples or JSON. The following example depicts a FOAF user represented as JSON serialization.


```
        {
  "@id": "#me",
  "@type": "foaf:Person",
  "foaf:family_name": "Raufi",
  "foaf:givenname": "Bujar",
  "foaf:homepage": {
  "@id": "http://bujarraufi.wordpress.com"
 },
  "foaf:knows": [{
    "@id": "_:N3a2be572e2c740b7a6a7ccb5a9fa693e"
   },
   {
    "@id": "_:Nf01bbdcb52694b1eaf2f804f9b8e99c2"
   }],
  "foaf:mbox_sha1sum": "637dd17cc2435eeabb90135e2b1817b90f26c966",
  "foaf:name": "Bujar Raufi",
  "foaf:schoolHomepage": {
   "@id": "http://www.seeu.edu.mk"
  },
  "foaf:title": "Dr.",
  "foaf:workInfoHomepage": {
   "@id": "Assiatant Professor at SEEU"
  },
  "foaf:workplaceHomepage": {
   "@id": "http://www.seeu.edu.mk/~braufi"
  }
```

},
{
"@id": "_:N3a2be572e2c740b7a6a7ccb5a9fa693e",
"@type": "foaf:knows",
"@type": "foaf:Person",
"foaf:mbox_sha1sum": "a4a2d17ca2cbb41635de5476d475971dce70cf28",
"foaf:name": "Florije"
}

Figure 1: JSON Serialization of a FOAF Profile

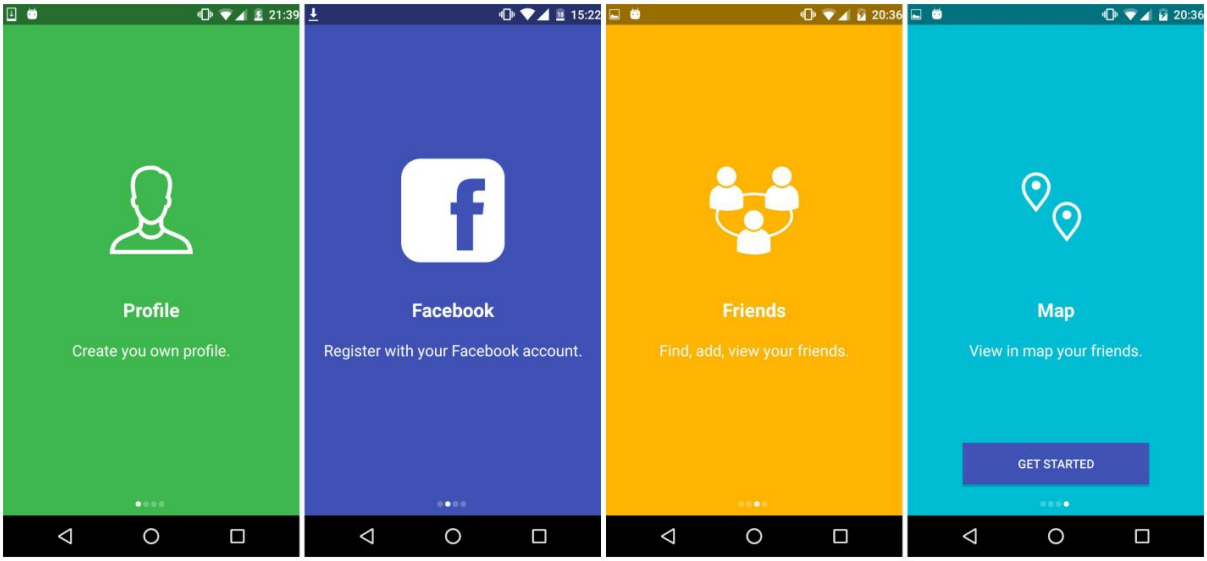
Context-aware systems and context aware computing represent a general class of mobile systems that can sense their physical environment, and adapt their behavior accordingly (Robles, 2010), (Thyagaraju, 2012). Context-aware is used in many applications for purpose to include nearby people, devices, lighting, noise level, network availability, and even in social situations, such as whether a person is with its family or friends or in finding people with similar preferences in vicinity (Robles, 2010). For the design of the application the context-awareness approach is used for allowing access to find similar profiles based on user geographical location as well as their profession.

3.1. Application Architecture

The main components of the application are treated in conjunction with user's interaction with the system. The overall logic of the application is given through the following main actions:

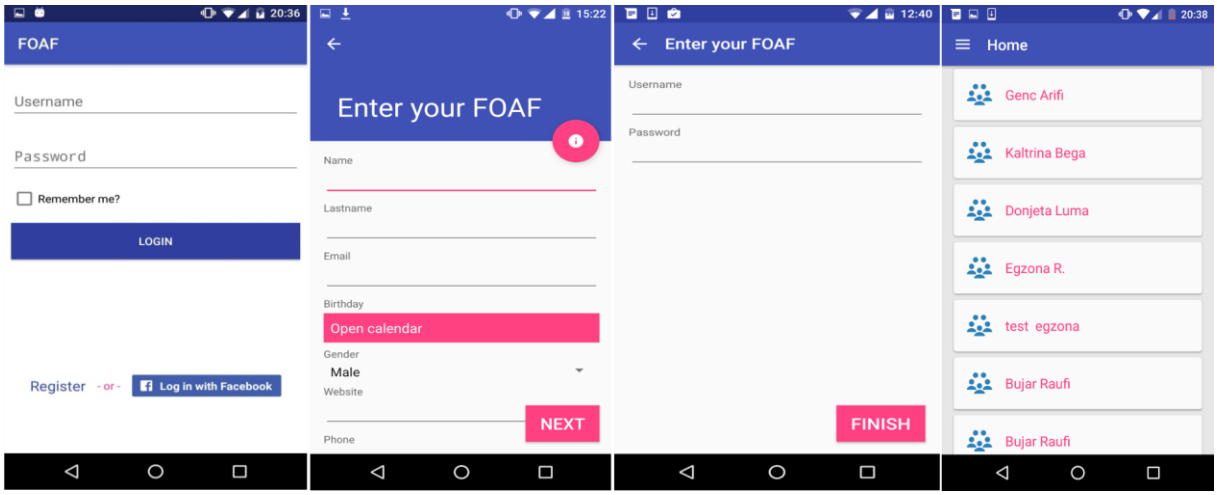
- If a user does not possess an actual FOAF profile. The "Find Professional" app will allow users to create new accounts, data can be also retrieved from their other social network accounts (such as Facebook). The data necessary for creating a Find professional FOAF profiles are:
  - Name
  - Last name
  - Facebook ID
  - Birthday
  - E-mail
- The information about user's birthday and e-mail address can be taken user's accounts if they are publicly available. If a user does not have a Facebook account, a regular registration of a profile goes in the application. Fig 2. Illustrates the initial four interfaces of the Find Professional mobile app.

Figure 2: Initial Interface of the Mobile App



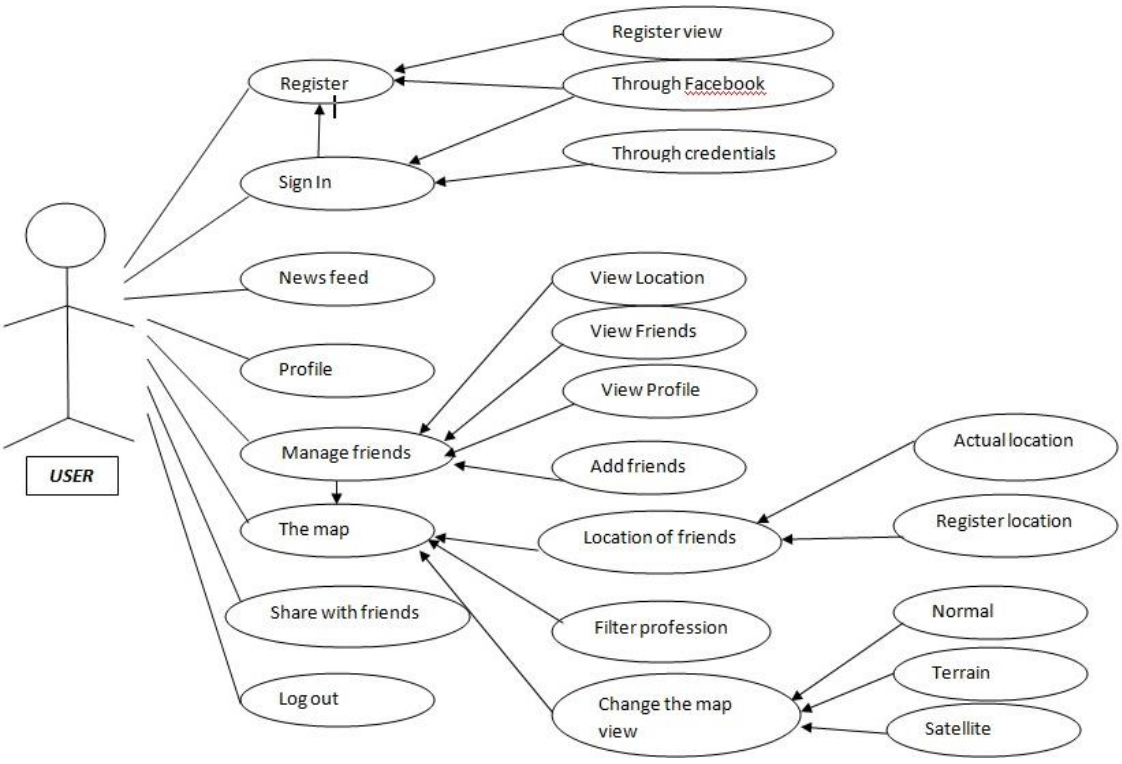
– By sliding the page on the left we will receive different pages' base on the stages of our registration process. I.e. either to login via Facebook or create a new account. If we already have e Facebook account, after login the initial data will be populated to a FOAF profile for creating our Find Professional profile as well as search for created profiles. These interfaces are illustrated in Fig. 3.

Figure 3: User Interface during Login and Search



The overall user activities are depicted in the use case diagram which is illustrated in Fig 4

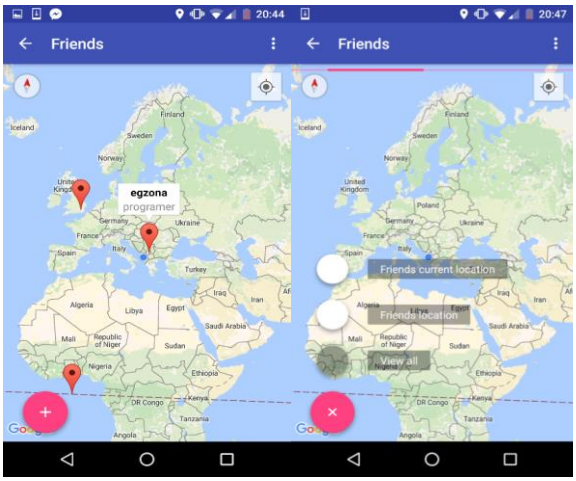
Figure 4: Use Case Diagram for "Find Professional" Mobile App



3.2. Context-Awareness in FOAF Application

The aspect of context-awareness has been achieved through the geographical location of users that share the same or similar profession and is registered via the application. The model is based on ontological context awareness because the application uses the foaf:based_near property. The application lists nearby friends, their names and professions. Fig. 5 illustrates the physical context-awareness as seen from the application interface.

Figure 5: Maps and Appearance of Friend's Locations



4. MOBILE PLATFORM EVALUATION

The evaluation process of the application was performed through a usability testing of the mobile platform. Usability testing represents a common tool which is used to evaluate the usability of a mobile application in a development phase. Usability testing are controlled in environments using a think aloud protocols initially set in a pioneering work of K. A. Ericsson and H. A. Simon's (1980, 1984) (Kaikonnen et al., 2005). For the evaluation purposes, we favoured the field tests against stricter laboratory test. We considered that field tests are more suitable when evaluating usability of mobile applications, considering that the context affects usage and performance. This was also hypothesised and favoured in (Kaikonnen et al., 2005). The procedure of evaluation of the mobile platform followed a strict procedure as described in (Sun et. al, 2013).

4.1. Setup

The study took place on the street by presenting the application to ordinary users and briefing them shortly about what an application was about. This user experience actions comprised tasks that were given after the application briefing occured and most of tested users were unfamiliar with each other. The user experience was therefore typical of that encountered during causal and unformal interviews. Information was collected on the site where a short survey was fullfilled.

4.2. Experimental Scenarios

The successful uilization of scenarios takes into account the diversity of contexts encountered by users. The scenarios where used in the application presented in the form of tasks that should be fullfilled by the user. Tasks that were given to users in different contexts are as follows:

1) Getting started: Download and Install FOAF Application
2) Finding and opening the FOAF application
3) Open a FOAF Profile through Facebook, if no account available Register via application
4) Navigating the main menu functionalities
5) Searching for other FOAF profiles
6) Looking up for nearby profiles based on the same profession
7) Utilizing the map for finding the similar users based on professions
8) Closing the application

Besides these tasks, we have also introduced some intermediary subtasks that varied from user to user.

4.2. Participants

Ten participants were recruited at a random basis. Their ages ranged from 24 to 55 years, with a mean of 29.7, and as much as equal gender split. A range of occupations were represented, including accountant, mechanics, computer scientist, architect, SEO Expert, and Higher Education professional; four participants were university students. Recruitment criterion was the possesion of a smartphone longer than a year.

4.3. Field Study Results

The field experiments generated substantial amounts of rich and valuable data in relatively short time. It is worth mentioning that participants in the testing emphasized the problems of mobile usage rather than simply application usability, and typically those problems were expressed in the form of occasional language. To the participants, the field experiment environment felt moderately informal, and the users talked freely about the use of the application and their sentiments. Users expressed how the field experiment granted them the relaxed look and feel and the possibility to communicate with the researcher as they undertook the evaluation scenarios. For example, users were concerned about the interface such as colors, background and colour vividness. Even though , there were cases where users generally expressed broader views and were able to give a wide range of evaluation-related information during the experiment, such as expressing ideas related to requirements concerning context awareness. A particular example this sort of data generated during the field study was the need for the mobile content and delivery to be highly integrated into the temporal flow of the searching proceess while finding similar matches based on profession. After the experiment, users filled a survey where they had to evaluate each task from 1 to 5, where 1- indicating poor and 5-indicating excellent experience. The detailed list of usability testing for each task i given as in Table 1.

Table 1: Usability Testing on Field Users, Their Task and Requirements

Name	Age	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8
User1	55	5.00	5.00	5.00	5.00	4.00	3.00	2.00	5.00
User2	25	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
User3	24	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
User4	25	5.00	3.00	2.00	5.00	5.00	4.00	3.00	4.00
User5	30	5.00	5.00	5.00	4.00	5.00	5.00	5.00	5.00
User6	26	5.00	2.00	3.00	4.00	5.00	4.00	5.00	5.00
User7	25	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
User8	24	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
User9	25	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
User10	38	4.00	3.00	5.00	5.00	5.00	5.00	5.00	4.00

The user experience assesment was conducted against a paper based detailed design of the mobile prototype application as control measurement. To ensure that the user interface itself (rather than prior knowledge of the user) was not majorly influencing the experiment results, the user-initiated interface tasks were evaluated by calculating the percentage of tasks completed by participants and analysing user comments and matching them against detailed design protoytype taken as a measurement control. Table 2 depicts the overall user experience assesment for each task given in experimental scnearios. The methodology followed is used in (Grawe, 2016).

Table 2: Overall User Experience Evaluation for Each Task

Task	Friedman (8 related Samples)	Multiple paired comparisons $\alpha = .05$
Getting started: Download and Install FOAF Application	$\chi^2(7) = 49.0,$ $P < .002$	$N = 10, \lvert R_{paper} - R_{prototype} \rvert = 1.0^*$
Finding and opening the FOAF application	$\chi^2(7) = 43.0,$ $P < .002$	$N = 10, \lvert R_{paper} - R_{prototype} \rvert = 7.0^*$
Open a FOAF Profile through Facebook, if no account available Register via application	$\chi^2(7) = 45.0,$ $P < .002$	$N = 10, \lvert R_{paper} - R_{prototype} \rvert = 5.0^*$

Navigating the main menu functionalities	$\chi^2(7) = 48.0, P < .002$	$N = 10, \lvert R_{paper} - R_{prototype} \rvert = 2.0^*$
Searching for other FOAF profiles	$\chi^2(7) = 49.0, P < .002$	$N = 10, \lvert R_{paper} - R_{prototype} \rvert = 1.0^*$
Looking up for nearby profiles based on the same profession	$\chi^2(7) = 46.0, P < .002$	$N = 10, \lvert R_{paper} - R_{prototype} \rvert = 4.0^*$
Utilizing the map for finding the similar users based on professions	$\chi^2(7) = 45.0, P < .002$	$N = 10, \lvert R_{paper} - R_{prototype} \rvert = 5.0^*$
Closing the application	$\chi^2(7) = 48.0, P < .002$	$N = 10, \lvert R_{paper} - R_{prototype} \rvert = 2.0^*$

* Indicates a insignificant difference between user experience tasks done in the experiment and those expected in the control values user activities in the paper detailed design

From the overhead data it can be seen clearely that the usability testing process did not yield any severe problems during application evaluation and comments are more related to user interface rather than difficulty in user task completion.

5. CONCLUSION AND FUTURE WORK

In this paper we introduced a social network based mobile application that utilized the FOAF ontology for profile creation, navigation and profile matchmaking and finding based on professions given in those profiles. The application was evaluated with a usability testing performed against ten field users. The results indicated no severe problems in using the mobile application.

Future research directions would involve:

- Creating FOAF profiles by migrating other social network accounts such as Twitter and LinkedIn and incorporating in the application. LinkedIn accounts would be the next logical step considering that it represents a social network where professionals share their expertise.
- Using and linking to other semantic data such as DBPedia and widening the search and reasoning over the semantic data retrieved.

We firmly consider that applications of this type in the future would yield semantically enriched data to be widely available on the web and as a result, it will bring the mobile platforms closer to Linked Open Data in particluar and Semantic Web in general.

REFERENCES

Allemang, D., & Hendler, J. (2011). Semantic Web for the Working Ontologist: Effective Modelling in RDFS and OWL. Elsevier.

Becker, C., & Bizer, C. (2008). DBpedia Mobile: A Location-Enabled Linked Data Browser. Ldow, 369, 2008.

Becker, C., & Bizer, C. (2009). Exploring the geospatial semantic web with dbpedia mobile. Web Semantics: Science, Services and Agents on the World Wide Web, 7(4), 278-286.

Berners-Lee, T., & Jaffe, J. (2016). W3C DATA ACTIVITY Building the Web of Data. Retrieved February 15, 2017, from https://www.w3.org/2013/data/

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data-the story so far. Semantic services, interoperability and web applications: emerging concepts, 205-227.

Brickley, D., & Miller, L. (2016). The Friend Of A Friend (FOAF) vocabulary specification, November 2007. URL http://xmlns. com/foaf/spec

Chen, G., & Kotz, D. (2000). A survey of context-aware mobile computing research (Vol. 1, No. 2.1, pp. 2-1). Technical Report TR2000-381, Dept. of Computer Science, Dartmouth College.

Dan, O., & Davison, B. D. (2016). Measuring and Predicting Search Engine Users' Satisfaction. ACM Computing Surveys (CSUR), 49(1), 18.

Ding, L., Finin, T., & Joshi, A. (2004). Analyzing social networks on the semantic web. IEEE Intelligent Systems (Trends & Controversies), 8(6), 815-820.

Edoh-Alove, E., Bimonte, S., Pinet, F., & Bédard, Y. (2016). New design approach to handle spatial vagueness in spatial OLAP datacubes: application to agri-environmental data. In Geospatial Research: Concepts, Methodologies, Tools, and Applications (pp. 1859-1880). IGI Global.

Gantz, J., & Reinsel, D. (2016). THE DIGITAL UNIVERSE IN 2020–Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East, IDC IView.

Grawe, P. H. (2016). Nonparametric Statistics for the Behavioral Sciences. Numeracy: Advancing Education in Quantitative Literacy, 9 (1)

Heath, T., & Bizer, C. (2011). Linked data: Evolving the web into a global data space. Synthesis lectures on the semantic web: theory and technology, 1(1), 1-136.

Henzinger, M. R., Motwani, R., & Silverstein, C. (2002, September). Challenges in web search engines. In ACM SIGIR Forum (Vol. 36, No. 2, pp. 11-22). ACM.

Höffner, K., Lehmann, J., & Usbeck, R. (2016, October). CubeQA—Question Answering on RDF Data Cubes. In International Semantic Web Conference (pp. 325-340). Springer International Publishing.

Jiang, C. (2007). A location-based service provides Bluetooth cell phone user rich content related to objects/people nearby (Doctoral dissertation, New York University).

Jung, J. J. (2010). Integrating Social Networks for Context Fusion in Mobile Service Platforms. J. UCS, 16(15), 2099-2110.

Kaikkonen, A., Kekäläinen, A., Cankar, M., Kallio, T., & Kankainen, A. (2005). Usability testing of mobile applications: A comparison between laboratory and field-testing. Journal of Usability studies, 1(1), 4-16.

Lei, Y., Uren, V., & Motta, E. (2006, October). Semsearch: A search engine for the semantic web. In International Conference on Knowledge Engineering and Knowledge Management (pp. 238-245). Springer Berlin Heidelberg.

Meng, R., Chen, L., Tong, Y., & Zhang, C. (2017). Knowledge Base Semantic Integration using Crowdsourcing. IEEE Transactions on Knowledge and Data Engineering.

Protocol, S. P. A. R. Q. L. (2016). RDF Query Language. RDF Data Access Working Group (DAWG) Std.

Rana, J., Kristiansson, J., Hallberg, J., & Synnes, K. (2009, July). An architecture for mobile social networking applications. In Computational Intelligence, Communication Systems and Networks, 2009. CICSYN'09. First International Conference on (pp. 241-246). IEEE.

Robles, R. J., & Kim, T. H. (2010). Review: context aware tools for smart home development. International Journal of Smart Home, 4(1).

Schaffert, S., Eder, J., Grünwald, S., Kurz, T., & Radulescu, M. (2009, May). KiWi–a platform for semantic social software. In European Semantic Web Conference (pp. 888-892). Springer Berlin Heidelberg.

Serena, F. D. (2017). U.S. Patent No. 9,536,268. Washington, DC: U.S. Patent and Trademark Office.

Sheth, A. (2009). Citizen sensing, social Signals, and enriching human experience. IEEE Internet Computing, 13(4).

Sun, X., & May, A. (2013). A comparison of field-based and lab-based experiments to evaluate user experience of personalised mobile devices. Advances in Human-Computer Interaction, 2013, 2.

Swaminathan, S. N., & Elmasri, R. (2016, June). Quantitative Analysis of Scalable NoSQL Databases. In Big Data (BigData Congress), 2016 IEEE International Congress on (pp. 323-326). IEEE.

Thyagaraju, G. S., & Kulkarni, U. P. (2012). Design and implementation of user context aware recommendation engine for mobile using Bayesian network, fuzzy logic and rule base. International Journal of Pervasive Computing and Communications, 8(2), 133-157.

Torres, B. P., & González, A. G. (2017). Evolution of the Semantic Web Towards the Intelligent Web: From Conceptualization to Personalization of Contents. In Media and Metamedia Management (pp. 419-427). Springer International Publishing.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

STUDY OF THE INFLUENCE OF WASTE OF PLASTIC ON THE MECHANICAL BEHAVIOR OF MODIFIED BITUMINOUS CONCRETE

DOI: 10.17261/Pressacademia.2017.622
PAP-WCTIE-V.5-2017(60)-p.443-448

N. Guessabi[1], K. Soudani[2], S. Saoula[3], S. Haddadi[4]
1.LEEGO, FGC, USTHB, Algiers, Algeria, nguessabi@yahoo.fr
2.LEEGO, FGC, USTHB, Algiers, Algeria, soudani.khedoudja@yahoo.com
3.LEEGO, USTHB, UMBB, Boumerdes, Algiers, Algeria,, ssaoula1@yahoo.fr
4.LEEGO, FGC, USTHB, Algiers, Algeria,, shaddadi@usthb.dz

ABSTRACT

The objective of this work is the study of the influence of the addition of plastic waste on the behavior of modified asphalt mixtures. Two mechanical tests were conducted on the modified bituminous concrete by the waste of plastic. Different contents of waste were used. The modification was performed by the dry method. The mechanical tests were : Marshall test and the indirect tensile strength test. This study showed that adding waste of plastic to asphalt mix improves their mechanical performances, such as increased stability, compactness, Marshall quotient and indirect tensile strength ITS and decreased Marshall flow and the deformation in the indirect tensile strength test. In addition to the interest of the use of plastic waste in road construction, there is a positive influence on the environment and human life by reducing storage areas of these plastics pollutants

Keywords: Plastic waste, marshall test, indirect tensile strength test.

1. INTRODUCTION

The Road pavement structures are a much newer design. The use of bituminous materials has emerged with the new problems posed by the movement of motor vehicles from the late nineteenth century. Traditional mixes, usually respond satisfactorily to traffic loads. However, changes in the traffic in terms of volume and aggressiveness, and application techniques, both increasingly demanding with regard to the quality of asphalt concrete, has shown some limitations in their jobs and pushed researchers to find solutions for improving the performance of asphalt concrete by adding various compounds. Of these, polymers dominate. For economic reasons related to oil shocks of the 1970s, the road techniques were oriented for over two decades to build roads with thinner layers. These structures must nevertheless ensure mechanical properties at least equivalent to those of traditional pavements. The appearance of special asphalt concrete, asphalt concrete draining, asphalt concrete high modulus, asphalt concrete resistant to hydrocarbons, etc., requires additives of performance in terms of cohesion, adhesion, resistance to chemicals whose oxidation and mechanical stress that is not available in all circumstances in asphalt mixtures with pure bitumen.

The rutting met remains the most damage on the pavement. The creep of asphalt mixture is responsible (Haddadi, 2007), (Proteau and Paquin, 2001). For reduce this phenomena, various techniques have been undertaken. Among them, modification of asphalt with the polymer became a promising technique (Frantzis, 2004) (Khosla and Zahran, 1989).

Modification of asphalt mixtures with compatible polymers or waste can be beneficial. It can be performed:

- by modifying of bitumen by addition of polymers, tires, chemical and other, it is the wet process.
- by modifying the asphalt mixture in the plant by adding polymer type additive, crumb from tires (Haddadi, 2007), plastic food packaging waste (Guessabi, 2011) and others, it is the dry process.

To satisfy the sustainability criteria of road surfaces, it should develop an effective bituminous mixture, which is modified bituminous concrete.

The objective of this work is to study the resistance of wearing rutting under the effect of traffic, by modifying the asphalt concrete with a polymer, which is the plastic food packaging waste.

2. MATERIALS AND METHODS

2.1.The Bitumen

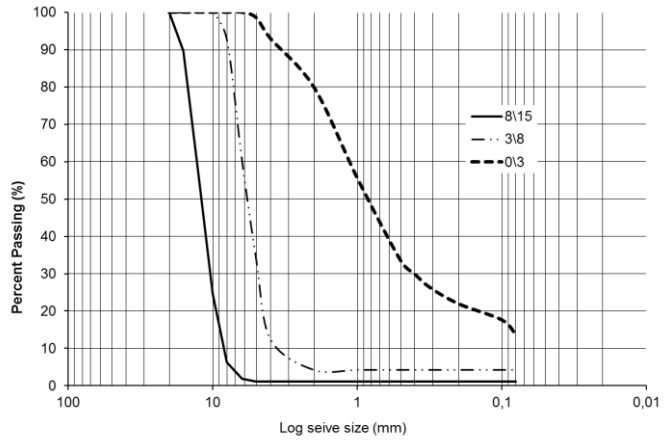
The bitumen (35/50) comes from refineries of NAFTAL (Algeria). Conventional tests are those of the penetration at 25 °C (NF EN 1426), of the ball and ring temperature (NF EN 1427) and ductility at 25 ° C. The results obtained are:

TBA: 51 ° C; Penetration: 40 (1 / 10mm); Ductility :> 100 (cm). These characteristics are in accordance with the specifications recommended for the class (35/50).

2.2.The Aggregates

Three crushed granular classes (0/3 sand, 3/8 and 8/15 gravel from the quarry of El Hachemia (wilaya of Boumerdes). Chemical analysis was performed on these aggregates where we found that they are natural limestone with high CaCO3 content of the three granular fractions. After various tests (LA, MDE, cleanliness, etc.) (NF EN 1097-1) (NF EN 1097-2). The selected materials present good qualities of intrinsic characteristics. The figure 1 shows the particle size analysis of the granular fractions.

Figure 1: Particle Size Analysis of Aggregates



2.3.Identification of the Modifier Used

The additive used in this study is a waste plastic of food; they are packing bags. The preparation of plastic waste is made in the laboratory of roads. The procedure is to cut the plastic bags in varying fiber length of 2 to 5 mm (Guessabi, 2011). Figure 2 shows the plastic waste used in this study.

Figure 2: The Plastic Waste

3. FORMULATION OF ASPHALT CONCRETE

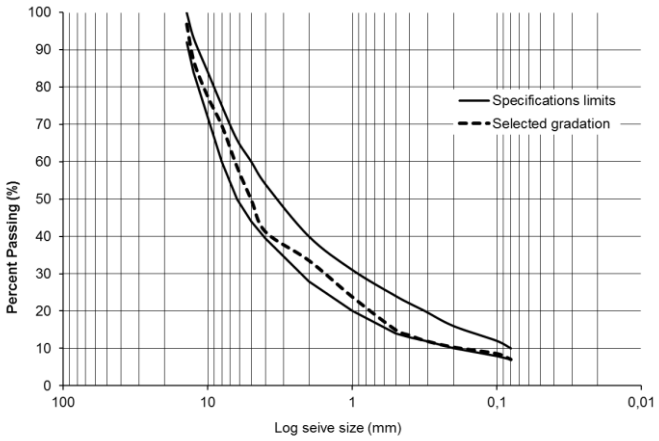
The particle size distribution curve shown in Figure 3 that is inserted in the specification curve is obtained for the following proportions (Guessabi, 2011):

- 40% of sand 0/3

- 30% of the fraction 3/8

- 30% of the fraction 8/15

Figure 3: Mixture Curve



In this formulation study, for the formula used (40% of 0/3, 3/8 30%; 30% 8/15), the optimum content is obtained for a bitumen content of 5.77% giving the best Marshall test characteristics. This content will be considered in this study as that of the bituminous concrete of control or reference without modification.

4. THE MODIFIED BITUMINOUS CONCRETE BY DRY METHOD

The bituminous mix is manufactured according to the previously defined proportions of the different fractions of aggregates and bitumen (NF P98-250-1, 1992), the amount of the additive is chosen as a weight percentage of the aggregate. The choosen percentages are 0.1; 0.2; 0.3; 0.5; 0.6; 0.7; 0.8 and 1% of plastic waste. The additive is incorporated with the aggregate skeleton during mixing operation, the bitumen is introduced thereafter. Once concrete bituminous Marshall type specimens are made.

To assess the impact of plastic waste on the mechanical performance of the mix, we performed the following Marshall Test (NF EN 12697-34) and indirect tensile strength test (NF EN 12697-23).

4.1. Results of the Marshall Test

The results of Marshall Test (stability, flow and quotient) are shown in Figures 4 to 6.

Figure 4. Marshall Stability versus Plastic Waste Content

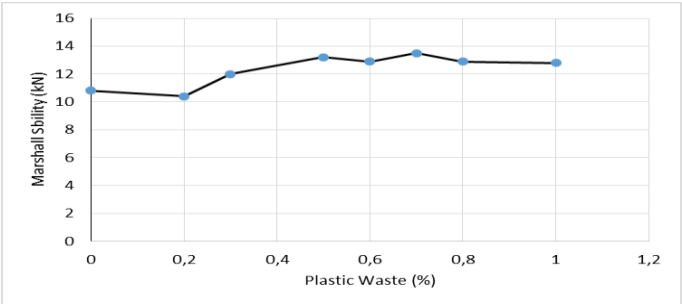


Figure 5: Marshall Flow versus Plastic Waste Content

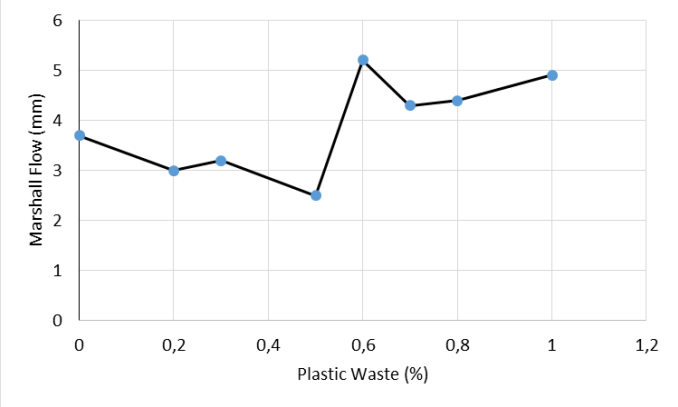
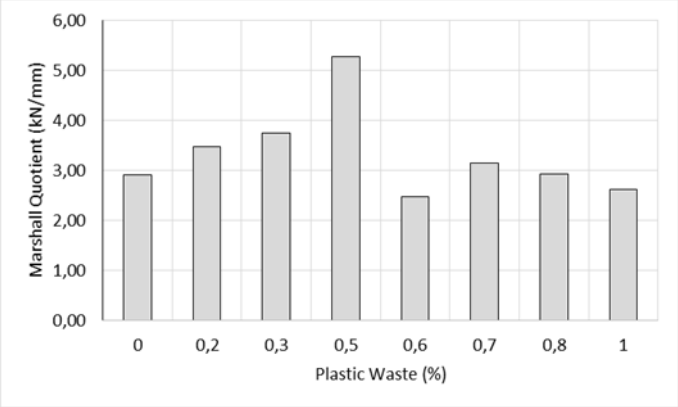


Figure 6: Marshall Quotient versus Plastic Waste Content



The curves of Figures 4 to 6 show that the incorporation of plastic waste improves the Marshall quotient in the bituminous mixture and improves so the resistance for permanent deformation (Haddadi al, 2008). The best results are obtained for the percentage of plastic waste 0.5% (stability, flow and quotient).

4.2. Test Results of the Indirect Tensile Strength

The indirect tensile strength is the maximum tensile stress calculated from the maximum load at the rupture. We chose the contents of 0%, 0.3%, 0.5% and 0.7%. The bituminous concrete to 0% is the reference mixture.

The results of the indirect tensile strength test for both temperatures 20 °C and 40 °C are shown in Figures 7 to 9.

Figure 7: The Maximum Breaking Load to the Indirect Traction Test

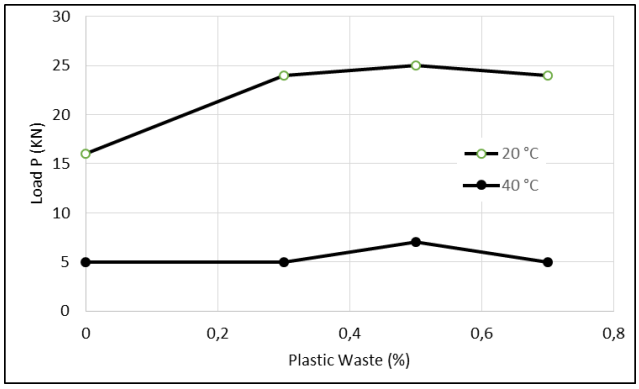


Figure 8: The Deformations According to the Plastic Waste Content to the Indirect Tensile Test

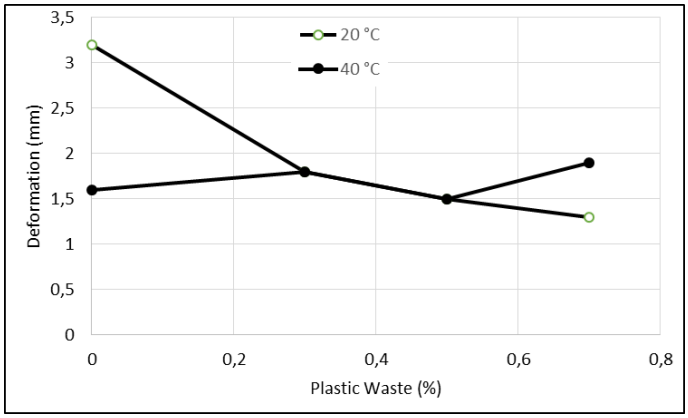
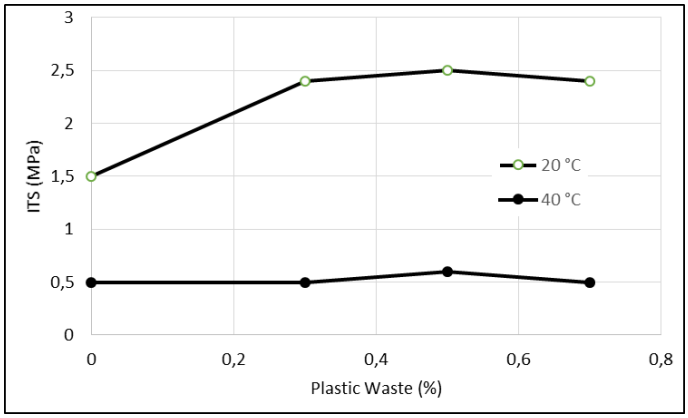


Figure 9: ITS According to the Plastic Waste Content



The curves of the test of indirect traction show that the incorporation of plastic waste substantially improves the resistance at 20 ° C but at 40 ° C the improvement is less.

5. CONCLUSION

The addition of plastic waste on asphalt highlighted the following experimental results (Guessabi, 2011):

5.1. Marshall Test

- The modification of the bituminous mix by the plastic waste improves its mechanical characteristics, increasing the Marshall stability and reduction of flow and consequently improves the Marshall quotient.

- The modified bituminous concrete at 0.5% give the best mechanical performances (stability and flow) for the Marshall test compared to the other mixtures.

5.2. Indirect Tensile Test

- The best resistance to indirect tensile is obtained at 20 °C and 40 °C by the asphalt concrete at 0.5% content of plastic waste

- The mixture of 0.5% of plastic waste gives greater rigidity than that of the bituminous concrete control.

REFERENCES

Haddadi S. (2007). Influence de la poudrette de caoutchouc sur le comportement au fluage des Enrobés bitumineux .*Thèse de Doctorat d'état*, FGC/USTHB FGC/USTHB 2007.

Proteau M., and Paquin Y. (2001). Contribution de différents bitumes purs et bitumes modifiés par ajout de polymères à la résistance à l'orniérage. *Revue Générale Routes et Aérodromes, 793, 34-39.*

Frantzis, P. (2004). "Crumb Rubber-Bitumen Interactions: Diffusion of Bitumen into Rubber." *J. Mater. Civ. Eng., 10.1061/(ASCE)0899-1561(2004)16:4(387), 387-390.*

Guessabi N. (2011). Guessabi.N, Etude de l'influence du déchet plastique sur le comportement à la traction indirecte des enrobés bitumineux modifies, M*émoire de magister, USTHB/FGC, Algerie.*

Khosla N.P. and Zahran S.Z. (1989). A mechanistic evaluation of mixes containing conventional and polymer modified (styrelf) asphalts. Proc Assoc Asphalt Paving Technol ; 58:274–302.

Haddadi S., Ghorbel E. and Laradi N. (2008). Effects of the manufacturing process on the performances of the bituminous binders modified with EVA. Construction and Building Materials, Volume 22, 2008, pp 1212–1219.

PressAcademia Procedia

2nd World Conference on Technology, Innovation and Entrepreneurship
May 12- 14, 2017, Istanbul, Turkey. Edited by Sefer Şener

EFFECT OF NATURAL POZZOLAN ON CONCRETE DURABILITY

DOI: 10.17261/Pressacademia.2017.623
PAP-WCTIE-V.4-2017(61)-p.449-452

Ahcene. Merida[1], Fattoum. Kharchi[2]
[1]University of Bab Ezzouar (USTHB), BP 32 Bab Ezzouar, EL Allia, Algiers, Algeria.merida.hacene@yahoo.fr
[2]University of Bab Ezzouar (USTHB), BP 32 Bab Ezzouar, EL Allia, Algiers, kharchifcong@yahoo.fr

ABSTRACT
Sulfate attack and its effects are important from both scientific and industrial viewpoints. It is perceived that cements containing pozzolan have better performance in sulfate solutions, since the pozzolanic reactions reduce the quantity of calcium hydroxide and increase calcium silicate hydrate. This paper investigates the physical and mechanical properties of concretes made by blended cement containing Algerian natural pozzolan of volcanic origin, and Portland cement. In order to better determine the pozzolanic effect of the pozzolan addition in the concrete, the analysis of the experimental results of the effect of the partial replacement of the cement by the natural pozzolan showed that it contributes positively to the improvement of its mechanical characteristics, its durability with respect to the permeability to the chlorine ions, the ultrasonic pulse velocity, the compressive strength as well as the sulphate resistance. The present study confirms the pozzolanic reactivity of the natural pozzolan used

Keywords: Natural pozzolan, durability, pozzolanic reactivity, sulphate.

1. INTRODUCTION

The durability of concrete structures is affected by many environmental factors, the sulphate corrosion being one of the most frequent and detrimental processes.Through the capillary pores of concrete due to the concentration gradient[1,2] and react with unhydrated components of the hardened cement paste. In consequence, these chemical reactions may lead to expansive reaction products such as ettringite ($C3A.3CaSO_4.32H_2O$**)** [3]. In turn, the ettringite may cause the overall expansion of a structural element and its extensive damage progressing from the outer surface towards the specimen inner core [4].This process may result in a gradual loss of concrete strength [5] accompanied by surface spalling and exfoliation [6].Pozzolanic materials improve the microstructure of concrete due to their particle size, and may alter chemical composition and hydration reactions. Pozzolan as an amorphous or glassy silicate material that reacts with calcium hydroxide formed during the hydration of Portland cement in concrete. The substance that contributes to the strength of the concrete called calcium silicate hydrates(C-S-H) [6]. Calcium hydroxide will reduce the strength of the concrete. Pozzolan contains silica that react with calcium hydroxide in concrete to form extra calcium silicate hydrates compound and diminish calcium hydroxide [7], further strengthening the concrete due to increase of C-S-H compound and making it stronger, denser, and durable during its service life. Many researches on the performance of concretes containing pozzolan in sulfate solutions have been performed [8-10].

The aim of this study is to experimentally investigate the effect of replacing 5% of cement by natural volcanic pozzolan in the mixture of high performance concrete (HPC) on the compressive strength, permeability to the chlorine ions, sulphate resistance and ultrasonic pulse velocity of specimens exposed to solutions of 5% sodium sulphate (Na_2SO_4)in comparison with traditional concrete (CC).

The specimens were stored for one year in drinking water (environment 1) and in aggressive solution containing 5% sodium sulphate (environment 2).

2.EXPERIMENTAL PROGRAM

2.1 Materials

The materials investigated in this paper are: aggregates, cement, natural volcanic pozzolan and superplasticiser.

2.1.1 Aggregates

Natural rolled sand obtained from a local river and crushed limestone with a maximum particle size of 16 mm made fine and coarse aggregates respectively. The size, the finesse modular (FM = 3.2), the sand equivalent value (SEV= 97%) and resistance to shocks (33%) confirmed suitability of both aggregates for high performance concrete (HPC).

Sand density = 2, 60 g/ cm^3

Coarse aggregates Density = 2, 50 g/ cm^3

2.1.2 Cement

Portland cement (CPA-CEM-I / A 42.5) conforming to the Algerian standard NA 443, EN 197-1 and the AFNOR 15-301/94 NFP1 (which is mainly based on the European EN197-1).

Specific gravity = 3100 Kg/m^3

Specific surface = 322 m^2/Kg

2.1.3 Natural Pozzolan

Natural volcanic pozzolan, extracted from the deposit Beni-saf (Algeria) was used as supplementary cementing material.

Specific gravity = 2660Kg/m^3
Specific surface = 960 m^2/Kg

Pozzolanic activity = 110 mg CaO/g.

The pozzolan content in the mix was fixed at 5% by weight of cement

2.1.4. Reducing Superplasticizer

The study of concrete composition is always to seek simultaneously two essential qualities: strength and workability, but these two qualities are linked to each other but vary in the opposite direction. The idea was to develop a dense concrete from a compact granular skeleton using cement and water and meeting the strength, durability and workability requirements. The optimized superplasticiser content was 2% at 0.3 W/C ratio giving a slump of about 21cm.

3. CONCRETES COMPOSITION

The study of the concrete composition is to define the optimal dosage of aggregates, cement and water to make a concrete with required qualities: strength and durability. This study used “Dreux Gorisse” method of mix proportioning which is based on the size analysis (sand and gravel different fractions) to investigate the use of natural pozzolan on the performance properties of concrete, two different concrete mixes were employed, details of which are given in table 1.The control mix and CC contained only Portland cement, mix of HPC the Portland cement was partially replaced with 5% natural pozzolan (by weight). All concrete mixtures were prepared according to ASTM C 192 standard. The super plasticizer was added at the time of mixing.

Table 1. Mixture proportions and properties of concrete

Concrete	W/ C	Cement, kg/m^3	Pozzolan kg/m^3	Water kg/m^3	Gravel 3/8 kg/m^3	Gravel 8/16 kg/m^3	SPa %	Sag cm	Density kg/m^3
CC	0,5	425	0	212,5	137	837	0	8	2430
HPC	0.3	403.75	21.25	107.66	137	837	2	21	2596

4. TEST METHODS

Compressive strength**:** This test was carried out in accordance with ASTM C39.

Chloride permeability: This test was performed using the procedures of ASTM C 1202.

Ultrasonic Pulse Velocity: This test was carried out in accordance with ASTM C597-02.

Sulphate resistance: This test was carried out in accordance with ASTM C1012.

5. RESULTS

Fig. 1. Evolution of compressive strengths

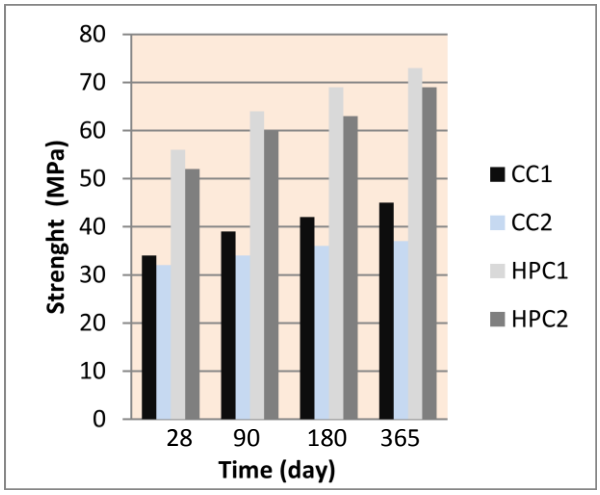


Fig. 2. Variation of charge passed

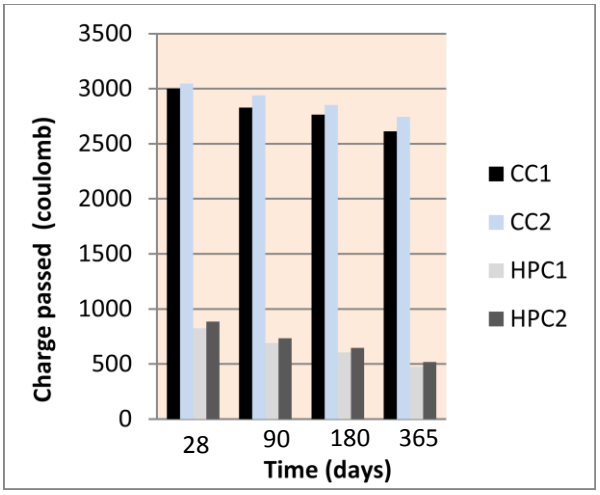


Fig. 3. Variation of velocities

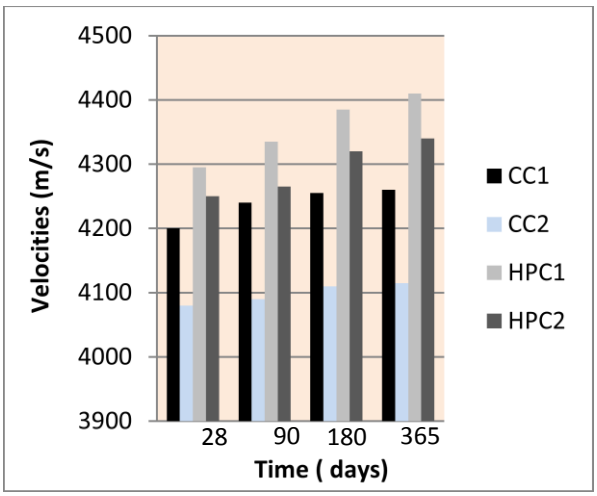
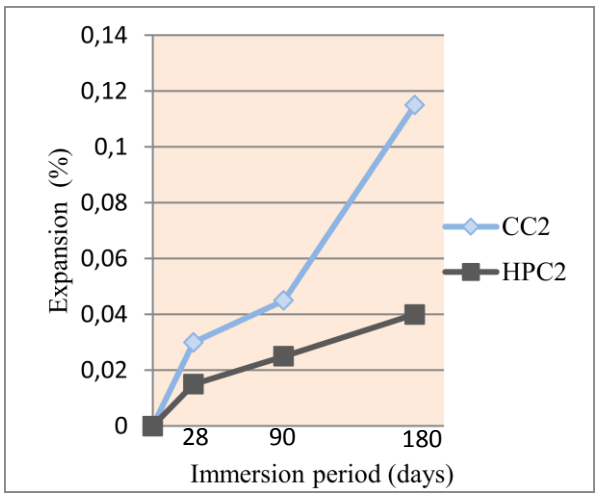


Fig . 4. Results of sulphate resistance



6. CONCLUSION

The specimens kept in water, the increase in compressive strength continuous as the duration of immersion increases. The specimen concretes curing in water, the strength of control concrete increase from 34 (MPa) to 45 (MPa), whereas the high performance concrete it increase from 56 (MPa) to 73 (MPa). The results indicates that pozzolan addition helps gain compressive strength.

The specimens kept in aggressive solution; the strength of the control concrete is reduced by 17.77(%) whereas the high performance concrete the reduction was by (5.48%) only.

The specimen concretes curing in water, the charge passed of the control concrete decrease from 3000 (coulomb) to 2614 (coulomb), whereas the high performance concrete it decrease from 825(coulomb) to 476 (coulomb).

The specimens kept in aggressive solution, the charge passed of the control concret is decrease by 4, 73 (%) whereas the high performance concrete the decreasing was by (8, 3%).The specimen concretes curing in water, the velocities of the control and pozzolan concretes increase from 4200 to 4260 (m/s) and 4295 to 4410 (m/s) respectively in aggressive solution, its decrease by 3, 4 (%) and 1, 58 (%) respectively.

The expansion rates are low at the beginning of the control and high performance concrete, and increases substantially after three months of curing for the first specimen concrete. The expansion of the control concrete and high performance concrete curing in aggressive solution decrease by 65, 21(%).

Based on the obtained data in this study, the use of natural volcanic pozzolan replacing (5%) by weight of cement in the mixture of high performance concrete influences positively the durability specimens concrete cured in sulphate environment. The pozzolan modifies the microstructure of the concrete in terms of its physical and chemical characteristics. It was observed that during the early stages, the filler effect results due to reduction in porosity. With aging, the pozzolanic action further evidence of densification and low porosity of the concrete due to the natural admixture by the formation of CSH with binding properties similar to those formed in mineral-based cements. I can be concluded that the mineral admixture improved the physical characteristics of concrete relativel to the control concrete sample.

REFERENCES

[1] Nadler, B. Schuss, Z.Singer, A. Eisenberg, R.S., Ionic Diffusion ThroughProtein Channels: From molecular description to continuum equations,Technical Proceedings of the Nanotechnology Conference and Trade Show, 2003,Vol. 3, pp. 439-442.

[2] Truc, O.,Prediction of Chloride Penetration into Saturated Concrete– Multispecies approach,Chalmers University of Technology, Gutenberg, Sweden, INSA, Toulouse, France2000.

[3] Atkinson, A. Haxby, A. and Hearne, J.A., The Chemistry and Expansion of Limestone-Portland Cement Mortars Exposed to Sulphate-Containing SolutionsNIREX Report NSS/R127, United Kingdom, 1988.

[4] Skalny,J. Marchand, J. and Odler., I, Sulphate Attack on Concrete, Spon Press, London, United Kingdom, 2002.

[5] Tixier, R.Mobasher, B., Modeling of damage in cement-based materials subjected to external sulfate attack – part 1: formulation. ASCE J Mater Eng,2003, Vol.15, N°.4, pp.305–313.

[6] Bentz, DP. Garbicz, EJ.,Simulation studies of the effects of mineral admixtures on the cement paste aggregate interfacial zone. ACI Mater J, 1991,Vol.88, N°.5, pp.518–529.

[7] Shamaran, M.Kasap, O.Duru, K. and Yaman, IO., Effect of mix composition and water cement ratio on the sulfate resistance of blended cements, Cement and ConcreteComposites, Vol. 29,N°.3, 2007, pp.159-167.

[8] Saricimen, H.Shameem, M. Barry, MS. Ibrahim, M. and Abbasi, TA.,Durability of proprietary cementitious materials for use in wastewater transport systems, Cement and Concrete Composites, Vol. 25, N°4-5, 2003, pp. 421-427.

[9] Chang, ZT. Song, XJ. Munn, R. and Marosszeky, M.,Using limestone aggregate and different cements for enhancing resistance of concrete to sulfuric acid attack, Cement and Concrete Research, Vol. 35, N°.8, 2005, pp. 1486-1494.

[10] Sersale, R.Frigione, G. and Bonavita L., Acid depositions and concrete attack: Maininfluences, Cement and Concrete Research, Vol. 28, N°.1,1998, pp.19-24.