



Balkan Journal of Electrical & Computer Engineering

An International Peer Reviewed, Referred, Indexed and Open Access Journal

www.bajece.com

Vol : 7
No : 4
Year : 2019
ISSN : 2147 - 284X



It is abstracted and indexed in, Index Google Scholarship, the PSCR, Cross ref, DOAJ, Research Bible, Indian Open Access Journals (OAJ), Institutional Repositories (IR), J-Gate (Informatics India), Ulrich's, International Society of Universal Research in Sciences, DRJI, EyeSource, Cosmos Impact Factor, Cite Factor, SIS Scientific Indexing Service, IJIF, iifactor. ULAKBİM-TR Dizin.

General Publication Director & Editor-in-Chief
Musa Yılmaz, Batman University, Turkey.

Vice Editor
Hamidreza Nazarpouya, University of California Riverside, USA

Scientific Committee
Abhishek Shukla (India)
Abraham Lomi (Indonesia)
Aleksandar Georgiev (Bulgaria)
Arunas Lipnickas (Lithuania)
Audrius Senulis (Lithuania)
Belle R. Upadhyaya (USA)
Brijender Kahanwal (India)
Chandar Kumar Chanda (India)
Daniela Dzhonova-Atanasova (Bulgaria)
Deris Stiawan (Indonesia)
Emel Onal (Turkey)
Emine Ayaz (Turkey)
Enver Hatimi (Kosovo)
Ferhat Sahin (USA)
Gursel Alici (Australia)
Hakan Temeltaş (Turkey)
Ibrahim Akduman (Turkey)
Jan Izykowski (Poland)
Javier Bilbao Landatxe (Spain)
Jelena Dikun (Lithuania)
Karol Kyslan (Slovakia)
Kunihiko Nabeshima (Japan)
Lambros Ekonomou (Greece)
Lazhar Rahmani (Algerie)
Marcel Istrate (Romania)
Marija Eidukeviciute (Lithuania)
Milena Lazarova (Bulgaria)
Muhammad Hadi (Australia)
Muhamed Turkanović (Slovenia)
Mourad Houabes (Algerie)
Murari Mohan Saha (Sweden)
Nick Papanikolaou (Greece)
Okyay Kaynak (Turkey)
Osman Nuri Ucan (Turkey)
Ozgur E. Mustecaplioglu (Turkey)
Padmanaban Sanjeevikumar (India)
Ramazan Caglar (Turkey)
Rumen Popov (Bulgaria)
Tarek Bouktir (Algeria)
Sead Berberovic (Croatia)
Seta Bogosyan (USA)
Savvas G. Vassiliadis (Greece)
Suwarno (Indonesia)
Tulay Adali (USA)
Yogeshwarsing Calleecharan (Mauritius)
YangQuan Chen (USA)
Youcef Soufi (Algeria)

Aim & Scope

The journal publishes original papers in the extensive field of Electrical-Electronics and Computer engineering. It accepts contributions which are fundamental for the development of electrical engineering, computer engineering and its applications, including overlaps to physics. Manuscripts on both theoretical and experimental work are welcome. Review articles and letters to the editors are also included.

Application areas include (but are not limited to): Electrical & Electronics Engineering, Computer Engineering, Software Engineering, Biomedical Engineering, Electrical Power Engineering, Control Engineering, Signal and Image Processing, Communications & Networking, Sensors, Actuators, Remote Sensing, Consumer Electronics, Fiber-Optics, Radar and Sonar Systems, Artificial Intelligence and its applications, Expert Systems, Medical Imaging, Biomedical Analysis and its applications, Computer Vision, Pattern Recognition, Robotics, Industrial Automation.



ISSN: 2147- 284X
Vol: 7
No : 4
Year: October 2019

CONTENTS

- Davut İzci;** Constructing an Electronic Circuitry for Label-free Hall Biosensors.....**366-372**
- Ş. Eraslan, S. Karabulut, M. C. Atalay, Y. Yeşilada;** Evaluation of Visualisation of Scanpath Trend Analysis (ViSTA) Tool.....**373-383**
- H. Gençoğlu, T. Yerlikaya;** Three Part Hybrid Encryption Schema.....**384-390**
- D. Köseoğlu, O.Chouseinoglou;** Opinion Analysis of Software Developers Working Onsite on Public Sector Software Projects: An Exploratory Study.....**391-398**
- M. F. Aslan, K. Sabancı, A. Durdu;** Comparison of Contourlet and Time-Invariant Contourlet Transform Performance for Different Types of Noises.....**399-404**
- B. Dogan;** An alignment-free method for bulk comparison of protein sequences from different species.....**405-416**
- M. Akgün;** An Active Genomic Data Recovery Attack.....**417-423**
- H. F. Carlak, E. Kayar;** Volt / VAR Regulation in Energy Transmission Systems Using SVC and STATCOM Devices.....**424-433**
- H. Ademgil;** Simultaneous Sensing of Dual Analyte Photonic Crystal Fiber Based Liquid Sensor.....**434-439**
- Y. Acar, S. Aldırmaz Çolak;** Novel OFDM System Using Orthogonal Pilot Symbols with Subcarrier Index Modulation.....**440-445**
- C. Bakır, V. Hakkoymaz, B. Diri, M. Güçlü;** Comparisons on Intrusion Detection and Prevention Systems in Distributed Databases.....**446-455**
- O. Akar, Ü. K. Terzi, B. K. Tunçalp, T. Sönmezocak;** Determination of the Optimum Hybrid Renewable Power System: A case study of Istanbul Gedik University Gedik Vocational School.....**456-463**

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING

(An International Peer Reviewed, Indexed and Open Access Journal)

Contact

Batman University
Department of Electrical-Electronics Engineering
Bati Raman Campus Batman-Turkey

Web: <http://dergipark.gov.tr/bajece>
<https://www.bajece.com>
e-mail: bajece@hotmail.com

All articles published by BAJECE are licensed under the Creative Commons Attribution 4.0 International License. This permits anyone to copy, redistribute, remix, transmit and adapt the work provided the original work and source is appropriately cited.



Constructing an Electronic Circuitry for Label-free Hall Biosensors

D. IZCI, J. HEDLEY


Abstract— Magnetic field has a huge potential of providing non-destructive and highly efficient detection platforms in biosensing field. It provides a low intrinsic background in biological systems since those systems have no comparable biological signal. Hall devices are currently dominating the market of the magnetic sensors due to several advantages such as allowing miniaturization, being compatible with electronics integration, cheaper fabrication and room temperature operation with high linearity. It is important to construct a suitable front-end circuitry in order to be able to achieve a portable and label-free working Hall biosensor. This paper presents the construction of an electronic circuitry that is required for actuating Hall based biosensors and obtaining the output since there are undesired effects which are causing less sensitive or less accurate results. The paper takes a closer look at those effects and presents the solution to the problem by introducing a developed circuitry on a designed printed circuit board along with the promising results it has achieved.

Index Terms— Hall effect, Hall biosensors, Offset removal, Spinning current circuitry.


I. INTRODUCTION

HALL EFFECT is one of the fundamental techniques used for characterizing electrical transport properties in metals and semiconductors. The basic idea of Hall effect is to drive current through a thin layer of metal sheet or semiconductor and apply a magnetic field perpendicular to the driven current. This exerts a force which is perpendicular to both driven current and applied magnetic field. The exerted force is also known as Lorentz Force [1] and is illustrated in Fig. 1. Transport properties of a material such as carrier mobility, carrier concentration and carrier type can be determined by measuring Hall voltage.

DAVUT İZCI, is with Department of Computer Engineering Batman University, Batman, Turkey, (e-mail: davut.izci@batman.edu.tr).

 <https://orcid.org/0000-0001-8359-0875>

JOHN HEDLEY, is with School of Engineering Newcastle University, Newcastle upon Tyne, UK, (e-mail: john.hedley@ncl.ac.uk).

 <https://orcid.org/0000-0002-2893-8776>

Manuscript received June 22, 2019; accepted September 01, 2019.
DOI: [10.17694/bajece.633908](https://doi.org/10.17694/bajece.633908)

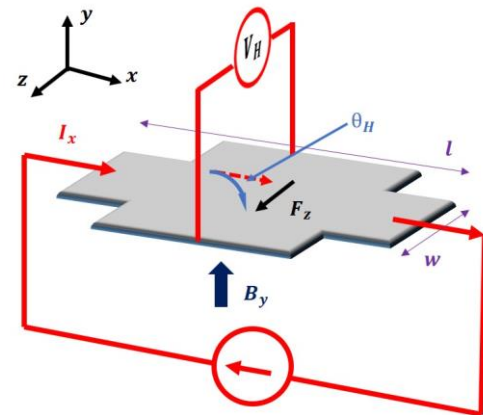


Fig. 1. A representation of Hall sensing principle.

The voltage exerted from the system is proportional with carrier mobility and inversely proportional with carrier concentration and thickness of the material as stated in Eq.(1). In Eq.(1), the term V_H represents the Hall voltage, n denotes the density of the charge carriers (carrier concentration) and it is given in cm^{-2} unit, μ is for carrier mobility with the unit of $\text{cm}^2\text{V}^{-1}\text{s}^{-1}$ and e stands for electron charge ($\sim 1.60 \times 10^{-19}$ coulombs). Additionally, t denotes the thickness of the material in metres, I_x represents the applied current that flows through the material sheet, E_x is the applied electric field and B_y is the term that represents the perpendicular magnetic field. G represents geometrical factor [2] which depends on carrier mobility and length to width ratio (l/w).

$$V_H = G \frac{1}{ne} \frac{B_y I_x}{t} \quad (1)$$

The Hall effect phenomenon has found important roles for many practical applications in last decades by utilizing Hall sensors. Those sensors are magnetic type sensors that are using the perpendicular magnetic field component to produce information and have been investigated for more than a hundred years and used for various purposes since then [3]–[5]. They are used widely in many applications as low-cost sensors. The applicability of Hall principle has also been investigated quite recently for biomedical purposes [6] and a considerable number of reports have been devoted to employability of Hall principle for biosensing applications [7]–[12]. Hall based biosensors have been shown to be successful with the aid of magnetic particles [13]–[16]. In this type of detection, magnetic beads are used to measure magnetic susceptibility [14]. Basically, a combination of

alternating and non-alternating magnetic fields is applied in-plane and perpendicular to the sensor surface, respectively (see Fig. 2). The detection is achieved via observing an output signal which has the same frequency of the in-plane excitation signal.

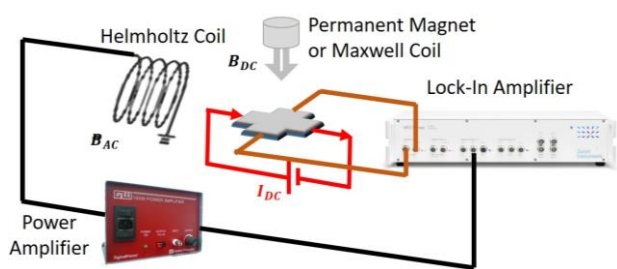


Fig. 2. Schematic view of the ac susceptibility measurements.

However, there are some restrictions that limit the sensitivity of the Hall sensor if magnetic beads are used. It is crucial to consider the separation between surface of the sensor and magnetic bead as magnetic field from the beads decreases proportional to $\sim 1/d^3$ where d represents the separation distance [17]. This parameter must be treated carefully for an optimized signal to noise ratio. Additionally, magnetic moment of a magnetic bead is another important parameter that must be considered since those parameters are affecting the magnetic field provided by magnetic beads. Moreover, it is quite hard to observe the output in real-time for a longer period in this type of detection scheme since it requires more power due to complexity of electronics. Running a coil for longer would also affect the produced magnetic output because of thermal heat being produced. Because of the complexity it requires in terms of electronics setup, as shown in Fig. 2, creating such a portable system is not feasible.

It is important to scale down the system due to necessity of real-time monitoring and point-of-care system requirements. Therefore, a label-free approach seems more attractive. However, an electronic front-end would be required to make a label-free system achievable and cost-effective along with good performance parameters. This paper investigates the steps and presents a solution to make the sensors portable with significantly improved performances.

II. DETECTION MECHANISM FOR LABEL-FREE APPROACH

A Hall effect sensor can be thought of as a combination of three major parts; the transducer itself, power supply to bias the transducer and data acquisition stage for acquiring relevant output. However, those parts require to be used in conjunction with specially designed electronics to perform properly and to present the data to end user. Adopting specially designed electronics significantly improves the system capability by maintaining the performance and providing application specific data, thus, leaving no concern to the end user.

A. Biasing Source

A Hall transducer can, in principle, be used simply by adding a stable power supply to it and biasing can be achieved

either with a voltage or a current source. A simple setup for biasing a Hall device is demonstrated in Fig. 3. In case of a voltage source, the temperature coefficient of the transducer sensitivity will be as high as $0.3\%/^{\circ}\text{C}$ [3], meaning that the output will vary significantly with respect to temperature change. Therefore, some additional front-end circuitry is required to limit the current that will be sourced from the voltage supply and provide more stable output in case of voltage mode of operation [3], [18]. Limiting the current is particularly important to reduce the power consumption, hence, preventing device from thermal heat effect. To obtain a stable output in case of temperature changes, the voltage source or the gain of the amplifier can be made temperature dependent [1] to maintain a stable output, thus, making the system less sensitive to variations of temperature. In terms of implementation, a buffer must be placed in between voltage reference and the Hall transducer for low current transducers where an additional transistor can be added in between the buffer and the device for the transducers requiring high currents.

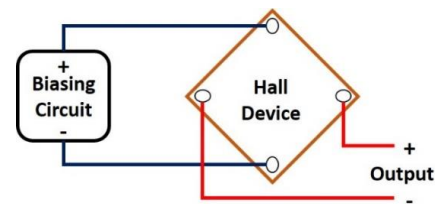


Fig. 3. An illustration of biasing a Hall plate.

Another way of reducing the output variations with respect to temperature changes is to bias the transducer via a current source. Current mode of operation provides a reduction as low as $0.05\%/^{\circ}\text{C}$ in temperature coefficient [3]. As a result, current source biasing is preferable in most cases as it provides more stable working conditions due to significant reduction in temperature coefficient [19]. In terms of current mode of operation, a power supply can simply be used with a combination of resistors for biasing devices. A constant current biasing circuit can be achieved by placing simply a resistor in between voltage reference and the transducer. It can also be obtained via using a transistor by connecting it to the transducer as shown in Fig. 4.

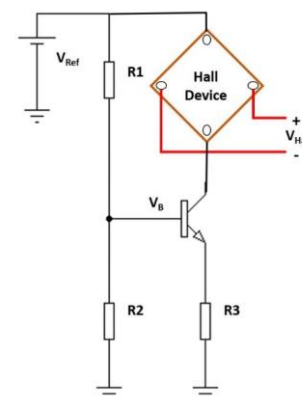


Fig. 4. Biasing Hall transducer via a constant current source using a transistor.

B. Amplification

In general, Hall voltages are quite small and therefore require amplification. Therefore, it is necessary to implement an amplification process since the output of such systems are in millivolt or microvolt ranges. Depending on the structure of the devices, the Hall voltage might be even smaller. The smaller ranges make it difficult to measure since it is below the detectable range of most of the acquisition systems. In the amplification stage, an instrumentation amplifier is desired since Hall devices have differential outputs as seen in Fig. 5.

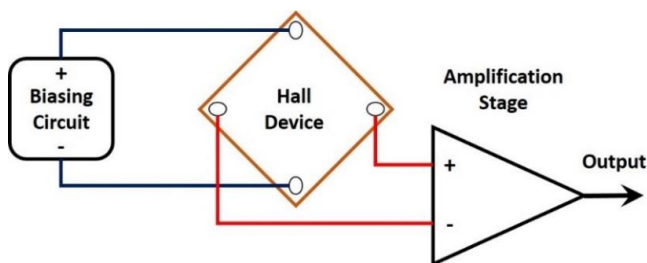


Fig. 5. Implementation of an amplification stage

C. Offset Removal

The offset voltage is an undesired effect which occurs due to contact misalignments, poor surface roughness or non-homogeneous device structure [20] and introduces a non-zero voltage value for the cases of zero magnetic field and keeps increasing with increased current under constant magnetic field. In most cases, this voltage could be even higher than the actual Hall voltage. Fig. 6 represents the parameters that causes the offset voltage. The amplification stage increases the output and help better understanding of the voltage change; however, it still has the limitations in terms of the ability of accurate measurement since the output is amplified together with offset voltage. That means the amplification process also causes the amplification of the offset voltage, as well.

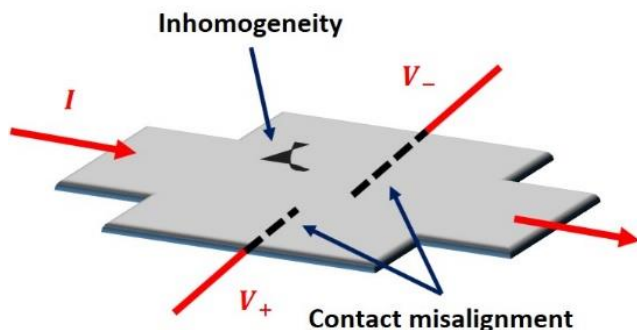


Fig. 6. The parameters that causes offset voltage.

A potentiometer can be used as a tool for eliminating offset voltage by being connected to the output of the amplification circuit as shown in Fig. 7 and could be arranged to provide zero offset. However, this will work for a specific biasing current and will require an adjustment if the amount of biasing current changes. Additionally, thermal effects will be dominant since a potentiometer is in use.

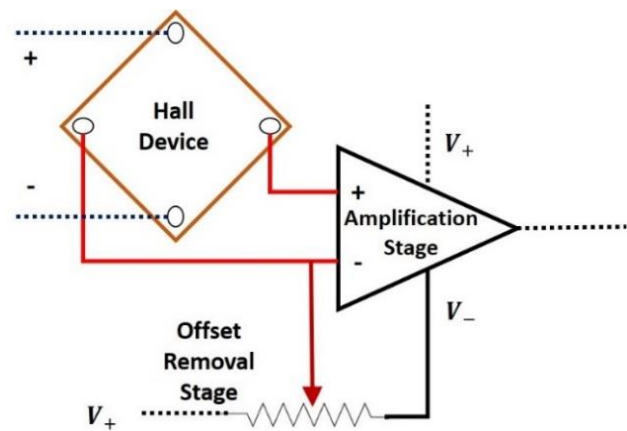


Fig. 7. A basic offset removal stage for Hall devices using a potentiometer.

D. Data Acquisition

Data acquisition can be performed after the amplification and offset removal stages. The amplification stage introduces another issue which might cause a deteriorating effect on actual signal of interest since it leads to a similar amount of amplification of noise signals. Therefore, the output of the amplifier must be filtered with a low pass filter for removal of undesired frequency components. Fig. 8 shows the implementation of a filtering stage. A filter having a specified cut-off frequency can simply be integrated to the circuit via passive components such as capacitors and resistors. Meanwhile, it can also be filtered using a software after acquisition.

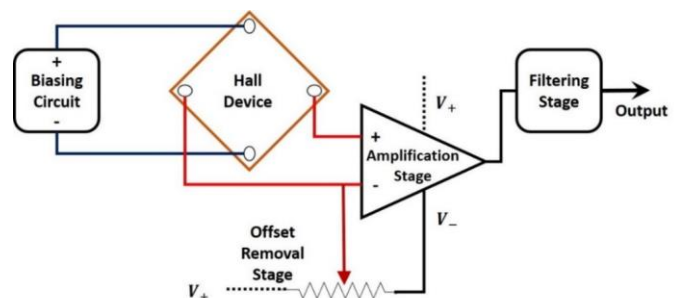


Fig. 8. Basic biasing, amplifying, offset removal and filtering stages.

III. DEVELOPMENT OF AN INTEGRATED CIRCUITRY

The approaches presented in above sections will not be able to remove the offset dynamically since any change in the biasing current leads to a change in output regardless of magnetic field. More importantly, the output is prone to thermal effects in such a design since it becomes very sensitive to temperature changes. To remove the offset voltage, a technique called 'current spinning' was employed [21] which can remove the offset voltage dynamically. Fig. 9 is a representation of a cross shaped Hall device which will be used to explain the spinning method between contacts. Basically, the idea can be explained by a simple approximation of a Hall device structure by using four resistors connected to each other in a Wheatstone bridge structure.

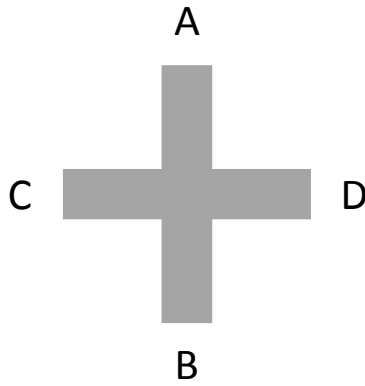


Fig. 9. A representation of cross shape Hall device.

In this case, an offset voltage can easily be introduced by using three resistors with the same values and one with a different value. Once the current is applied from one connection to the other non-neighboring connection a voltage will be measured between other two non-neighboring connection points. If this setup is rotated by 180°, the measured voltage will change in sign but remain as the same value. In case of a Hall device, the Hall voltage would not change since it depends on the exerted force due to applied current and magnetic field. Once the outputs for 0° and 180° are obtained, the Hall voltage can easily be calculated by doing a simple mathematical operation on them. This method is summarized in TABLE 1 by considering the current rotation between all contacts.

TABLE 1. BIASING AND MEASUREMENT CONFIGURATIONS FOR A CROSS SHAPE HALL DEVICE

Current	Measure	Obtained Output
From A to B	Between C and D	$V_H + V_{offset1}$
From C to D	Between B and A	$V_H + V_{offset2}$
From B to A	Between D and C	$V_H - V_{offset1}$
From D to C	Between A and B	$V_H - V_{offset2}$

To address the issue, a bespoke circuitry was designed and constructed on a PCB board, by employing a ‘current spin model’ [20], and integrated with the graphene Hall effect sensors for biasing and processing the output. The circuitry was used to correct the output by eliminating non-desired offset voltages and reducing noise levels thus providing an improvement on the sensitivity of the overall system. The schematic in Fig. 10 demonstrates the steps that were implemented for driving devices and reading the outputs. Fundamentally, the current is driven from one of the contacts to a non-neighboring contact and simultaneously the produced output voltage is measured across the remaining two contacts (e.g. current flows from contact A to B and the voltage difference between C and D is measured (Fig. 9 and TABLE 1). In such a case, the produced output includes both Hall voltage (V_H) generated by applied perpendicular magnetic field (B_y) and offset voltage (V_{off}). To be able to remove the offset voltage, the current flow is rotated 180° and passes through two new non-neighboring contacts whilst the voltage across

the remaining contacts is measured. In this case, the measured output includes Hall voltage subtracted by offset voltage. For a more robust elimination, this procedure continues for one complete cycle meaning that the current is driven between all contacts for a complete cycle following 90° rotation steps. Averaging the obtained outputs eliminates the offset and reveals the Hall voltage.

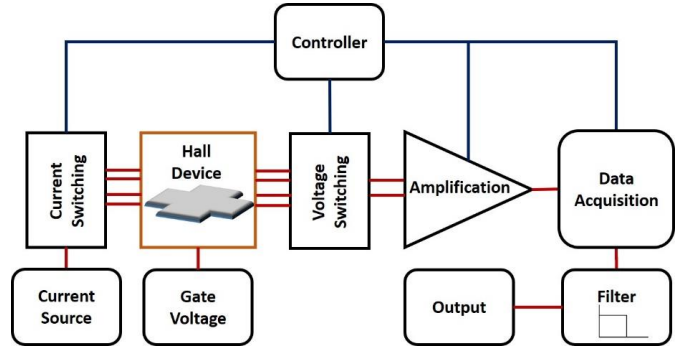


Fig. 10. Schematic of developed driving and processing circuitry.

The constructed PCB (Fig. 11) has an on-board constant current source, which can easily be adjusted from 1 μ A to 10 mA and can bias the sensor in floating or grounded mode of operation. Due to the controllable sensitivity of materials such as graphene, an optional input for gate voltage was also included. Current and voltage switching mechanisms were connected to the current source and the amplifier along with the connections to device contacts. A microcontroller simultaneously manipulates the current rotation and output reading for various frequencies up to 1 MHz. Amplification was performed via a high-performance instrumentation amplifier (LMP8358). The sensors can be integrated by being mounted on the tip of the board as shown in the figure. A rare earth magnet such as neodymium with relevant dimensions can easily be fitted to the tip of the developed portable system given in Fig. 11.



Fig. 11. Constructed system board PCB.

IV. RESULTS AND DISCUSSION

The Hall voltage is not affected from the misalignment of the contacts along the length of the device. Because the electrons would feel the same amount of resistance to their

flow and a voltage value due to the exerted force would occur at the location where the transverse contacts are placed. A typical output obtained from a Hall device without using the developed front-end circuitry is shown in Fig. 12 as an example.

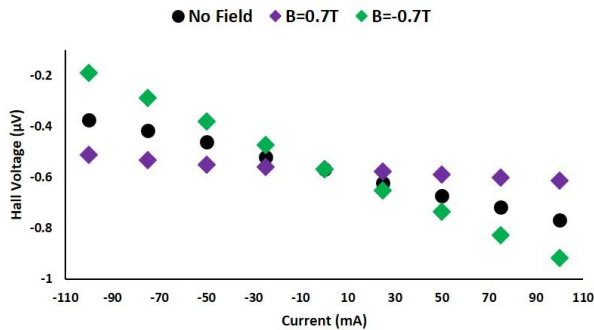


Fig. 12. A typical output obtained from a Hall device without correction.

The zero-magnetic field line should have been on zero-horizontal line instead of being a negative value with a voltage gradient. This was because of the offset voltage effect as explained in previous sections. However, if the no-field line is taken as a reference, the response of the device is highly linear for both positive and negative magnetic field strengths since a voltage gradient is observed for positive and negative values of current and magnetic field. A typical processed output of the device for one cycle (using the developed front-end circuitry) is given in Fig. 13.

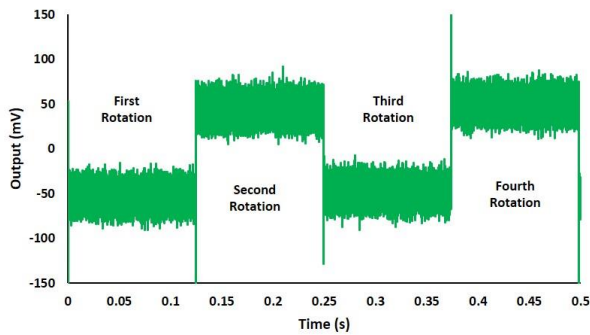


Fig. 13. A Typical output obtained after one cycle.

Each region indicated in the figure corresponds to the specific cases related to current rotation sequences, e.g., first region is the obtained output for the first case given in TABLE 1 and the second region is the obtained output for the second case given in TABLE 1, and this sequence keeps going on. The response of the device shown in Fig. 12 was corrected using the developed circuitry. Fig. 14 presents the corrected version of the results. As seen clearly, the voltage gradient is removed, and the voltage of interest is presented.

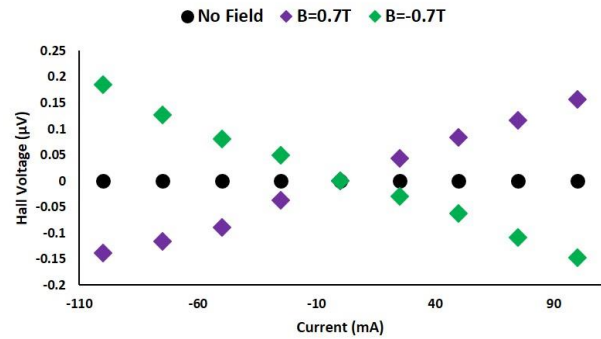


Fig. 14. Corrected Hall output after removing offset voltage.

The developed driving and processing board was tested for its dynamic offset removal ability. The results obtained from the output of the devices with and without circuit implementation was initially assessed for no magnetic field with a variable current source Fig. 15. Direct driving of the Hall sensor under no magnetic field has shown significant reduction in offset voltage. The data shown with purple provides the reduction ratio in percentage.

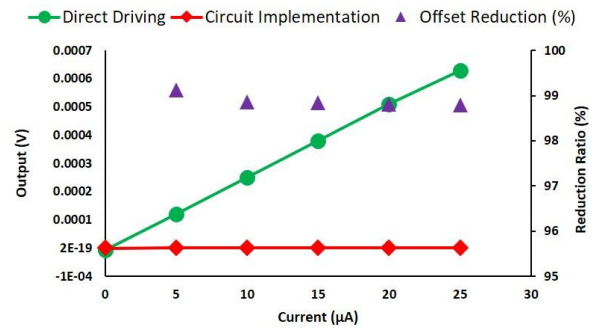


Fig. 15. Demonstration of offset removal.

The circuitry was also tested for constant current flow with variable magnetic field as shown in Fig. 16. The implementation of the circuitry helped reducing the offset voltage by 99%. The negligible residual offset shown in Fig. 17 corresponded to an offset equivalent magnetic field value of 100 nT thus providing a significant improvement in terms of the system accuracy.

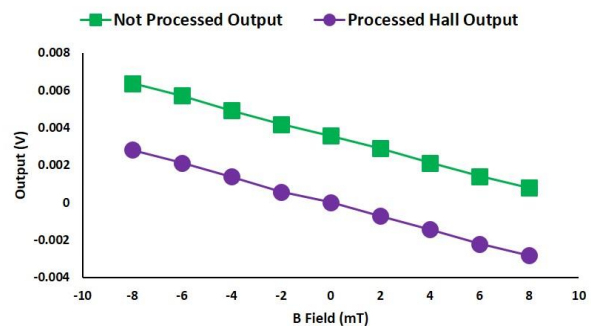


Fig. 16. Processed and not processed outputs for variable magnetic fields.

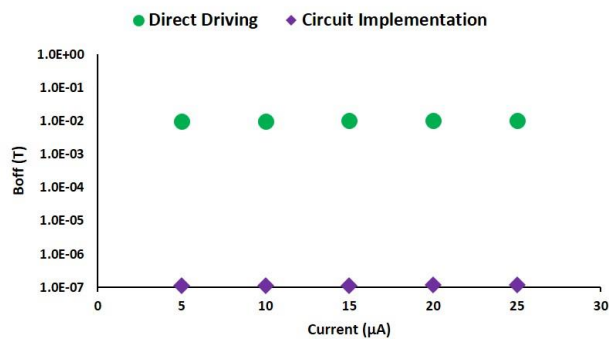


Fig. 17. The comparison of residual magnetic offset values.

Also, fast Fourier transform measurements with and without biasing and driving circuitry showed that the noise level is reduced considerably by employing the developed circuitry. The power spectral density figures showed that the higher frequency operation leads to lower noise effect, thus, providing devices with better sensitivities. Fig. 18 shows power spectral density measurements with respect to frequency of rotation.

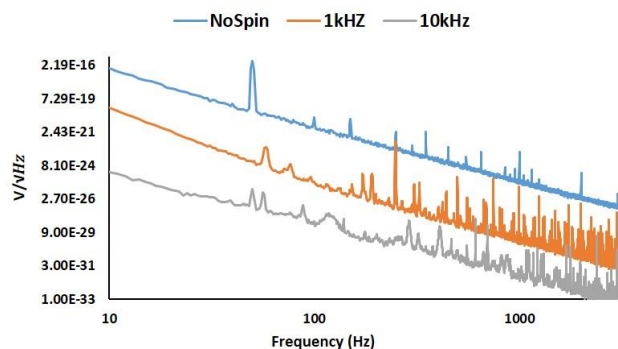


Fig. 18. Power spectral density measurements.

V. CONCLUSION

In this work, an electronic front-end circuitry has been designed and interfaced with Hall biosensors in order to obtain a portable system. The system was proved to have merit for potential point-of-care systems as it does not require to use labels such as beads, thus, eliminates complex electronics setup. In addition, the figures showed that the system has the capacity of correcting undesired effects such as offset voltage and reducing noise, therefore, making the system available for detecting lower field strengths. Being placed on a small printed circuit provides a compact and functional structure which can pave the way for future wearable biosensing systems.

REFERENCES

[1] R. S. Popović, *Hall effect devices*, 2nd ed. Philadelphia: Institute of Physics Pub., 2004.

[2] G. Boero, M. Demierre, P. A. Besse, and R. S. Popovic, "Micro-Hall devices: Performance, technologies and applications," *Sensors Actuators, A Phys.*, vol. 106, no. 1–3, pp. 314–320, 2003.

[3] E. Ramsden, *Hall-Effect Sensors*, 2nd ed. Elsevier, 2006.

[4] H. Xu *et al.*, "Batch-fabricated high-performance graphene Hall elements," *Sci. Rep.*, vol. 3, p. 1207, 2013.

[5] D. Izcı, C. Dale, N. Keegan, and J. Hedley, "The Construction of a Graphene Hall Effect Magnetometer," *IEEE Sens. J.*, vol. 18, no. 23, pp. 9534–9541, Dec. 2018.

[6] S. Johnstone, "Is there potential for use of the Hall effect in analytical science?," *Analyst*, vol. 133, no. 3, pp. 293–296, 2008.

[7] K. Skucha, P. Liu, M. Megens, J. Kim, and B. Boser, "A compact Hall-effect sensor array for the detection and imaging of single magnetic beads in biomedical assays," *16th Int. Solid-State Sensors, Actuators Microsystems Conf. TRANSDUCERS'11*, pp. 1833–1836, 2011.

[8] T. Ishikawa, "Immunoassay on silicon chip," *2014 29th Symp. Microelectron. Technol. Devices Chip Aracaju, SBMicro 2014*, 2014.

[9] P. Manandhar *et al.*, "The detection of specific biomolecular interactions with micro-Hall magnetic sensors," *Nanotechnology*, vol. 20, no. 35, p. 355501, 2009.

[10] A. Sandhu and H. Handa, "Practical hall sensors for biomedical instrumentation," *IEEE Trans. Magn.*, vol. 41, no. 10, pp. 4123–4127, 2005.

[11] K. Togawa *et al.*, "High sensitivity InSb hall effect biosensor platform for DNA detection and biomolecular recognition using functionalized magnetic nanobeads," *Japanese J. Appl. Physics, Part 2 Lett.*, vol. 44, no. 46–49, pp. L1494–L1497, 2005.

[12] D. Issadore, H. J. Chung, J. Chung, G. Budin, R. Weissleder, and H. Lee, "μHall chip for sensitive detection of bacteria," *Adv. Healthc. Mater.*, vol. 2, no. 9, pp. 1224–1228, 2013.

[13] D. Issadore *et al.*, "Magnetic sensing technology for molecular analyses," *Lab Chip*, vol. 14, no. 14, pp. 2385–2397, 2014.

[14] T. Takamura, P. J. Ko, J. Sharma, R. Yukino, S. Ishizawa, and A. Sandhu, "Magnetic-particle-sensing based diagnostic protocols and applications," *Sensors (Switzerland)*, vol. 15, no. 6, pp. 12983–12998, 2015.

[15] T. A. P. Rocha-Santos, "Sensors and biosensors based on magnetic nanoparticles," *TrAC - Trends Anal. Chem.*, vol. 62, pp. 28–36, 2014.

[16] J. Wang, A. N. Kawde, A. Erdem, and M. Salazar, "Magnetic bead-based label-free electrochemical detection of DNA hybridization," *Analyst*, vol. 126, no. 11, pp. 2020–2024, 2001.

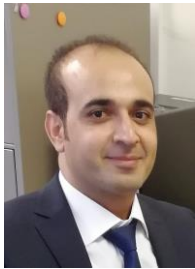
[17] A. Manzin and V. Nabaei, "Modelling of micro-Hall sensors for magnetization imaging," *J. Appl. Phys.*, vol. 115, no. 17, 2014.

[18] M. A. Paun, J. M. Sallese, and M. Kayal, "Hall effect sensors design, integration and behavior analysis," *J. Sens. Actuator Networks*, vol. 2, no. 1, pp. 85–97, 2013.

[19] S. Sanfilippo, "Hall probes: Physics and application to

- magnetometry,” in *CAS 2009 - CERN Accelerator School: Magnets, Proceedings*, 2010, pp. 423–462.
- [20] R. Steiner, C. Maier, A. Häberli, F. P. Steiner, and H. Baltes, “Offset reduction in Hall devices by continuous spinning current method,” *Sensors Actuators, A Phys.*, vol. 66, no. 1–3, pp. 167–172, 1998.
- [21] X. Chen, Y. Xu, X. Xie, Y. Guo, and Y. Huang, “A novel Hall dynamic offset cancellation circuit based on four-phase spinning current technique,” *China Semicond. Technol. Int. Conf. 2015, CSTIC 2015*, pp. 1–3, 2015.

BIOGRAPHIES



DAVUT IZCI He received his BSc degree from Dicle University, Turkey, in Electrical and Electronic Engineering and his MSc and PhD degrees from Newcastle University, England, UK, in Mechatronics and Microsystems, respectively. He is currently a member of academic staff at Batman University, Turkey. His research interests are in microsystems development, sensing applications utilizing graphene nanomaterial and robotics along with instrumentation and control systems.



JOHN HEDLEY He received the B.Sc. degree in astronomy and astrophysics and the Ph.D. degree in atomic physics from Newcastle University, U.K., in 1991 and 1996, respectively. After three post-doctoral researcher positions in MEMS design, fabrication, and testing, he obtained an academic position with the School of Engineering, Newcastle University, where he is currently a Senior Lecturer. His research interests are in microsystems development and sensing applications utilizing graphene.


Evaluation of Visualisation of Scanpath Trend Analysis (ViSTA) Tool

Ş. ERASLAN, S. KARABULUT, M.C. ATALAY and Y. YEŞİLADA


Abstract—Eye tracking plays a key role in user behaviour understanding and usability studies. We previously proposed an algorithm called STA (Scanpath Trend Analysis) that analyses multiple individual scanpaths on a web page to discover their trending path in terms of the areas of interest (AOIs). This algorithm provides the most representative path of multiple users compared to other algorithms (i.e., provides the most similar path to individual scanpaths). However, its current implementation has no graphical user interface and provides a sequence of characters that represent AOIs. Some external modules should also be installed in advance to run it. In our previous work, we presented the first web-based visualisation tool for the STA algorithm called ViSTA along with its initial evaluation. This tool allows to visualise individual scanpaths on a particular web page with gaze plots, visually draw AOIs, apply the STA algorithm, and visualise the result of the algorithm. In this paper, we present the extended version of ViSTA with a follow up user evaluation. The first version of ViSTA uses the STA algorithm which identifies trending AOIs based on all individual scanpaths. However, the extended one uses the STA algorithm with the tolerance level parameter which means trending elements can be identified based on a subset of individual scanpaths for discovering a more representative path. Both of our initial and follow up evaluations show that the workload in terms of NASA Task Load Index (TLX) is lower with ViSTA compared to the current implementation of the STA algorithm.

Index Terms—Eye Tracking, Areas of Interest, Web Pages, Trending Path, Web-based, User Interface.


ŞÜKRÜ ERASLAN, Middle East Technical University, Northern Cyprus Campus, 99738 Kalkanlı, Güzelyurt, Mersin 10, Turkey (e-mail: seraslan@metu.edu.tr).

 <http://orcid.org/0000-0002-9277-8375>


SERKAN KARABULUT, Middle East Technical University, Northern Cyprus Campus, 99738 Kalkanlı, Güzelyurt, Mersin 10, Turkey (e-mail: karabulut.serkan@metu.edu.tr).

 <https://orcid.org/0000-0002-6694-7219>

MEHMET CAN ATALAY, Middle East Technical University, Northern Cyprus Campus, 99738 Kalkanlı, Güzelyurt, Mersin 10, Turkey (e-mail: atalay.can@metu.edu.tr).

 <https://orcid.org/0000-0002-9029-9880>

YELİZ YEŞİLADA, Middle East Technical University, Northern Cyprus Campus, 99738 Kalkanlı, Güzelyurt, Mersin 10, Turkey (e-mail: yyeliz@metu.edu.tr).

 <http://orcid.org/0000-0001-8501-0205>

Manuscript received November 30, 2018; accepted October 24, 2019.
DOI: [10.17694/bajece.490601](https://doi.org/10.17694/bajece.490601)

I. INTRODUCTION

Eye tracking has been commonly used for better understanding of how users interact with web pages, especially which areas are frequently used and which paths are usually followed in terms of these areas to complete certain tasks. Based on the analysis of eye tracking data, web pages can be further processed to support users in constrained environments for completing their tasks. In particular, web pages can be re-engineered for these users to allow them to directly access the most frequently used areas without being distracted by other inappropriate areas. This is especially useful for visually disabled users who access the web with their screen readers and have to listen to the whole page from top to bottom [1].

While users are interacting with web pages, their eyes make quick movements called saccades. Their eyes also make fixations on certain points. The series of saccades and fixations show their scanpaths on these pages. There are different techniques to visualise individual scanpaths [2] (see Section 2). The most popular technique for scanpath visualisation is a gaze plot that represents saccades and fixations with straight lines and circles respectively. The radius of a circle is directly proportional to the duration of a fixation and each circle is numbered to show the sequence. Figure 1 shows an example of a gaze plot on the home page of the Babylon website.

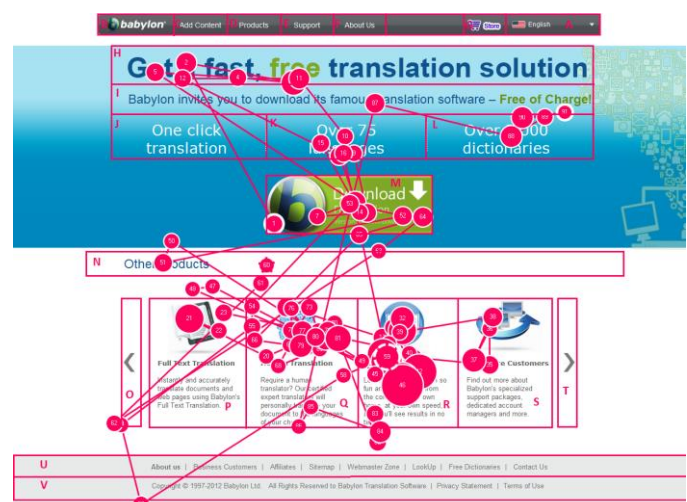


Fig. 1. A gaze plot on the home page of the Babylon website segmented by the extended VIPS algorithm [3]

To conduct scanpath analysis on the web, web pages are typically divided into their areas that can interest or attract users (AOIs) and then fixations are represented with their

corresponding AOIs. For example, Figure 1 shows how the home page of the Babylon website is automatically divided into its AOIs with the extended Vision-based Page Segmentation (VIPS) algorithm [3]. If a user looks at the AOIs M, H and H respectively, then his/her path will be represented as MHH. After representing scanpaths in terms of AOIs, they are further analysed for different purposes, especially calculating a similarity score between two scanpaths, computing transition probabilities between AOIs, detecting patterns in a set of scanpaths, or discovering a single representative scanpath of multiple users [4].

Even though there are different algorithms which aim to detect patterns in multiple scanpaths or discover a representative scanpath for a group of scanpaths, their results typically have low similarities to individual scanpaths [4]. Compared to these algorithms, Scanpath Trend Analysis (STA) discovers the most representative path of multiple users on a web page as its result is the most similar to individual scanpaths [5] (see Section 3). The STA algorithm firstly identifies the AOIs that will be included in the resultant path. The first version of this algorithm identifies these AOIs based on all individual scanpaths. Since high variance in individual scanpaths could negatively affect the representativeness of the resultant path, a new parameter called tolerance level was added to the STA algorithm to allow the identification of the AOIs to be included in the resultant path based on a subset of individual scanpaths instead of all of them. The value of the tolerance level parameter can be adjusted both manually and automatically to achieve the most representative path.

The current implementation of the STA algorithm¹ has no visual user interface and its output is only a sequence of alphanumeric characters that represent AOIs (e.g., ABCD). This algorithm is novel in this field and is being increasingly used. For example, it has been used for identifying common patterns while reading programming code [6]. It has also been used to investigate differences in web interaction of people with autism and neurotypical people [7, 8]. Although the STA algorithm is developed for the analysis of eye tracking data, it has also been adapted to analyse sequenced binary motion sensor data to recognise daily activities (meal preparation, eating, dish washing, etc.) of elderly people who live alone [9]. To understand and interpret the output, researchers and practitioners should remember which alphanumeric character corresponds to which AOI. If a visual user interface is developed for the STA algorithm and its output is also visualised on a corresponding page, then we eliminate the burden of remembering and understanding the sequence of letters. Furthermore, some platform and modules are also needed to be installed in advance to run the current implementation.

To address the limitations of the current implementation of the STA algorithm and make it more usable and functional, we presented the first web-based visualisation tool for this algorithm called Visualisation of Scanpath Trend Analysis or shortly ViSTA along with its initial evaluation in our previous

paper [10]. It uses the STA algorithm without the tolerance level parameter. This tool does not need any specific platform and modules installed. It allows users to upload an eye tracking dataset and select one of the web pages included in the dataset for further analysis. Since gaze plots are commonly used to visualise scanpaths, this tool visualises the individual scanpaths on the selected page by using gaze plots [2]. It then allows users to visually draw AOIs on the page and apply the STA algorithm for discovering the trending path based on the drawn AOIs. The trending path is then visualised by using an AOI graph. Since an AOI graph shows both AOIs and transitions between these AOIs, it would be easier to recognise which AOIs are trending among users and in which order these AOIs are visited. We evaluated the ViSTA tool by comparing its workload with the workload of the current implementation of the STA algorithm. Our evaluation showed that the STA algorithm can be used with less workload in terms of NASA Task Load Index (TLX) with the ViSTA tool.

In this paper, we present the extended version of the ViSTA tool (See Section 4) along with its follow up evaluation (See Section 5). The extended version of the tool uses the STA algorithm with the tolerance level parameter. In the follow up study, the same methodology of the previous study was used with different users. The results of the follow up study were consistent with the results of the first study which means that the ViSTA tool decreases the workload of using the STA algorithm.

The ViSTA tool can be used by eye tracking researchers, usability evaluators, data scientists and psychologists. Although it is mainly designed for visualising the STA algorithm, it has an open-architecture such that other similar algorithms can also be integrated into this tool. Similarly, other visualisation techniques can also be integrated so that the outputs of the algorithms can be visualised in different ways. We conclude our paper with a discussion of such possible future work (see Sections 6 and 7).

II. RELATED WORK

A large number of techniques have been proposed in the literature to visualise eye tracking data on 2D and 3D visual stimuli in context or not in context [2]. Heat maps and gaze plots have been commonly used in eye tracking research for visualisation. Heat maps use spatial information of eye tracking data, specifically x, y and possibly z coordinates of fixations. These maps use different colours to differentiate which areas are commonly used and which areas are rarely used by users. For example, Tobii Studio Software [11] generates heat maps which show the commonly used areas with the red colour and the rarely used areas with the green colour. These maps can be generated based on different features, such as fixation counts or fixation durations. Although the most commonly used areas can easily be recognised with heat maps, these maps do not take sequential information into consideration. Specifically, heat maps cannot be used to determine in which order these areas are used. Even though gaze plots use both spatial and sequential information

¹ STA's current implementation: <https://github.com/SukruEraslan/sta>

as illustrated in Figure 1, they will overlap each other when there are many scanpaths. Therefore, the visualisation of many scanpaths with gaze plots at the same time would be a problem because it would be difficult to analyse them. There are also some visualisation techniques which are designed to visualise eye tracking data based on given AOIs [2, 12]. In particular, each AOI can be used as a node (represented with circles) in a graph and directed edges (represented with lines) between these nodes are used to illustrate transitions between AOIs [2]. The sizes of nodes can be determined based on different features, such as fixation count or dwell time [2]. The thickness of edges can also be determined based on different features, such as the number of transitions [13, 2]. This visualisation technique is called an AOI graph [2].

Even though some of these visualisation techniques have a web-based interface, for example [14, 15], most of them do not have a web-based interface which would allow researchers and practitioners to directly access and use them in their studies without installing extra platform and modules on their computers. There are also different algorithms available to process and analyse eye tracking data, especially scanpaths [4]. However, the results of these algorithms are typically not visualised with any visualisation technique on the web platform, with some exceptions [16, 17, 15]. For example, the ScanGraph tool² finds similarities between individual scanpaths based on one of three methods (Levenshtein, Needleman-Wunsch and Damerau-Levenshtein) and produces a graph where similar scanpaths are connected to each other [16]. Another example is from [15] which generates an AOI graph based on transitions between AOIs where a ticker edge show more transitions.

To the best of our knowledge, none of the algorithms which discover a representative scanpath for multiple scanpaths has a tool to visualise their outputs on the web. In our previous paper, we presented a web-based visualisation tool for the STA algorithm which is able to identify the most representative path for multiple users, and then compared the workload of this tool with the current implementation of the STA algorithm with a user study [10]. In this paper, we present the extended version of this tool by integrating the tolerance level parameter of the STA algorithm and also present a follow-up user study with more users for workload comparison between the ViSTA tool and the current implementation of the STA algorithm. The STA algorithm is briefly explained in the following section where the full description can be found in [5, 18].

III. STA: SCANPATH TREND ANALYSIS

The STA algorithm is a multi-pass algorithm with three main stages which are responsible for preparing individual scanpaths, identifying trending AOIs and constructing the trending path with the trending AOIs respectively [5, 18].

A. Preliminary Stage

The first stage represents each fixation with its

corresponding AOI and each individual scanpath is represented a series of AOIs, such as P [250 ms] Q [120 ms] R [300 ms] where P, Q and R are the AOIs fixated by a particular users for 250 ms, 120 ms and 300 ms respectively.

Eye trackers may experience some inaccuracy in storing the exact positions of fixations, and therefore their degree of accuracy are typically provided. Since this situation can cause some problems in identifying the corresponding AOI for each fixation, the STA algorithm takes the degree of accuracy into account and calculates an error offset by using the following formula (Equation 1). It then uses this error offset to extend the AOIs horizontally and vertically. If a fixation falls into more than one AOI because of the extensions, its corresponding AOI will be the closest one to the fixation.

$$E = \tan(x) \cdot D \cdot P \quad (1)$$

where:

E: Error offset

x: Degree of accuracy

D: Distance between a participant and a screen

P: Pixels per inch for the screen which is calculated based on the screen resolution and the screen size.

B. First Pass

The second stage identifies trending AOIs to be used for the construction of the trending path. An AOI can be visited more than once by a particular user consecutively (such as, PQQR) and/or non-consecutively (such as, PQRQ). Each non-consecutive visit to a particular AOI is referred to as an instance. For example, PQRQP contains three instances of P, two instances of Q and one instance of R. These instances are numbered to be differentiated based on the dwell time (i.e., total fixation duration on each instance) where the first number is given to the longest instance of a particular AOI (such as, P1 Q2 R1 R1 P2 Q1 P3).

The first version of the STA algorithm considers all individual scanpaths for identifying trending AOI instances. Specifically, if a particular AOI instance gets at least the same attention as the fully shared AOI instances in terms of the total fixation duration and the total number of fixations, it is considered as trending AOI. However, the second version of the STA algorithm does not have to identify trending instances based on fully shared instances. In the second version of the STA algorithm, the following conditions should be satisfied by an instance to be trending.

- The total number of fixations on the instance is greater than or equal to the minimum total number of fixations on the instances which are shared by a subset (determined based on the tolerance level parameter) of the individual scanpaths.
- The dwell time on the instance is greater than or equal to the minimum dwell time on the instances which are shared by a subset (determined based on the tolerance level parameter) of the individual scanpaths.

² ScanGraph: www.eyetracking.upol.cz/scangraph

The value of the tolerance level parameter can be between zero and one. When the tolerance level is set to 0.5, the algorithm identifies trending instances based on the instances which are shared by 50% of individual scanpaths. The algorithm does not only allow to adjust this parameter manually, but also adjusts it automatically to identify the most representative trending path (i.e, the most similar path to individual scanpaths).

The instances which are not trending among users are removed from the individual scanpaths as they will not be used for constructing the trending path.

C. Second Pass

The last stage constructs the trending scanpath by combining the trending AOI instances based on their overall positions in the individual scanpaths. The same instances in each scanpath are firstly merged (such as, P1 [250 ms] P1 [150 ms] Q1 [200 ms] R1 [270 ms] → P1 [400 ms] Q1 [200 ms] R1 [270 ms]), and the sequential priority value is computed for each instance with the following formula (Equation 2):

$$\psi_i = 1 - P_i \cdot z, \text{ where } z = \frac{\max_i - \min_i}{L - 1} \quad (2)$$

where:

ψ_i : Priority value of i^{th} instance

P_i : Position of the instance in the scanpath (start: 0)

L : Length of the scanpath

\max_i : Maximum priority value (default: 1)

\min_i : Minimum priority value (default: 0.1)

Once the individual priority values of the instances are computed in each of the scanpaths, the total priority value (Ψ) for each instance is computed with the following formula (Equation 3):

$$\Psi = \sum_{i=1}^n \psi_i \quad (3)$$

where:

Ψ : Total priority value

n : The number of scanpaths

The trending path is then constructed by combining the trending instances and sorting them according to their total priority values in descending order. The dwell time and the total number of fixations on the instances are also considered in case of the same total priority value. In the end, the instance numbers are removed (such as, P1 → P) and the consecutive repetitions are deleted (such as, PQQR → PQR) such that the trending path is represented in terms of the AOIs.

IV. VISTA: VISUALISATION OF SCANPATH TREND ANALYSIS

The ViSTA tool³ can be used to upload an eye tracking dataset, visualise individual scanpaths with gaze plots, visually draw AOIs on stimuli, apply the STA algorithm to discover the trending scanpath and then visualise the trending path with an AOI graph. As illustrated in Figure 2, the ViSTA tool is comprised of three main modules explained below which are as follows: (1) Data Processing, (2) Algorithmic Processing and (3) Visualisation Processing. This tool was mainly developed by using jQuery which is a JavaScript library. The current implementation of the STA algorithm was also slightly modified by using Flask⁴ to run it as a server which accepts POST requests from the ViSTA tool.

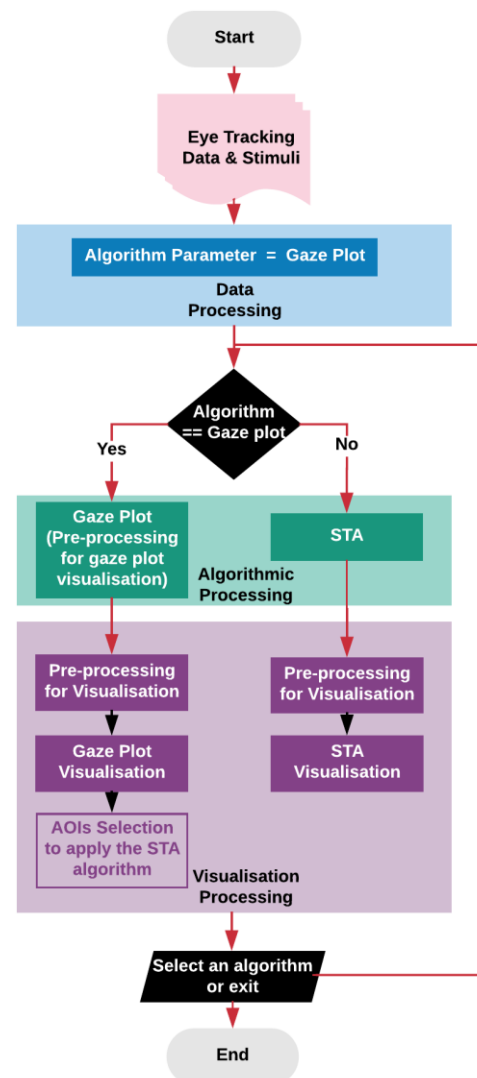


Fig.2. The work flow of the ViSTA tool [Selection of AOIs is needed when the STA algorithm will be applied.]

A. Data Processing

The first module is responsible from taking eye tracking data and converting this data into an internal storage format for further analysis. There should be a different data file for

³ ViSTA: <http://iam.ncc.metu.edu.tr/vistatool/demo/index.html>

⁴ Micro web framework for Python: <http://flask.pocoo.org/>

each participant that contains the details of their fixations (index, timestamp, duration, x and y coordinates and stimuli name). These data files should include web page links as stimuli name since the ViSTA tool is currently designed for scanpath analysis on the web. This tool accesses each page by using their links to retrieve some information, especially their titles. When the data files are uploaded and they are converted into an internal storage format, this module becomes ready to send the data to the next module for further processing.

In this tool, two algorithms are currently available: (i) the gaze plot algorithm which further processes the data to visualise individual scanpaths with gaze plots; and (ii) the STA algorithm which further processes the data to identify the trending scanpath. Since the ViSTA tool firstly visualises individual scanpaths with gaze plots, the algorithm parameter is set to the gaze plot algorithm internally here so that individual scanpaths are visualised with gaze plots in the first round. The ViSTA tool also allows setting the parameters of the STA algorithm as shown in Figure 3 (see the parameters of the STA algorithm in Section 3). Specifically, when the “Highest Fidelity” option is chosen, the STA algorithm automatically adjusts the tolerance level to identify the most representative path.

Fig.3. The settings page of the ViSTA tool to set the parameters of the STA algorithm

B. Algorithmic Processing

When the Algorithmic Processing module takes the data, it further processes the data based on the given algorithm parameter. If the algorithm parameter is set to the gaze plot algorithm, this module further processes the data to be able to visualise individual scanpaths with gaze plots. However, if the algorithm parameter is set to the STA algorithm, this module further process the data to discover the trending path. Therefore, the algorithm processing modules either prepares the data to be visualised with gaze plots or applies the STA algorithm and prepares the result to be visualised with an AOI graph based on the algorithm parameter.

C. Visualisation Processing

The Visualisation Processing module takes the results of the algorithms and visualises them with their visualisation

techniques by using vis.js⁵. In particular, when the results of the gaze plot algorithm are received, researchers and practitioners should firstly select which participants' scanpaths on which web pages will be visualised with gaze plots. By default, all the participants are selected. When the participants and the web page are selected, this tool takes a screen shot of the page with html2canvas⁶ by using its link and the individual scanpaths are visualised on the page with gaze plots. In case of any problem with taking a screen shot of the page, the tool allows to upload an image file for the page as shown in Figure 4 (the red button). This is also an improvement to the first version of the ViSTA tool that was presented in [10]. Figure 5 shows how the ViSTA tool visualises individual scanpaths on the home page of the Babylon website with gaze plots. Researchers and practitioners can then visually draw AOIs to be used with the STA algorithm. If no AOI is drawn, the STA algorithm cannot be selected for further processing the data because this algorithm discovers the trending scanpath in terms of the AOIs. As an example, Figure 5 also shows three AOIs identified by a user on the home page of the Babylon website.



Fig.4. An option to upload an image file for a web page

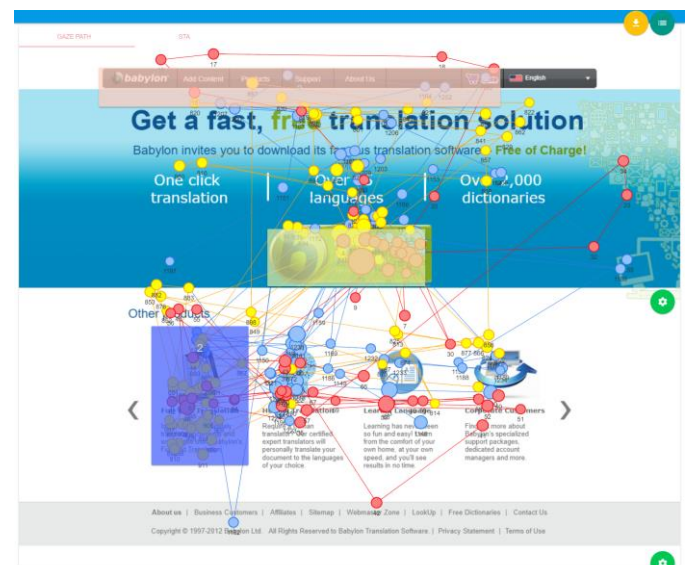


Fig.5. Identification of AOIs on the home page of the Babylon website with gaze plots

When the STA algorithm is selected after the visualisation of individual scanpaths with gaze plots and the result of the STA algorithm is received, this module firstly pre-processes

⁵ Dynamic, browser based visualisation library: <http://visjs.org/>

⁶ Screenshots with JavaScript: <https://html2canvas.hertzen.com/>

the result for visualisation, and then visualises it by using an AOI graph. An example of this visualisation is shown in Figure 6. Since the AOI that includes the link to download the free version of Babylon is the most frequent AOI in the trending path, its node is larger than other two AOIs' nodes.

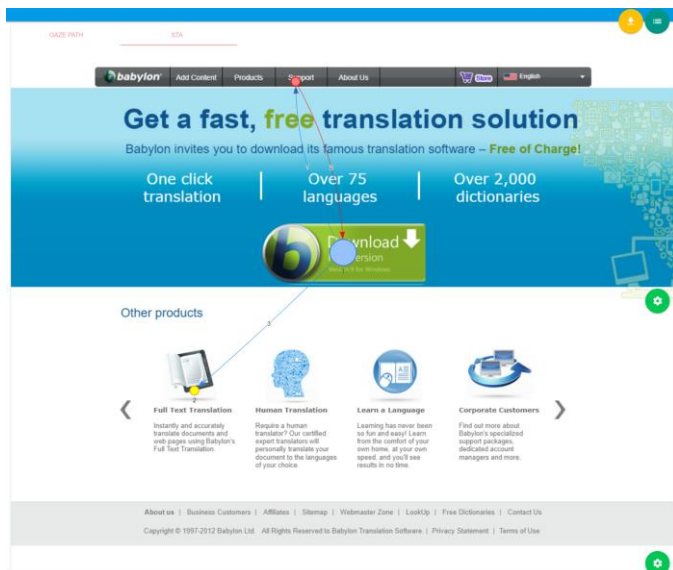


Fig.6. Visualisation of the trending scanpath discover by the STA algorithm on the home page of the Babylon website

V. EVALUATION

In order to assess the effectiveness of the ViSTA tool, two user studies were conducted by using the NASA Task Load Index (TLX). User studies are a reliable way to understand how products are perceived by users [19] and NASA TLX has widely been used to perform workload assessments for different kinds of human-machine systems [20]. This metric is composed of six attributes: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration.

In this study, our main research question was “Does the ViSTA tool have better TLX attributes in comparison with the current implementation of the STA algorithm?”. We investigated this research question with the following questions associated with each of the TLX attributes:

- Mental Demand: “How mentally demanding was the task?”
- Physical Demand: “How physically demanding was the task?”
- Temporal Demand: “How hurried or rushed was the pace of the tasks?”
- Performance: “How successful were you in accomplishing what you were asked to do?”
- Effort: “How hard did you have to work to accomplish your level of performance?”
- Frustration: “How insecure, discouraged, irritated, stressed, and annoyed were you?”

Each attribute is assessed by the participants by giving a value between one and 20. Higher values represent negative scores. For example, for mental demand one means very low

and 20 means very high mental load. This is also the case for the performance attribute - a higher score means the users did not think they were successful in accomplishing the tasks they were asked to do. In the given scale, one means perfect performance and 20 means failure in terms of performance.

We conducted the first study with 12 participants from the Middle East Technical University Northern Cyprus Campus, and presented the results in [10]. To extend the evaluation of the ViSTA tool, we conducted a follow-up study by using the same methodology but with different 12 participants from various universities in Turkey. We present the results of the follow-up study below along with the comparison with the results of the first study. We also combine the results from the first and second studies to present an overall evaluation of the tool. We were able to do this because both studies followed the same methodology.

A. Methodology

A detailed description of the dataset used in our user studies for the evaluation of the ViSTA tool can be found in [5]. In both of these user studies, we used a subset of that dataset. First of all, we randomly divided the dataset into two halves called dataset A and dataset B. We then asked our participants to apply the current implementation of the STA algorithm with one half of the dataset and the ViSTA tool with the other half to identify their trending paths on a particular web page. A half of the participants used the current implementation of the algorithm and then the ViSTA tool whereas another half used the ViSTA tool and then the current implementation of the algorithm. Besides this, a half of the participants applied the ViSTA tool with the dataset A and the current implementation of the algorithm with the dataset B whereas another half applied the ViSTA tool with the dataset B and the current implementation of the algorithm with the dataset A. Therefore, we ensured that all the cases were counterbalanced to deal with any possible familiarity and order effects. When the participants identified the trending scanpath for a particular dataset, they were given a printed copy of the page to draw the trending path on the page and then they were asked to fill in the English version of the NASA TLX Form⁷.

We also applied to the dependent T-Test or its non-parametric alternative Wilcoxon signed-rank test to statistically compare the workload of the ViSTA tool and the current implementation of the STA algorithm based on NASA TLX attributes. Since we expected to have lower workload with the ViSTA tool, we applied these tests as one-sided. The statistical comparison was not applied in our previous paper [10], so it is also a new analysis for the first user study.

1) Participants

The first study was conducted at Middle East Technical University Northern Cyprus Campus with four female and eight male students. The participants were between the age of 18-24, apart from one of them who was between the age of 25-

⁷

<https://humansystems.arc.nasa.gov/groups/TLX/downloads/TLXScale.pdf>

34. All of the participants were daily web users. They were also asked to assess their computer skills by using a 5-point Likert scale. The mean value was 3.42 with the standard deviation 1.00.

The second study was conducted with three female and nine male students from different universities in Turkey. Nine of these participants were between the age of 18-24 whereas three of them were between the age of 25-35. All of them were also daily web users. They ranked their computer skills as 3.17 out of 5 (standard deviation: 0.72)

2) Procedure

The participants firstly read the information sheet to understand the main objectives of the study and their rights and then they signed a consent form. After that, they were asked to complete a short questionnaire to collect their basic demographic information (gender, age groups, web usage, general computer skills, and departments). A training session was then given to them to illustrate how they can use the current implementation of the STA algorithm and the ViSTA tool. This training session mainly used a sample dataset (not the one used in the evaluation). After the training session, they started their evaluation session.

3) Materials

The original eye tracking dataset includes six web pages with varying level of visual complexities. We used the home page of the Babylon website for the training sessions, and the home page of the Apple website for the evaluation sessions.

B. RESULTS

This section presents the results from the first and the second studies respectively and then combines the results of these studies to provide an overall evaluation of the ViSTA tool.

1) Results of Evaluation Study I

Figure 7 and Table I shows the results from the first study. Figure 7 shows the comparison of the current implementation of the STA algorithm and the ViSTA tool based on the mean values of the NASA TLX attributes where the error bars illustrate the standard deviation. Besides this, Table I shows a detailed descriptive analysis of the NASA TLX attributes for the current implementation of the STA algorithm and the ViSTA tool.

The results of the first study shows that the ViSTA tool needs less mental demand, physical demand and temporal demand in comparison with the current implementation of the STA algorithm. The participants also stated that they needed less effort to complete the given task with the ViSTA tool. Moreover, they were less frustrated when they used the ViSTA tool. However, they stated that they performed slightly better (better satisfied with their performance in achieving the task) with the current implementation of the STA algorithm.

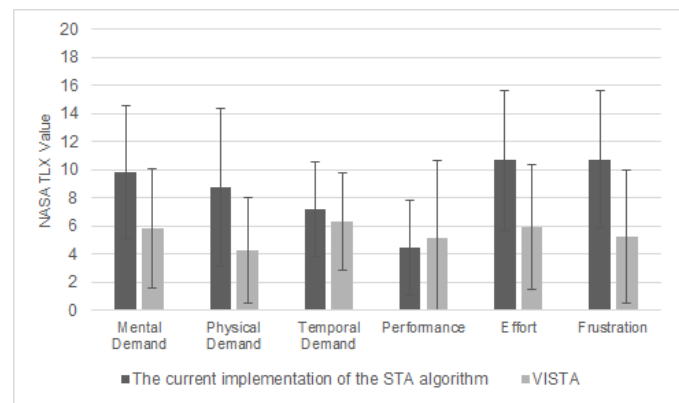


Fig.7. NASA TLX attributes for the current implementation of the STA algorithm and the ViSTA tool from the first study

TABLE I
THE MEAN (M), MEDIAN (MD) & STANDARD DEVIATION (SD) OF THE NASA TLX ATTRIBUTES IN THE FIRST STUDY

NASA TLX Attributes	The current imp. of the STA alg.			ViSTA		
	M	MD	SD	M	MD	SD
Mental Demand	9.8	9.5	4.7	5.8	4.5	4.2
Physical Demand	8.8	10.0	5.6	4.3	3.0	3.8
Temporal Demand	7.2	8.0	3.4	6.3	5.5	3.5
Performance	4.5	3.5	3.4	5.2	2.5	5.5
Effort	10.7	12.5	5.0	5.9	4.5	4.4
Frustration	10.8	10.5	4.9	5.3	3.5	4.7

The statistical comparison shows that the ViSTA tool has a significantly lower workload in terms of mental demand ($t(11)=-2.65$, $p<0.05$, $d=0.8$), physical demand ($t(11)=-2.8$, $p<0.05$, $d=0.84$), effort ($t(11)=-3.06$, $p<0.05$, $d=0.92$) and frustration level ($t(11)=-3.45$, $p<0.05$, $d=1.04$) with large effect size compared to the current implementation of the STA algorithm. Although the temporal demand is also lower with the ViSTA tool, it is not statistically different ($t(11)=-1.22$, $p=0.12$, $d=0.37$). Besides, there is no significant difference between the ViSTA tool and the current implementation of the STA algorithm in terms of their performance ($t(11)=0.55$, $p=0.3$, $d=0.16$).

2) Results of Evaluation Study 2

Figure 8 and Table II illustrates the results from the second study. Figure 8 illustrates the comparison of the current implementation of the STA algorithm and the ViSTA tool in based on the mean values of the NASA TLX attributes where the error bars illustrate the standard deviation. Moreover, Table II illustrates a detailed descriptive analysis of the NASA TLX attributes for the current implementation of the STA algorithm and the ViSTA tool.

Similar to the results of the first study, the results of the second study also illustrates that the participants completed their tasks with less mental demand, physical demand, temporal demand, effort and frustration with the ViSTA tool in comparison with the current implementation of the STA

algorithm. However, in contrast to the first study’s participants, the second study’s participants stated that they performed slightly better with the ViSTA tool instead of the current implementation of the STA algorithm.

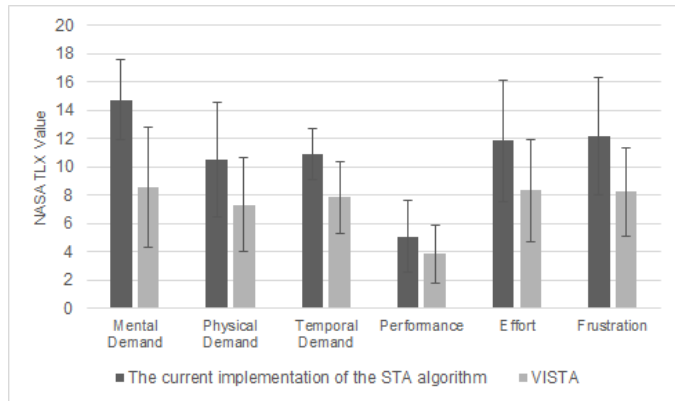


Fig.8. NASA TLX attributes for the current implementation of the STA algorithm and the ViSTA tool from the second study

TABLE II

THE MEAN (M), MEDIAN (MD) & STANDARD DEVIATION (SD) OF THE NASA TLX ATTRIBUTES IN THE SECOND STUDY

NASA TLX Attributes	The current imp. of the STA alg.			ViSTA		
	M	MD	SD	M	MD	SD
Mental Demand	14.8	16.0	2.8	8.6	7.5	4.2
Physical Demand	10.5	12.0	4.0	7.3	7.5	3.3
Temporal Demand	10.9	11.0	1.8	7.8	7.0	2.6
Performance	5.1	5.0	2.6	3.8	3.0	2.0
Effort	11.8	14.0	4.3	8.3	9.0	3.6
Frustration	12.2	13.0	4.2	8.3	6.5	3.1

The statistical comparison shows that the ViSTA tool has significantly better values for mental demand ($t(11)=-4.21$, $p<0.05$, $d=1.27$), physical demand ($t(11)=-2.7$, $p<0.05$, $d=0.81$), temporal demand ($t(11)=-3.99$, $p<0.05$, $d=1.2$), effort ($t(11)=-4.31$, $p<0.05$, $d=1.3$) with large effect size and frustration ($z= 2.09$, $p<0.05$, $r=0.43$) with medium effect size in comparison with the current implementation of the STA algorithm. Although the ViSTA tool has a better performance value at this time, it is not statistically different ($t(11)=-1.26$, $p=0.12$, $d=0.38$).

3) Overall Results

Figure 9 and Table III shows the overall results by combining the results of both the first and second studies. The overall results suggest that the ViSTA tool scores significantly better in mental demand ($t(23)=-4.83$, $p<0.05$, $d=1.01$), physical demand ($t(23)=-3.9$, $p<0.05$, $d=0.81$), effort ($t(23)=-4.76$, $p<0.05$, $d=0.99$) and frustration ($t(23)=-4.5$, $p<0.05$, $d=0.94$) with large effect size and temporal demand ($t(23)=-3.52$, $p<0.05$, $d=0.73$) with medium effect size in comparison with the current implementation of the STA algorithm. Even though the ViSTA tool compared to the current implementation of the STA algorithm scores slightly better in

performance in the second study, its performance score is slightly lower in the first study. However, when we combine the results of both of the studies to provide an overall evaluation, the ViSTA tool performance was perceived better (i.e., the overall performance score was lower for ViSTA which means better performance), even though the difference is not statistically significant ($t(23)=-0.37$, $p=0.36$, $d=0.08$).

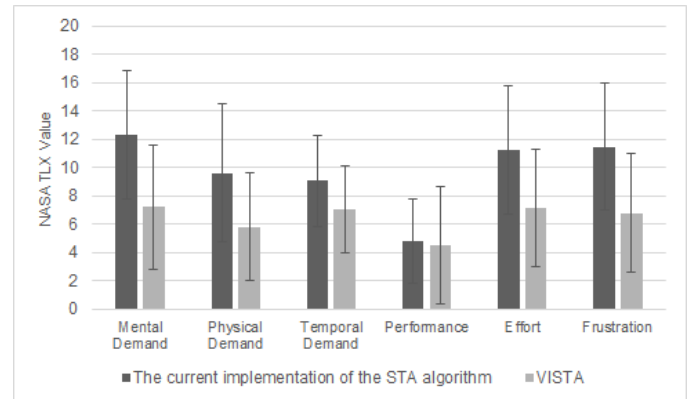


Fig.9. Overall NASA TLX attributes for the current implementation of the STA algorithm and the ViSTA tool from both the first and second studies

TABLE II

THE OVERALL MEAN (M), MEDIAN (MD) & STANDARD DEVIATION (SD) OF THE NASA TLX ATTRIBUTES

NASA TLX Attributes	The current imp. of the STA alg.			ViSTA		
	M	MD	SD	M	MD	SD
Mental Demand	12.3	13.0	4.5	7.2	6.0	4.4
Physical Demand	9.6	11.0	4.8	5.8	5.0	3.8
Temporal Demand	9.0	10.0	3.3	7.1	6.5	3.1
Performance	4.8	4.0	2.9	4.5	3.0	4.1
Effort	11.3	13.0	4.6	7.1	7.5	4.1
Frustration	11.5	12.0	4.5	6.8	6.0	4.2

VI. DISCUSSION

The ViSTA tool provides a web-based interface for the STA algorithm. It allows researchers and practitioners to directly access and use the STA algorithm for their studies without installing extra platform and modules on their computers. Similar to other algorithms, the STA algorithm can also be used by different researchers in different fields. Therefore, it should not be expected that all of these researchers easily manage to download and run their implementations on their computers. With the ViSTA tool, researchers and practitioners can easily upload their eye tracking datasets and visually draw their AOIs to apply the STA algorithm. When the trending path is generated with the algorithm, it is visualised with an AOI graph so that researchers can easily interpret the trending path without the need for remembering which alphanumeric character represents which AOI as they will see the trending AOIs and the order in which the AOIs are used. Our overall evaluation results show that the ViSTA tool decreases the workload of using the STA algorithm.

The current version of the ViSTA does not include automatic AOI detection based on fixation clusters or the source code of web pages. The tool allows researchers to visually draw their own AOIs on the fly. They may detect AOIs by using different techniques and then they can draw the AOIs in the ViSTA tool. As the tool has an open-architecture, it can be easily extended in the future to support automatic AOI detection. For example, the VIPS algorithm has been extended and implemented as a web service [3], so we can use this web service to automatically segment web pages into their areas based on their source code and visual representation. Appropriate clustering algorithms can also be implemented as part of this tool to allow the identification of AOIs by clustering fixations [21]. Another visualisation level can also be added to the ViSTA tool for AOIs which will be shared by other components of the tool.

The ViSTA tool currently has only the STA algorithm, apart from the generic algorithm to pre-process the data for visualising individual scanpaths with gaze plots. However, it has an open architecture which means that other algorithms can easily be integrated. For example, the SPAM algorithm has been used to identify sequential patterns in eye tracking data [22]. It can also be integrated into this tool and its output can also be visualised with an AOI graph. However, some studies need to be conducted. Similarly, if there are other online web-based eye tracking tools, the ViSTA tool can also easily be integrated to those tools. It is a web application so it can easily be added to other web-based applications.

We are still improving this tool by exploring appropriate algorithms and visualisation techniques to make this tool more functional and usable. In the future, we are planning to add different visualisation techniques to the Visualisation Module (see Figure 2), such as a time plot [23]. As a consequence, when a particular algorithm provides an output, a list of appropriate visualisation techniques can be listed to be selected for visualising the output. The visualisation of the outputs of the algorithms can also be animated to make them more informative. For example, when the output of the STA algorithm is visualised with an AOI graph, its nodes and edges can be shown one at a time based on the time sequence which allows researchers to focus on a particular trending AOI at a time during analysis. The ViSTA tool can also be improved in a way that a particular algorithm can be applied to different groups of participants (at the moment, the STA algorithm is only applied to one group) and the results of these groups can be visualised on the same visual stimulus for a comparison purpose. Additionally, an account system can be created such that the users of the tool can have their workbenches. This would allow them to store their datasets and revisit their workbenches.

Our work presented here is not without limitations. We evaluated the ViSTA tool with 24 people in total who were mostly not very experienced in eye tracking research and studies. This situation could have an impact on the scores given by our participants. However, we believe if we had eye-tracking specialists, our results would be even better for the

ViSTA tool. However, to confirm this, in the future, we are planning to conduct further studies with more, and more experienced participants for detailed feedback.

VII. CONCLUSION

The aim of this paper is to introduce the extended version of the first web-based interface of the STA algorithm and also present its extended user evaluation where the participants were asked to assess the workload of the VISTA tool and the current implementation of the STA algorithm for the comparison purpose. The STA algorithm is novel in this field and it is being increasingly used in different studies. Its visualisation, called ViSTA, will allow researchers and practitioners to use the STA algorithm with a lower workload and understand its output directly without the need of remembering the names of AOIs and relating the AOIs to the output in their minds. Since there is a limited number of web-based tools for analysing eye tracking data, our tool also makes an useful contribution to eye tracking research. This tool is open to be further improved by integrating other algorithms and visualisation techniques.

REFERENCES

- [1] Y. Yesilada, S. Harper and S. Eraslan, "Experiential Transcoding: An EyeTracking Approach," in *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility*, New York, NY, USA, 2013.
- [2] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf and T. Ertl, "Visualization of Eye Tracking Data: A Taxonomy and Survey," *Computer Graphics Forum*, vol. 36, pp. 260-284, 2017.
- [3] M. E. Akpınar and Y. Yeşilada, "Vision Based Page Segmentation Algorithm: Extended and Perceived Success," in *Current Trends in Web Engineering: ICWE 2013 International Workshops ComposableWeb, QWE, MDWE, DMSSW, EMotions, CSE, SSN, and PhD Symposium, Aalborg, Denmark, July 8-12, 2013. Revised Selected Papers*, Q. Z. Sheng and J. Kjeldskov, Eds., Cham, Springer International Publishing, 2013, pp. 238-252.
- [4] S. Eraslan, Y. Yesilada and S. Harper, "Eye tracking scanpath analysis techniques on web pages: A survey, evaluation and comparison," *Journal of Eye Movement Research*, vol. 9, 2015.
- [5] S. Eraslan, Y. Yesilada and S. Harper, "Scanpath Trend Analysis on Web Pages: Clustering Eye Tracking Scanpaths," *ACM Trans. Web*, vol. 10, pp. 20:1--20:35, 11 2016.
- [6] C. Tablatin and M. M. Rodrigo, "Identifying Common Code Reading Patterns using Scanpath Trend Analysis with a Tolerance," in *Proceedings of the 26th International Conference for Computers in Education (ICCE 2018)*, Metro Manila, Philippines, 2018.

- [7] S. Eraslan, V. Yaneva, Y. Yesilada and S. Harper, "Do Web Users with Autism Experience Barriers When Searching for Information Within Web Pages?," in *Proceedings of the 14th Web for All Conference on The Future of Accessible Work*, New York, NY, USA, 2017.
- [8] S. Eraslan, V. Yaneva, Y. Yesilada and S. Harper, "Web users with autism: eye tracking evidence for differences," *Behaviour & Information Technology*, vol. 38, pp. 678-700, 2019.
- [9] H. Y. Yatbaz, S. Eraslan, Y. Yesilada and E. Ever, "Activity Recognition Using Binary Sensors for Elderly People Living Alone: Scanpath Trend Analysis Approach," *IEEE Sensors Journal*, 2019.
- [10] Ş. Eraslan, S. Karabulut, M. C. Atalay and Y. Yeşilada, "ViSTA: Visualisation of Scanpath Trend Analysis (STA) / Scanpath Trend Analysis (STA)'in Görselleştirilmesi," in *Proceedings of the 12th Turkish National Symposium on Software Engineering (12. Ulusal Yazılım Mühendisliği Sempozyumu, UYMS 2018)*, İstanbul, Turkey, 2018.
- [11] Tobii Technology AB, "Tobii Studio™ 2.X User Manual (Sep. 2010)," 2010.
- [12] M. Burch, A. Kull and D. Weiskopf, "AOI rivers for visualizing dynamic eye gaze frequencies," in *Computer Graphics Forum*, 2013.
- [13] K. Holmqvist, J. Holsanova, M. Barthelson and D. Lundqvist, "Reading or scanning? A study of newspaper and net paper reading.," J. R. In Hyönä and H. Deubel, Eds., Elsevier, 2003, pp. 657-670.
- [14] L. Herman, S. Popelka and V. Hejlova, "Eye-tracking Analysis of Interactive 3D Geovisualization," *Journal of Eye Movement Research*, vol. 10, 2017.
- [15] M. Burch, A. Kumar and N. Timmermans, "An Interactive Web-based Visual Analytics Tool for Detecting Strategic Eye Movement Patterns," in *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, New York, NY, USA, 2019.
- [16] J. Dolezalova and S. Popelka, "ScanGraph: A Novel Scanpath Comparison Method Using Visualisation of Graph Cliques," *Journal of Eye Movement Research*, vol. 9, 2016.
- [17] G. Topić, A. Yamaya, A. Aizawa and P. Martínez-Gómez, "FixFix: Fixing the Fixations," in *Proceedings of the Ninth Biennial ACM Symposium on ETRA*, New York, NY, USA, 2016.
- [18] S. Eraslan, Y. Yesilada and S. Harper, "Engineering web-based interactive systems: trend analysis in eye tracking scanpaths with a tolerance," in *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems, EICS 2017*, New York, NY, USA, 2017.
- [19] S. Eraslan and C. Bailey, "End-User Evaluations," in *Web Accessibility: A Foundation for Research*, Y. Yesilada and S. Harper, Eds., London, : Springer London, 2019, pp. 185-210.
- [20] A. Cao, K. K. Chintamani, A. K. Pandya and R. D. Ellis, "NASA TLX: Software for assessing subjective mental workload," *Behavior Research Methods*, vol. 41, pp. 113-117, 01 2 2009.
- [21] A. Santella and D. DeCarlo, "Robust Clustering of Eye Movement Recordings for Quantification of Visual Interest," in *Proceedings of the 2004 Symposium on ETRA*, New York, NY, USA, 2004.
- [22] P. Hejmady and N. H. Narayanan, "Visual Attention Patterns During Program Debugging with an IDE," in *Proceedings of the 2012 Symposium on ETRA*, New York, NY, USA, 2012.
- [23] K.-J. Rähä, A. Aula, P. Majaranta, H. Rantala and K. Koivunen, "Static Visualization of Temporal Eye-Tracking Data," in *Human-Computer Interaction - INTERACT 2005: IFIP TC13 International Conference, Rome, Italy, September 12-16, 2005. Proceedings*, M. F. Costabile and F. Paternò, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 946-949.

BIOGRAPHIES



Sükrü Eraslan is currently working as a research associate at the University of Manchester. He completed his BSc in Information Systems Engineering at Cyprus International University in 2009. After that, he received his MSc in Information Technology (Data Engineering) at the University of Wolverhampton in 2011. He then completed his Ph.D. in Computer Science at the University of Manchester in 2016. After his Ph.D., he worked as a post-doctoral researcher at Middle East Technical University Northern Cyprus Campus. His main research area is eye tracking on the Web. He is interested in analysing eye movements of users to understand how they interact with web pages to make the web more accessible in constrained environments. Besides, he has started to work on learning analytics & software engineering.



Serkan Karabulut received his B.Sc. in Computer Engineering from Middle East Technical University Northern Cyprus Campus (METU NCC) in 2012. He is currently working in AYESAŞ as a Software Developer. His research interests include eye tracking and human-computer interaction.



Mehmet Can Atalay received his B.Sc. in Computer Engineering from Middle East Technical University Northern Cyprus Campus (METU NCC) in 2012. He is currently doing his M.Sc. in Cognitive Science at Middle East Technical University (METU) and working as a Software Engineer in Turkey. His research interests include deep learning, neuro-linguistic programming and eye tracking.



Yeliz Yeşilada is an Associate Professor at Middle East Technical University Northern Cyprus Campus (METU NCC) and an honorary research fellow in the School of Computer Science at the University of Manchester. She received her B.Sc. in Computer Engineering from the Eastern Mediterranean University in 1999, her M.Sc. in Computer Science from the University of Manchester in 2000 and also completed her Ph.D. in 2005 in Computer Science at the University of Manchester. Her primary research interest is centred around Human Computer Interaction and User Experience; in particular the mobile Web, Web accessibility and eye tracking research to improve user experience. Further information about Dr. Yeşilada can be found at: <http://www.yelizyesilada.info>.

Three Part Hybrid Encryption Schema

H. GENÇOĞLU, T. YERLİKAYA

Abstract—Hybrid encryption schemes are consist of two parts. First part for encrypting the text with symmetric algorithm called DEM (Data Encryption Mechanism) and the second one is for encrypting the symmetric algorithm key with asymmetric algorithm called KEM (Key Encryption Mechanism). If we think to expand the KEM packet with e-sign, message digest and extra security data for validation and authentication, how will be the encryption of KEM packet and also the performance KEM mechanism. Our study try to answer that question and also offer a new hybrid encryption mechanism for information security.

Index Terms—Cryptography, Hybrid Cryptography, Data Security

I. INTRODUCTION

IT IS NEVER safe to do business or personal correspondence on the Internet unless it is possible to protect the information. Information security is provided by eliminating threats such as being listened by another one, change of information, imitation of identity. The basic tool used for securing information security is cryptography. Cryptography is a science that analyzes information security and makes the understandable one incomprehensible. Topics such as confidentiality, reliability, data integrity, authentication, authenticity, and irrefutability are important areas of work for cryptography. [1-4]

Encryption techniques can be used to prevent unauthorized access to data content during data security. By using symmetric or asymmetric algorithms, the encrypted data will be meaningless even if it is captured by unauthorized persons. However, by using asymmetric algorithms and processing the data with the keys of the sender, it is possible to obtain unique values. If these values are related to the data and the sender, the sender of the data will be guaranteed if it is added to the data, which is a digital signing process.

Nowadays, communication security is not only to secure the data content. It is now important for communication security to associate the sender with the data. It should also be proved

that the received data is received without any modification on the recipient's side.

Complex security solutions are generally provided by hybrid encryption mechanisms. Symmetric algorithms only have the ability to encrypt data. They are fast and safe. Key management of asymmetric algorithms is successful. Symmetric algorithms are much slower when compared to symmetric algorithms in encryption processes. Hybrid mechanisms are built to complement the missing aspects of symmetric and asymmetric algorithms.

II. ENCRYPTION ALGORITHMS

A. Asymmetric Encryption Algorithms

Asymmetric algorithms use two different keys, one for encryption and the other for decryption. These keys mathematically depend on each other, but one cannot get a key from the other. The decryption key is known only by the recipient, but the encryption key can be known by anyone. The fact that the keys are interlinked and one cannot get a key from the other makes it possible to use asymmetric algorithms rather than encryption. Today, asymmetric algorithms are used for digital signatures. [5]

B. Symmetric Encryption Algorithms

Symmetric algorithms use a single key for encryption and decryption. This key must be kept confidential by both the sender and the recipient, otherwise unauthorized persons can easily decrypt the encrypted data and access the actual data. They have a widespread use, because they perform encryption and decryption fast.[6]


C. Hash Algorithms

Hash algorithms are used to create a strict representation of the given data. Calculated hash value is unique. The smallest change in the actual data substantially changes the hash value. Since they are one-way functions, the actual data cannot be obtained. Because of these features, Hash algorithms play a big role in data integrity processes and in authentication. [14 - 15]

III. PREVIOUS WORKS

One of the important works on this subject was made by Masayuki Abe, Rosario Gennaro, Kaoru Kurosawa and Victor Shoup. In this study, the hybrid mechanism is composed of two modules. It is called DEM Data Encapsulation Mechanism for the Hybrid Mechanism's encryption of the data itself, and KEM-Key Encapsulation Mechanism for the mechanism managing encryption keys. In principle, the data is

HAKAN GENÇOĞLU, is with Department of Computer Engineering İstanbul Sabahattin Zaim University, İstanbul, Turkey, (e-mail: hakan.gencoglu@izu.edu.tr).

 <https://orcid.org/0000-0003-2968-1615>

TARIK YERLİKAYA, is with Department of Computer Engineering Trakya University, Edirne, Turkey, (e-mail: tarikyer@trakya.edu.tr).

Manuscript received September 07, 2019; accepted October 08, 2019.

DOI: [10.17694/bajece.616893](https://doi.org/10.17694/bajece.616893)

encrypted with DEM and the keys are encrypted with KEM. [7]

In the tag-KEM / DEM architecture, when a KEM is generated, first a random key is selected and then this key is encrypted with a tag value. Signing procedures are not available in this architecture.

Fujisaki-Okamoto's KEM-DEM architecture is an improved version of the "Tag-KEM / DEM" architecture. But this architecture does not have signing procedures either.

Another study was conducted by Kerim YILDIRIM and H. Engin DEMİRAY. The previous hybrid mechanisms have been improved in their study called "SİMETRİK VE ASİMETRİK ŞİFRELEME YÖNTEMLERİNE METOTLAR: ÇIRPILMIŞ VE BİRLEŞİK AKM-VKM". [8]

Three algorithms are mentioned in this study. In the "Scrabbled KEM-DEM" architecture, KEM and DEM constructions were mixed using the receiver's open key and sent to the user in a single locale. It is stated that the purpose of this mixing process is to make sure that the attacker does not know if it works with KEM or DEM. In the cascaded KEM-DEM architecture, which is the other algorithm, the shuffle algorithm was considered to be safer to do with the random key instead of the receiver's open key. However, in both cases, it will be seen that there is no difference between the DEM keys if they are thought to be encrypted and stored in the KEM. In the "Combined KEM-DEM" structure, the server was first logged on and secure communication was established with the "Scrabbled KEM-DEM" architecture. The logon key encryption specified by the server in the logon process was performed as if it were in the other two systems.

Although this work enhances security measures, signing and verification procedures are incomplete. Plain or encrypted text is not signed in the system, only the key that generated the key which is used to encrypt the text is signed. However, the signing of the text instead of the key will be more meaningful. In addition, first a key is generated in the system with a string expression, then the encryption key is generated using this key, and encryption is performed. This process will increase the work load on the system and increase the working time of the system. If it is desired that the encryption key is independent of the sender, the system can generate the key without any string entry and encryption can be performed.

In both cases, data security is provided by encrypting the plain text and obtaining a summary to guarantee that it has not changed. Signaling by the sender to verify the sender of the message is also an important security parameter and it should be used for authentication purposes if the attack techniques are considered to be diverse today.

Another study, "A Novel Idea on Multimedia Encryption using Hybrid Crypto Approach", describes that multimedia files can be used safely without leaving the hybrid algorithm architecture.[9]

In the study named "Dik eşleştirme arayış yöntemi ile hibrit veri sıkıştırma ve optiksel kriptografi ", the processed signal is secured with the hybrid mechanism and transmission is proposed.[10]

In the study "A Password-Protected Secret Sharing Based on Kurosawa-Desmedt Hybrid Encryption", the hybrid mechanism has been used to eliminate defects in the secret sharing schemes. [11]

In another study, implementation solution was presented for hybrid mechanism cloud storage systems. RSA is used as the asymmetric algorithm and AES is used as the symmetric algorithm in the work named "The hybrid encryption algorithm of lightweight data in cloud storage".[12]

An idea is proposed for the hybrid architecture to work on the FPGA in the study called "Implementing a hybrid cryptocoding algorithm for an image on FPGA".[13]

As it is seen, the vast majority of studies on hybrid encryption algorithms focus on the implementation of the architecture or performance enhancement rather than the development of the architecture.

IV. ENCRYPTED COMMUNICATION

The encrypted communication between two points is made using the keys of the point. There will be 4 keys at each point. These keys are; symmetric algorithm key for encryption of the data, public and private keys of asymmetric algorithm for signing, validation, and security packet encryption, and public key belonging to the receiver.

Create 3 packages during the communication. The first package, the message package, contains encrypted data. The second package, the security package, contains signing and authentication values. The third package is the key package that contains the keys of encrypting message package and the security package.

While there are two parts in classical hybrid cryptography architectures, we have created the architecture in three parts in our study. We designed the security package as a separate package and increased the security parameters to make the security of the message stronger. Because increasing the security parameters increases the packet size, we encrypted the security package with a symmetric algorithm. We have completed the hybrid architecture by encrypting two symmetric keys using asymmetric algorithm.

A. Package –Message (Data) Package

The plain data goes in two processes:

- Encryption process
- Sign the message by the sender

The plain data is encrypted using the symmetric algorithm with the randomly generated key. The encrypted message goes to shuffle algorithm and produces two outputs:

1. Shuffled encrypted message
2. Coefficients after the shuffling operation

The purpose of re-mixing the encrypted message is to make the prediction more difficult against the known text attack against the symmetric algorithms. To make this process meaningful, the coefficients must be sent to the receiver and the mapping must be reversed to determine the original locations of the data blocks. One of the inputs to the security package is these coefficients.

The digest value is calculated by entering the plain data in the digest algorithm. This value is signed using the sender's secret key, and a message signature is generated which means that the message is approved by the user. The message signature is one of the inputs of the security package.

- $\delta_{K1}(m) = y$
- $Sh(y) = (z, q)$
- $q \rightarrow \delta_{K2}$
- $z \rightarrow Message\ Package$
- $\delta_{K1} = K1$ Symmetric encryption algorithm using K1 Key
- $\delta_{K2} = K2$ Symmetric encryption algorithm using K2 Key
- $Sh =$ Shuffle Algorithm
- $K1 =$ Encryption Key (Symmetric Algorithm)
- $z =$ Shuffled Encrypted Message
- $q =$ Coefficient Array
- $y =$ Encrypted Message (Message Package)

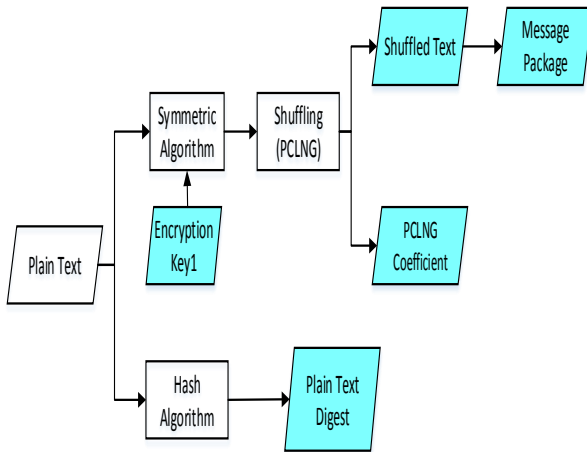


Fig.1. Message Encryption and Message Packet Creation Flowchart

B. Package2 - Security Package

The message signature is one of the inputs of the security package.

The digest of the plain message and the digest of the scrambled message are created in a data packet, first by signing with the device secret key and then by signing with the user secret key. Coefficients of the data packet from shuffle algorithm are also added to the signed data and encrypted with a symmetric algorithm.

A randomly generated key is used for encryption, then this key goes to the package.

- $H(m) = h_m$
- $H(z) = h_z$
- $SG_{SKC}(h_m) = I_{mC}$
- $SG_{SKK}(I_{mC}) = I_{mK}$
- $SG_{SKC}(h_z) = I_{zC}$
- $SG_{SKK}(I_{zC}) = I_{zK}$
- $(I_{mC}, I_{mK}) = I_m$
- $I_m \rightarrow \delta_{K2}$
- $q \rightarrow \delta_{K2}$
- $\delta_{K2}(I_m, q) = \tau$

$\delta_{K2} = K2$ Symmetric encryption algorithm, using K2 Key, pre-security package

$SG_{SK} = SK$ Signature Algorithm using SK Key (Device-Person)

$H =$ Hash Algorithm

$SKC =$ Sender Device's Secret Key (Asymmetric Algorithm)

$SKK =$ Sender's Secret Key (Asymmetric Algorithm)

$I_{mC} =$ Signature of the message created by the device

$I_{mK} =$ Signing a device signature for a message created by a person

$I_{zC} =$ Signature of the scrambled encrypted message created by the device

$I_{zK} =$ Signing the device signature of the scrambled encrypted message created by the person

$K2 =$ Security Packet Encryption Key (Symmetric Algorithm)

$m =$ Plain Text

$h_m =$ Message Digest

$h_z =$ Summary of scrambled encrypted text

$q =$ Coefficient Array

$\tau =$ Security Package

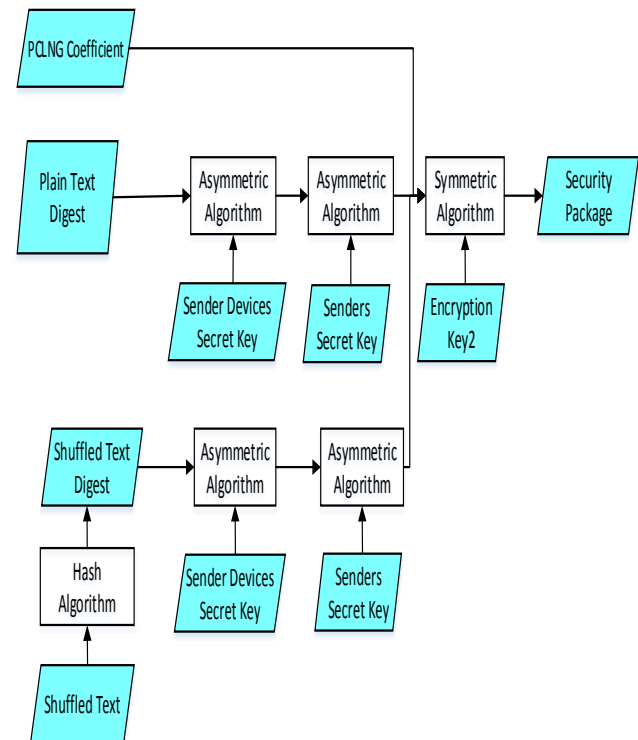


Fig.2. Signing and encryption of security package entries by device and person and creation of security package

C. Package3 - Key Package

Asymmetric algorithm is used in the last packet. Encryption keys for the first two packages in the encrypted state are moved. Encryption keys of symmetric algorithms used in message package and security package are encrypted by Asymmetric Algorithm to create Packet3.

$\gamma_{PK}(K1, K2) = Ap$
 γ = Asymmetric encryption algorithm
 K1 = Message Encryption Key (Symmetric Algorithm)
 K2 = Security Packet Encryption Key (Symmetric Algorithm)
 PK = Receiver's Public Key (Asymmetric Algorithm)

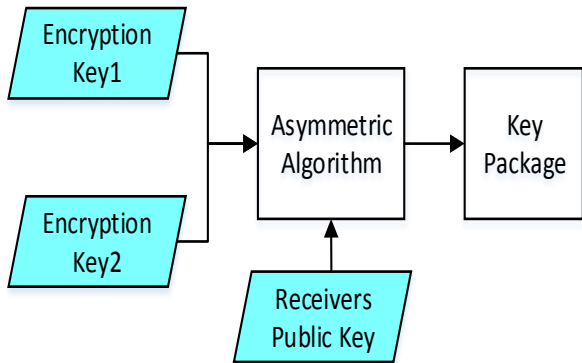


Fig.3. Key Package Creation Process

$Vp (Mp, Gp, Ap)$
 Ap = Key Package
 Gp = Security Package
 Mp = Message Package

The data packet is formed by combining the message package security package and the password package.



Fig.4. Combination of Password Pack and Packages

After receiving the recipient message, the keys needed to display the message content are encrypted in the key package. In this case, the key package must only be opened by the receiver. For this, the key package must be encrypted with the receiver's public key. Public keys must be mutually shared for the encryption to take place.

V.DECRYPTION PROCESS

When the data package arrives at the recipient, it is sufficient for the open message to reverse the order of the transactions in order to be able to access the signature information. The same key in the symmetric algorithm and the second key in the asymmetric algorithm must be used.

The packages are separated from each other in the very beginning. Each package follows its own sequence of operations within the architecture. First the security package is opened, and the shuffle coefficients and signatures are obtained, then signatures are checked to obtain hash values. Finally, the Message Pack is opened and a plain text is obtained. The shuffle algorithms for the plain text are compared with the hash values obtained from the security packet. If the values obtained in this comparison are the same, it is guaranteed that the message is not changed.



Fig.5. Combination of Password Pack and Packages

A. Step 1. Unpacking the Key Package

The key package is encrypted with the recipient's public key. The keys of the symmetric encryption algorithm used for message encryption and security packet encryption are obtained by unpacking the receiver's secret key.

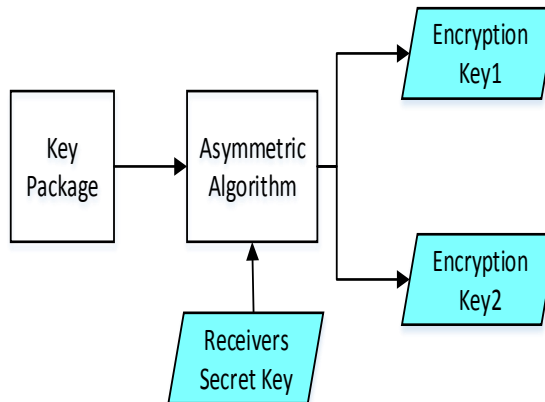


Fig.6. Unpacking the Key Package and Obtaining the Encryption Keys

$\gamma_{SKK}(Ap) = (K1, K2)$
 γ = Asymmetric encryption algorithm
 K1 = Message Encryption Key (Symmetric Algorithm)
 K2 = Security Packet Encryption Key (Symmetric Algorithm)

SKK = User's Secret Key (Asymmetric Algorithm)

Encryption Key2 will be used to decrypt security package, Encryption Key1 will be used to obtain the clear message.

B. Step 2: Unpacking Security Package

The security packet obtained from the data packet is decrypted using the Encryption Key2 obtained by decrypting the key packet. The obtained data are a hash of the message signed by the user and the device and the coefficients used in the process of shuffling the encrypted message. This hash value will be compared with a hash value of the message obtained in the next step, and the device and the person to whom the message is sent will be verified.

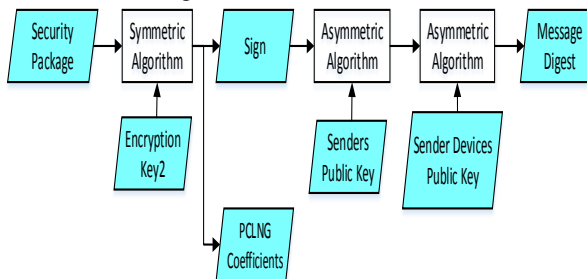


Fig.7. Unpacking of Security Package and Retrieving Hash Value and AKDU Coefficients

$$\begin{aligned} \delta_{K2}(\tau) &= (I_K, q) \\ SG_{SKK}(I_K) &= I_C \\ SG_{SKC}(I_C) &= (h_m, h_z) \\ H(m) &= h_m \\ H(z) &= h_z \\ I_K &\rightarrow \delta_{K2} \\ q &\rightarrow \delta_{K2} \end{aligned}$$

δ_{K2} = K2 Symmetric encryption algorithm using K2 Key
 SG_{SK} = SK Signature Algorithm using Key K (Device-Person)

H = Hash Algorithm
 I_C = Signature of the message created by the device
 I_K = Signing a device signature created by a person
 K2 = Security Packet Encryption Key (Symmetric Algorithm)
 m = Plain Text
 h_m = Message Digest
 h_z = Hash of scrambled encrypted text
 q = Coefficient Array
 τ = Security Package

The signature is calculated by processing the digest value of the message data first with the device's secret key then sender's secret key by the asymmetric algorithm. The signature verification process is also performed by asymmetric algorithm using first the device's secret key then the sender's secret key. As a result of this process, a hash value of the signature data should be obtained.

C. Step 3: Unpacking Message Package

The string that was passed in the shuffling process after the message package is encrypted. Therefore, the cryptographic message must first be obtained by reversing the shuffling process, and the resulting cryptographic message must be decrypted with the Encryption Key 1 coming from the key packet with the symmetric algorithm. The coefficients that will be processed to recover the shuffling process come from the security package.

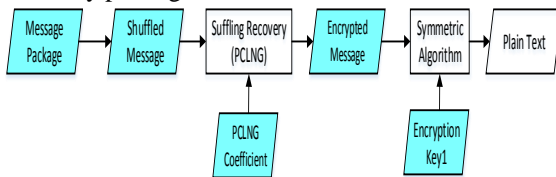


Fig.8. Unpacking Message Pack and Obtaining Open Message

$$\begin{aligned} Sh(z)_q &= y \\ \delta_{K1}(y) &= m \end{aligned}$$

δ_{K1} = K1 Symmetric encryption algorithm using Key K1
 Sh = Shuffle Algorithm
 K1 = Message Encryption Key (Symmetric Algorithm)
 z = Shuffled Encrypted Text
 q = Coefficient Array
 y = Encrypted Text (Message Packet)

D. Step 4: Verification

At the end of the second phase, the hash value was obtained. This summary value consisted of a message digest mixed with a plain message digest. Because it is not possible to obtain the original data from the summary value, the plain text is again processed to the shuffling and hashing operations to obtain the hash value which is verified by comparing the summary value with the data packet.

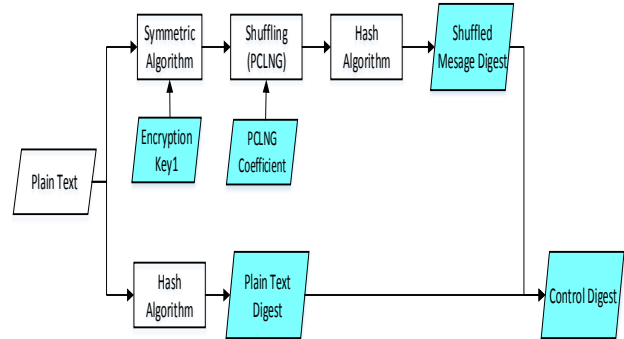


Fig.9. Confirmation of Value Compared with Security Package and Open Messaging Values

The control hash value is compared with the summary from the data packet. If the results are correct, the device and sender are verified.

VI. PRIME COEFFICIENT LINEAR NUMBER GENERATOR- PCLNG

A. Shuffling Process:

As $A[n]=\{ a_1, a_2, a_3, \dots, a_n \}$ and $1 \leq x \leq n$ are, algorithm is like this:

Select p,q primes as $2 \leq p \leq n$ and $2 \leq q \leq n$.

1. Define Sh(x) function
 $Sh:[1,n] \rightarrow [1,n]$
 $Sh(x)=px+q \text{ mod } n$

2. $(T[x] = A[Sh(x)])_{x=1}^n$ place all values of array A to temporary array T using mapping function.
3. $(A[x] = T[x])_{x=1}^n$ transfer the temporary T array values to array A.

Using this algorithm we have shuffled array A.

Main question is if $Sh:[1,n] \rightarrow [1,n]$ $Sh(x)=px+q \text{ mod } n$ function is $1 \rightarrow 1$? .Is $Sh(x_1) = Sh(x_2)$, for $x_1 \neq x_2$.

Proof:

Select two elements as $x_1 \neq x_2$ in different places.

We assume that $px_1+q=px_2+q \text{ mod } n$ $px_1+q=px_2+q$
 + kn

$$px_1 = px_2 + kn$$

$$px_1 - px_2 = kn$$

$$p(x_1 - x_2) = kn \text{ is found.}$$

- i. If $k=0$, because of $p \neq 0$, $x_1 - x_2 = 0 \Rightarrow x_1 = x_2$ is found, and it is against our approval
- ii. If $x_1 - x_2 = \frac{nk}{p}$, because of $\gcd(p,n)=1$, $p \nmid nk$ is found. That means p cannot divide n . We assume that p divide k ($p|k$): $\frac{k}{p} = k'$. At this time $x_1 - x_2 = nk' \Rightarrow x_1 = nk' + x_2$ is found. That means $x_1 > n$ and that situation is against $x_1 \leq n$.

Therefore $Sh(x) = px+q \text{ mod } n$ function is $1 \rightarrow 1$ function. And for all $x_1 \neq x_2$, becomes $Sh(x_1) \neq Sh(x_2)$.

B. Recovery Process:

To recover the original array after shuffling process, we must use reverse of the $Sh(x)=px+q$ function.

Reverse of $Sh(x)=px+q \text{ mod } n$ function is $Sh' = \frac{x-q}{p}$. Because of this function's being linear the reverse of p and q to $\text{mod } n$ can be used instead of them. Lets call p' and q' to the reverse of p and q . If $p' = \frac{1}{p} \text{ mod } n$ and $q' = -q \text{ mod } n$ are calculated $Sh'(x)$ becomes, $Sh'(x) = (x + q')p' \text{ mod } n$.

The place of each element of the array is calculated in $Sh(x) = px + q \text{ mod } n$ function, and their new place is found. To recover the original array we must calculate $Sh'(x) = (x + q')p' \text{ mod } n$ for all places in the shuffled array and find their original places.

VII. CONCLUSION

Encryption and decryption times at different data sizes are measured by developing an application for performance evaluation.

RSA is used as the asymmetric algorithm in signing operations and ciphers, and AES-256 is chosen as the symmetric algorithm. The P and Q prime numbers required for the RSA algorithm are selected to be 402 digits - 1334 bits and 401 digits - 1331 bits.

The numbers E and D selected using P and Q are 802 digits - 2662 bits and 801 digits - 2658 bits.

The values measured by the application are given in the table.

TABLE 1
THE VALUES MEASURED BY THE APPLICATION

Data (Byte)	Encryption (Sec)	Decryption (Sec)
8	2,9021564	1,6582876
16	2,3145948	1,8892842
24	2,2512891	1,7815471
32	2,2989085	1,7368394
40	2,3057441	1,7242867
80	2,3866955	1,7844664
160	2,3197742	1,6806094
320	2,3422487	1,6720422
640	2,2734561	1,6412479
1280	2,361179	1,6358062
4000	2,2584238	1,5706789
12000	2,4123994	1,7610501
24000	2,6840385	1,8357856
48000	3,7703456	2,1218189
72000	5,5372788	2,7813307
96000	8,3044108	3,7025864
144000	16,4070262	7,4168919

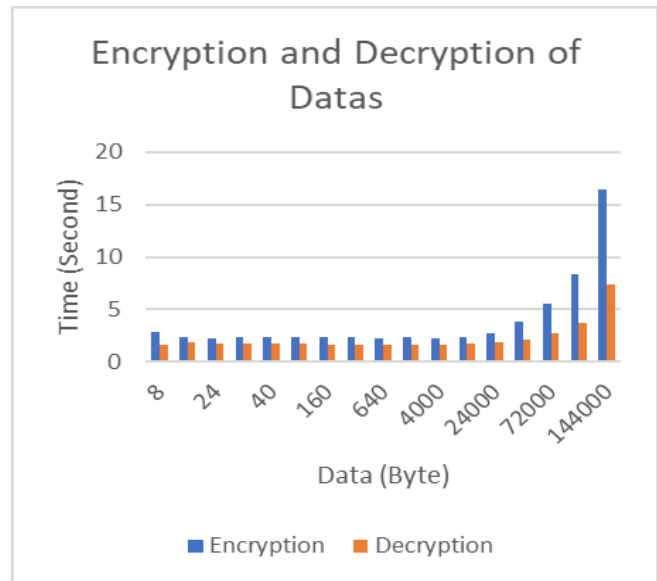


Fig.10. Encryption and Decryption Time of Datas

Certificate authorities do not prefer 1024 bit keys in signing and strategic applications anymore. Nowadays, 2048 or 4096 bit keys have begun to be preferred for longer security choice

Absolute security is ensured in the transmission of thumbnails and textual data as well as problems during transmission of audio and video data. Our next work will be on ensuring faster delivery of larger data.

REFERENCES

- [1] H. Kodaz "Veri İletiminde Güvenlik İçin Şifreleme", Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 2002
- [2] T. Yerlikaya, E. Buluş, N. Buluş, "Asimetrik Şifreleme Algoritmalarında Anahtar Değişim Sistemleri", Akademik Bilişim 2006 (Ab2006), 9-11 Şubat 2006, Denizli
- [3] M. Krishnamurthy, E.S. Seagren, R. Alder, A.W. Bayles, J. Burke, S. Carter, E. Faskha, "Basics of Cryptography and Encryption, How to Cheat at Securing Linux", 2008, 249- 270.
- [4] T. Stapko, "Security Protocols and Algorithms, Practical Embedded Security", 2008, 49-66.
- [5] H. Kodaz, F. M. Botsali "Simetrik Ve Asimetrik Şifreleme Algoritmalarının Karşılaştırılması", Selçuk-Teknik Dergisi ISSN 1302-6178 Journal of Selçuk-Technic Cilt 9, Sayı:1-2010 Volume 9, Number:1-2010 10
- [6] T. Yerlikaya, E. Buluş, N. Buluş "Asimetrik Şifreleme Algoritmalarında Anahtar Değişim Sistemleri" AKADEMİK BİLİŞİM 2006 + BilgiTek IV 9-11 Şubat 2006 Pamukkale Üniversitesi Denizli
- [7] M. Abe, R. Gennaro, K. Kurosawa, V. Shoup, "Tag-KEM/DEM: A New Framework For Hybrid Encryption And New Analysis Of Kurosawa-Desmedt KEM", Advances in Cryptology – Eurocrypt 2005, Lncs 3494, Pp. 128–146, 2005.
- [8] K. Yıldırım, H. E. Demiray, "Simetrik Ve Asimetrik Şifreleme Yöntemlerine Metotlar: Çırpılmış Ve Birleşik Akm-Vkm", Gazi Üniv. Müh. MİM. Fak. Der. Cilt 23, No 3, 539-548, 2008
- [9] S. C. Iyer, R.R. Sedamkar, S. Gupta, "A Novel Idea On Multimedia Encryption Using Hybrid Crypto Approach", 7th International Conference On Communication, Computing And Virtualization 2016, Procedia Computer Science 79 (2016) 293 – 298.
- [10] E. Atar, O. K. Ersoy, L. Özyılmaz, "Dik Eşleştirme Arayış Yöntemi İle Hibrit Veri Sıkıştırma Ve Optiksel Kriptografi", Journal Of The Faculty Of Engineering And Architecture Of Gazi University 32:1 (2017) 139-147
- [11] T. Arai, S. Obana, "A Password-Protected Secret Sharing Based On Kurosawa-Desmedt Hybrid Encryption", Fourth International Symposium On Computing And Networking (CANDAR) CANDAR Computing And Networking (CANDAR), 2016 Fourth International Symposium On. :597-603 Nov, 2016
- [12] L. Chengliang, Y. Ning, R. Malekian, W. Ruchuan, "The Hybrid Encryption Algorithm Of Lightweight Data in Cloud Storage", 2nd International Symposium On Agent, Multi-Agent Systems And Robotics (ISAMSR) Agent, Multi-Agent Systems And Robotics (ISAMSR), 2016 2nd International Symposium On. :160-166 Aug, 2016
- [13] B.V. Srividya, S. Akhila, "Implementing A Hybrid Crypto-Coding Algorithm For An Image On FPGA", Information And Communication Technology For Intelligent Systems, ICTIS 2017. (Smart Innovation, Systems And Technologies, 2018, 84:72-84)
- [14] F. Yavuzer-Aslan, M. T. Sakallı, B. Aslan, "Önemli Blok Şifrelerde Kullanılan Doğrusal Dönüşümlerin İncelenmesi", Akademik Bilişim '12 - XIV. Akademik Bilişim Konferansı Bildirileri 1 - 3 Şubat 2012 Uşak Üniversitesi 49
- [15] M Özbek, "Adli Bilişim Uygulamalarında Orijinal Delil Üzerindeki Hash Sorunları", 1st International Symposium on Digital Forensics and Security (ISDFS'13), 20-21 May 2013, Elazığ, Turkey
- [16] <http://primes.utm.edu/curios/index.php?start=301&stop=1000>

BIOGRAPHIES



HAKAN GENÇOĞLU Bursa, Türkiye, in 1978. He received the B.S. in Mathematics from İstanbul University in 2002 and M.S. degrees in Computer Engineering from the İstanbul Aydın University, İstanbul, in 2012 and the Ph.D. degree in Computer Engineering from Trakya University, Turkey, in 2018. Research topics are Asymmetric Encryption Algorithms performance and

applications.

BIOGRAPHIES



TARIK YERLİKAYA İpsala, Edirne, in 1977. He received the B.S. Electronics and Communication Engineering from Yıldız Technical University in 1999, M.S. and Ph.D. degrees in computer engineering from the Trakya University in 1999 and 2006.

From 1999 to 2007, he was a Research Assistant with the Trakya University Computer Engineering Department. He is Assistant Professor since 2007. His research interests focus Encryption Algorithms and their Cryptanalysis

Opinion Analysis of Software Developers Working Onsite on Public Sector Software Projects: An Exploratory Study

D. KÖSEOĞLU and O. CHOUSEINOĞLOU


Abstract— Among critical stakeholders of software projects developed in public institutions are software developers and project managers affiliated to sub-contractor companies, who are working onsite at the public institutions. The aim of this study is to identify, determine and have an initial understanding regarding how software developers are affected and what are the factors that effect software developers and project managers while working onsite in public software projects. The research is designed as an exploratory and qualitative study. The study participants were 10 software developers, one technical team leader and four project managers, all having long experiences in working on public software projects on client premises. Semi-structured interview technique is employed as the data collection method and “content analysis” was applied on the obtained participant responses. Following this analysis, “key content” approach was used to identify the issues to be further considered. It has been determined that the productivity of software developers working onsite at public institutions is low due to intense requests, shortened time to fulfill these requests, and long duration of authorization processes. Moreover, a finding of the study is that the quality of the developed projects and the motivation of the developers are considered to be low for the same aforementioned reasons. Among the foremost causes for these results are the fact that employees work under two different hierarchical structures one in the public institution and one in the software company, more than one managers assign work and there is an authority confusion between organizational units and people in the public institutions.

Index Terms— Content analysis, Exploratory study, Onsite employment, Public sector software projects, Software developer motivation.


I. INTRODUCTION

A NUMBER OF RESEARCHERS [1-3], being motivated by the fact that “people” are an important factor in software development processes, have argued that one of the

DOĞANCAN KÖSEOĞLU, is with Ekinoks Software, Ankara, Turkey, (e-mail: dogancan.koseoglu@ekinoksoftware.com).

 <https://orcid.org/0000-0002-4559-280X>

OUMOUT CHOUSEINOĞLOU, is with the Department of Industrial Engineering Hacettepe University, Ankara, Turkey, (e-mail: uhus@hacettepe.edu.tr).

 <https://orcid.org/0000-0002-8513-351X>

Manuscript received January 24, 2019; accepted July 30, 2019.
DOI: [10.17694/bajece.517419](https://doi.org/10.17694/bajece.517419)

most effective methods to improve software developers’ performance is focusing on the person and the work environment. In spite the fact that software development processes are controlled with engineering principles and approaches, software developers are still individuals who are affected by the physical and emotional environment and the conditions they work in and around them, and unavoidably their professional and work performance may be influenced by all these [4,5].

Government and public sector institutions and agencies (public institutions) generally meet their software and/or IT requirements by outsourcing these projects to external software development organizations [6]. Moreover, as the number of operations performed by these software systems in the public institutions is high and critical, in order to assure that the software meets the required quality characteristics such as security, reliability and maintainability, numerous employees are employed in the public institutions with titles as “software developer”, “software analyst”, “software project manager” and “software system manager”. These employees can be either civil servants employed directly by the public institution or employees of the software development company which is developing the software in question, who are working within the public institution on a contract in accordance with the agreement reached between the public institution and the software development company. The utilization of external resources as aforementioned is referred to as “onsite outsourcing”, where there is a governance structure requiring that all processes are carried out at the client’s premises [7], in this case the client being a public institution. Onsite outsourcing in public institutions as defined within the scope of this research is briefly displayed in Figure 1, where the software developers are shown within the domain of the public institution but are organically associated with the software development company. In the rest of this paper, for sake of consistency, software developers working in public institutions with onsite outsourcing contracts will be briefly referred to as “client side working”, client side referring to the public institution.

Onsite outsourcing arrangement offers clients with a greater potential to exert control over the project being developed, especially when the client requires involvement in all steps of the development process, onsite outsourcing is the only solution. Moreover, onsite outsourcing is considered to be advantageous for the client when the client plans to increase

its workforce with capable employees through augmenting staff, or by bringing in consultants for transferring, sharing and creating knowledge within the organization [8]. However, contrary to these benefits to the client organization, it is observed that such a work agreement between a public institution and a software development organization may lead to highly dissatisfied employees and poor performance.

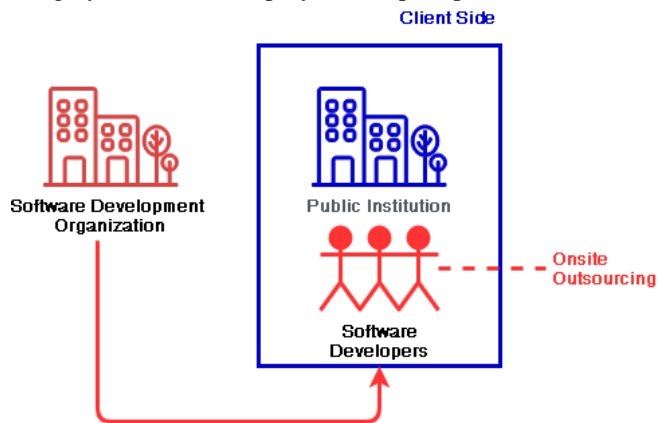


Fig.1. Onsite outsourcing in public institutions

In this study, the factors that affect the productivity, motivation and will to work of software developers working in software projects structured as onsite outsourcing projects are examined, together with the factors that affect the project plan, quality and development schedule of these. The study is a qualitative one and has been conducted by employing the semi-structured interview technique. The employed methodology is the exploratory research, as the aim of the study is to conduct an initial research to better understand the happenings on onsite outsourcing projects in public institutions, so that this initial understanding would facilitate new insights for future research. According to this aim, the research group consists of 15 subjects. The responses collected were analyzed with the use of “qualitative content analysis” technique and the findings were formed and prioritized with the use of the “inductive category formation” technique [9]. The work scheme of the software developers employed in onsite outsourcing projects in public institutions was examined with respect to the following seven research questions (RQ):

- RQ1. What are the reasons for selecting/preferring to work on the client side as part of an onsite outsourcing contract?
- RQ2. What characteristics of client side working affect the productivity of software developers?
- RQ3. What characteristics of client side working affect project quality?
- RQ4. What characteristics of client side working affect the motivation of software developers?
- RQ5. What are the factors that affect the work scheme preference of software developers?
- RQ6. What characteristics of client side working affect project schedule?
- RQ7. What are the important features of being associated to two different organizational hierarchies on software developers working on client side?

The rest of the paper is organized as follows: First, the

related work is summarized to identify the gap that this work tries to address. In the third section the details of the research methodology are introduced, and in the fourth section the findings and the key concepts related to each research question are given. The paper ends with a discussion of the findings, contributions and limitations of the present study, threats to validity and suggestions for future work.

II. RELATED WORK

Onsite outsourcing has been mostly researched and discussed in connection and as a delivery model of offshore outsourcing and global software development. The common research topics in this area with respect to onsite employees and teams are mostly focusing on knowledge sharing and codification [10-15], work design [16, 17], project cost, project schedule, project cost [18,19] and cultural differences [20]. Betz, Oberweis and Stephan [10] discuss the best practices of knowledge transfer in offshore outsourcing and globally distributed projects, by considering both onsite and offshore project teams. They argue that knowledge transfer is a critical success factor for software development projects that have an onsite and offshore structure, identify problem areas and discuss how onsite teams and offshore teams respond. Regarding knowledge, Kotlarsky, Scarbrough and Oshri [11] examine the coordination of effort between client staff and onsite and offshore vendor personnel, with respect to codification of knowledge as a process that supports coordination. Similarly, Gopal et al. [12] try to reveal the influence of knowledge codification factors that global software development teams designed as offshore and onsite have on the outcomes of the projects. Kumar and Thangavelu [13] examining the relationship of offshore and onsite teams in global software development projects, argue that knowledge sharing and knowledge transfer have a significant association with the project outcome. They argue that knowledge sharing and transfer among offshore and onsite development teams is an important issue and propose a model to address this. Mishra and Mahanty [14] try to identify the knowledge requirements and the effect of knowledge on both onsite and offshore project work division in the outsourcing software industry, concluding that there is a need of higher onsite presence in development projects. Vlaar, van Fenema, and Tiwari [15] investigate how knowledge and experience asymmetries between onsite and offshore team members force them to engage in different acts of sensegiving, sensedemanding, and sensebreaking, concluding that these acts allow them to jointly explore and generate value. Mishra and Sinha [16], by proposing a new metric named as Onsite-Ratio and examining the interrelationships between the elements of work design, project integration challenges, and project performance conclude that managers can actively lower integration glitches by increasing the levels of onsite ratio in software development projects. In a different research, Mishra and Mahanty [17] study manpower dynamics with the purpose of finding good values of onsite-offshore team strength and the number of hours of communication between onsite and offshore teams for smooth transfer of project. Their findings

show that the project of work transfer should start with initial onsite team strength higher than that of required. Finally, the effects that onsite-offshore employment has on software projects with respect to cost, schedule and quality is investigated by Mishra and Mahanty [18,19].

To the best of our knowledge, no study has been undertaken to understand how software developers are affected while working onsite on the client side in public software projects. The current study is an initial attempt to better understand how the work characteristics, professional decisions and career path of software developers is affected by this employment model. However, the existing literature has guided us on better understanding the specificities of the onsite employment model and how such an exploratory study should be conducted.

III. METHOD

A. Qualitative Research and Semi-Structured Interviews

The current study is designed as a qualitative study. Qualitative studies generally are employed to understand the individual's experiences and how they interpret these experiences [9,21]. Accordingly, interviews have been conducted with a number of employees from the software industry and as their feelings, thoughts and perceptions may differ no limitations have been imposed on their answers and therefore a more extensive content analysis has been conducted.

To conduct the qualitative study, the semi-structured interview technique has been used [22]. Semi-structured interviews are widely used in the software engineering domain in qualitative and/or exploratory studies. In a similar study Li, Ko and Begel [23] have used the semi-structured interviews to discover the interaction types between software engineers and employees from other disciplines and the important characteristics of these interactions. The main rationale for employing semi-structured interviews is to extract detailed information and to allow for further and franker discussion with the participants. Even though the questions are defined and ordered before the interview, they are not asked in the predefined order but rather are being asked as the interview proceeds, typically in an impromptu way.

In the conducted semi-structured interviews, the time-glass model has been employed [24]. In the time-glass model the interviews start with open questions that aim to relax the participant and enhance the sincerity, so that they would answer the interview questions more truthfully and without any reservations. Accordingly, the interviews were concluded by asking more generalized questions. Another reason for employing the time-glass model was that it allowed the interviewer to ask the interview questions in any order, based on the relaxation level of the participant.

B. Content Analysis

As defined by Haggarty [25] content analysis is a research method used to make generalizations with respect to specific categories of interest by systematically and reliably analyzing

the qualitative data collected in the research. In this study qualitative content analysis was employed together with the phenomenological method, which is concerned with the individual's personal perception or account of an event and aims to examine participant's lifeworld in detail by attempting to explore personal experiences [26,27]. In the current study a perspective focusing on the individuals, that is the software developers, and their responses has been chosen. The qualitative content analysis has been implemented by coding the data, identifying the themes in the coded data, by organizing the data codes and themes, and finally by identifying and interpreting the findings [28]. The themes in the coded data were formed and prioritized with the use of the "inductive category formation" technique, by summarizing and analyzing the content to develop categories gradually from the obtained semi-structured interview transcripts [9].

C. Study Group

The study group consists of 15 professionals actively employed in the software development industry and being client side working individuals; in detail 10 software developers, one team leader and four project managers. The study participants are employed in 10 different software development organizations, working according to the onsite outsourcing employment structure in 10 different public institutions and in 12 different projects. All participants are residing and working in Ankara, Turkey. The participants have been selected according to the maximum variation / heterogeneous purposive sampling in order to provide a diverse range of cases relevant to client side working software developers. The demographic details of study participants are given in Table I.

TABLE I
STUDY GROUP DEMOGRAPHICS

		N	%
Gender	Female	5	33.33
	Male	10	66.67
Title	Software Developer	10	66.67
	Team Leader	1	6.67
	Project Manager	4	26.67
Graduated University Departments	Computer Science	13	86.67
	Statistics	1	6.67
	Computer Technology / Information Systems	1	6.67
Experience	0-9	9	60.00
	10-19	6	40.00
Years Working in Client Side	1-3	11	73.33
	3-5	4	26.67

It has been observed that analogous number of participants have taken part in similar exploratory studies when semi-structured interviews are employed: Betz, Oberweis, and Stephan [10] explore the problems and best practices of knowledge transfer in offshore outsourcing by conducting semi-structured interviews with 13 experts, whereas Mishra and Mahanty [14] by conducting semi-structured interviews with 14 experts try to identify the knowledge requirements and the effect of knowledge on both onsite and offshore project work division. Winkler, Dibbern, and Heinzl [20] have conducted interviews with a total of nine project managers to explore the nature of cultural differences in offshore outsourcing and to analyze the relationship between those

cultural differences and offshore outsourcing success, with respect to onsite-offshore employment model.

D. Data Collection

Semi-structured interviews have been used as the data collection technique in order to ease the participants and allow them to give sincerer responses, as they are not as strict and fixed in organization as the structured interviews and not as loose as the fully unstructured interviews. The interviews were conducted face-to-face separately between the period of April 26, 2018 – May 10, 2018. The interview locations were independent locations and were carefully selected such as they would allow the interviewed participant to feel comfortable and relaxed, and consequently to answer the questions frankly. Before the interviews and during them, all required information about the research and the interview questions were given to the participants.

In order to avoid any loss, the responses of the participants were recorded with the use of a voice recording device. The participants were informed that a voice recording device would be used and their consent was obtained before the start of the interviews. Following the conclusion of the interview, the voice recordings were shared with the participants and they were informed that a part or the whole recording may be removed from this research.

E. Data Analysis

As the first step of the analysis, the interview transcripts were decoded and parsed by the researcher by typing them in a tabular form. In order to ensure the confidentiality of the participants shortened code names were used in the data forms (e.g., MSD1). The forms were filled separately by listening to the voice recordings, were shared individually with the participants with the use of different sharing media (e.g. e-mail, message) in order to avoid any errors while recording the interview responses and were analyzed only after the respondents confirmed the transcripts.

The data extracted from the responses were grouped based on the interview questions and themes were constructed based on the data and the groups, in accordance with the inductive category formation technique. The formed themes were named as “key contents” and were grouped with respect to respondents and interview questions, in order to calculate their frequencies. These key contents were used to identify the factors and factor categories that affect the productivity, motivation and will to work of the respondents, and the factors that affect the plan, quality and development schedule of the project, in accordance with the stated motivation of this study.

The significance and integrity of the findings has been evaluated by the authors consistently. In order to ensure the integrity of the findings, the factors that constitute the themes were checked within themselves and with the other themes, and cross-checked to test if they form a meaningful whole. Finally, the findings were reviewed by other software practitioners and they have been characterized as realistic.

Moreover, during the identification of these key concepts and factors, the positive or negative perception of the

respondents regarding each one of these factors was recorded by the interviewer either based on the respondent’s statements or by explicitly asking questions. These perceptions are given next to the associated factors in the Findings section as (+) or (-), denoting positive or negative perceptions, respectively.

IV. FINDINGS

– RQ1. What are the reasons for selecting/preferring to work on the client side as part of an onsite outsourcing contract?

The most common answer (7 responses) by the participants was that they were “appointed by the management of the software development organization to work on client side”. This, together with the “fear of losing job” (2 responses) and “conflicts with previous employers and managers” (5 responses) displays that involuntary responses (14, shown with -) are higher than responses that refer to voluntary reasons (4, shown with +). All responses for RQ1 are given in Table II.

TABLE II
RESPONSES FOR RQ1

	N
Company appointment & compulsory employment	- 7
Conflicts with previous employers and management	- 5
Wanting to acquire client-side work experience	+ 3
Fear of being unemployed	- 2
Wanting to work at a more corporate organization	+ 1

– RQ2. What characteristics of client side working affect the productivity of software developers?

The participants have provided a number of characteristics regarding the client side working scheme that affect their productivity, which are given in detail at Table III.

TABLE III
RESPONSES FOR RQ2

	N
Excessive requirements by the client	- 4
Shorter periods/deadlines given to complete work	- 4
Lag/tardiness in authorization and approval processes	- 4
Lack of technical knowledge by the public institution employees	- 3
Unwillingness of the public institution employees to work	- 2
Public institution employees abstaining from innovations	- 2
The necessity of communicating one-to-one with the public institution employees	- 2
The corporate organization structure	+ 1
Complexity of the existing legacy systems	- 1
Dealing directly with the customer	+ 1
Public institution employees preferring to communicate via telephone	- 1
The existence of outdated technologies	- 1
Being part of more than one projects at the client side	- 1
Increased effect of laws and regulations on the developed project	- 1
Projects in public institutions have a longer development lifecycle	- 1
The organizational culture of the public institution	- 1
Negative attitude of the public institution employees	- 1
The work commitment of the public institution employees	- 1
The projects have attention getting / spectacular contents	+ 1
Use of subcontractors instead of project team	- 1
The indifference of the company towards its employees	- 1
The extensive length of work history with the public institution	+ 1
The experience level of the project manager	+ 1
Working as a team on the client side	+ 1
The coordination between the software development company and the project team working at client side	- 1
Having to work both in the software development company and the client side (at the same time)	- 1
Working with the highest ranking responsible person at the client side	+ 1

The answers that have been given most are “excessive

requirements by the client”, “shorter periods/deadlines given to complete work” and “lad/tardiness in authorization and approval processes”. The majority of these characteristics, as stated by them, are affecting their productivity negatively (34, shown with –) and only few positively (7, shown +).

– RQ3. What characteristics of client side working affect project quality?

The respondents have stated that “focusing on having the job finished (because of excessive requirements and short deadlines)” and “public institution employees just demand that the work s finished” are the most common factors (both appearing 6 times) that affect project quality. The complete list of these characteristics is given in Table IV with factors that affect the project quality negatively shown with a – (a total of 31 responses) and the ones affecting quality positively shown with a + (a total of 7 responses).

TABLE IV
RESPONSES FOR RQ3

		N
Focusing only on having the job finished because of excessive requirements and short deadlines	-	6
Public institution employees just demand that the work is finished	-	6
Public institution employees lacking the knowledge to assess the work quality	-	4
Using outdated technologies on client side and not risking modernization	-	4
Software development in accordance with the requirements of a corporate organization	+	2
The unwillingness of the project team to work at client side	-	2
Project team members thinking that their liberty has been compromised	-	2
Modernization of client side technology is very expensive	-	2
The necessity of a very solid project plan and analysis	+	2
The experience level of the project manager	+	1
Being in an environment isolated from the organizational culture	+	1
The continuous administrator/manager change in the public institution	-	1
Public institution employees abstaining from innovations	-	1
The processes of the public institution are not discussable	-	1
The project team having few and highly qualified members	+	1
The hierarchical structure and "ego wars" at public institution	-	1
Not developing in accordance to the quality standards	-	1

– RQ4. What characteristics of client side working affect the motivation of software developers?

Regarding the factors that affect their motivation the respondents mentioned that discussions and quarrels in the public institution work environment (5 responses) mostly upset them. This factor has appeared in the responses to be related with the second most common item (“ego wars”, 4 responses). As it can be seen in the results in Table V, the majority of the responses are related with negative factors (a total of 44) and only 7 responses mention factors that affect their motivation positively.

– RQ5. What are the factors that affect the work scheme preference of software developers?

When the factors that affect software developers work scheme preferences were examined, it is observed that three major responses come forward, namely: “career change possibilities” (10), “the contents of the project” (7) and “the existence of hierarchies” (5). Contrary to the responses that were collected for the other research questions, it was observed in Table VI that respondents have mentioned positive factors more (total of 26) with respect to negative factors (9).

TABLE V

RESPONSES FOR RQ4

		N
The discussion (quarrel environment) that exists because of work overload and time limitations	-	5
The hierarchical structure and "ego wars" at public institution	-	4
Interferences with the work schedule	-	3
The bureaucracies at public institution	-	3
Having to work overtime	-	3
Public institution employees acting as managers	-	2
Environmental factors regarding the public institution	-	2
The software development company decides and plans the onsite working	-	2
Not having the expected working environment at client side	-	2
The projects have attention getting / spectacular contents	+	2
Restrictions on social media, video, music	-	2
Knowing that client side working is temporary	+	2
Concern that there is no chance for self-improvement	-	2
Lack of job satisfaction	-	2
The continuous administrator/manager change in the public institution	-	2
The affiliation with the software development organization is just on paper	-	2
The project plan is a good one	+	1
The processes take longer to complete	-	1
More facilities and resources at the software development organization	-	1
Being held responsible for external bugs / defects	-	1
Working with a highly qualified team	+	1
Being motivated with the work	+	1
Negative attitude of the public institution employees	-	1
The work commitment of the public institution employees	-	1
Working with subcontractors	-	1
Not following project plans or not having plans at all	-	1
Having to communicate with public institution employees who lack process knowledge	-	1

TABLE VI
RESPONSES FOR RQ5

		N
The possibilities of career change with respect to professional capabilities	+	10
The projects have attention getting / spectacular contents	+	7
The existence of hierarchies and bureaucracies	-	5
To feel more comfortable and more free	+	3
More capable / better colleagues	+	3
The use of outdated technologies	-	2
The prospect of starting your own family	-	2
One-to-one contact with the public institution employees	+	2
Positive attitude of the managers	+	1

– RQ6. What characteristics of client side working affect project schedule?

The study participants have stated that the “intensity of requests and limited time given to complete these requests” is the most common factor (appearing 5 times) that affect the project schedule. As it can be observed from the complete list of these factors in Table VII, the majority are factors that affect the project schedule adversely (22 responses, shown with a -) and only 5 responses mention a positive affect (shown with a +).

TABLE VII
RESPONSES FOR RQ6

		N
Intensity of requests and limited time given to complete these requests	-	5
The continuous administrator/manager change in the public institution result to change in processes	-	3
More or less a project plan is formed	+	3
Hierarchy and bureaucracy engagement to project	-	3
The existence of numerous project management approaches	-	2
Lag/tardiness in authorization and approval processes	-	2
The software product is associated with other institutions	-	1
Requests are initiated by more than one sources	-	1
The move from legacy system to the new system	-	1
Time for fulfilling requirements	-	1
Public institution employees' off-topic requests	-	1
The conformance of the subcontractors to the planned processes	+	1
Lack of technical knowledge by the public institution employees	-	1
"Ego wars" between public institution employees	-	1
Number of in-team employees related to contract	+	1

– RQ7. What are the important features of being associated to two different organizational hierarchies on software developers?

Regarding the last research question, respondents were asked to state the most important features of being associated to two different organizational hierarchies (one in the software organization and one in the public institution). As it is shown in Table VIII, participants have only mentioned negative features regarding such a working scheme.

TABLE VIII
RESPONSES FOR RQ7

		N
Works are assigned by more than one person	-	6
Requests are initiated by more than one sources	-	5
"Ego wars" between public institution employees	-	5
No connection between employees and software development organization	-	4
Public institution employees affecting the project manager with their bureaucratic attitude	-	3
Public institution employees considering themselves in a higher position compared to software development organization	-	2
Negative approach of the public institution managers	-	2
Interference by public institution and software development organization to software development process	-	1
The attitude of public institution employees towards the software development organization because of their own process lag/tardiness	-	1
Not adopting the public institution culture	-	1
Not taking into consideration any person or event other than work	-	1
Irregularities in the software development organization management levels	-	1

V. DISCUSSIONS

A. Validity Threats

The identified and addressed validity threats of this qualitative and exploratory study are detailed below. The most important construct validity threat is that the participants may have not correctly and fully understood the characteristics and related issues of working on the client side in a public institution, and may have not correctly communicated their thoughts on the subject during the interviews. In order to minimize this, the selected participants were software practitioners having long years of experience working on public institution software projects on the premises of the client. Regarding internal validity and instrumentation, the questions asked to the respondents were formed with respect to the set research questions and they were asked to all participants. However, in accordance to the characteristics of the semi-structured interviews, the questions were not asked in a specific order, but the order would change with respect to the responses of the participants. Moreover, to ensure internal validity the responses of the participants were recorded with the use of a voice recording device. Only after the participant was informed about this, and had consented on the use of this voice recording device the interview was initiated. Following the conclusion of the interview, the collected responses after being processed and the related voice recordings were accordingly shared with the participants and only after receiving their approval were further processed in order to obtain the key concepts. This study was conducted with participants working in Ankara, Turkey and in public institutions located in Ankara, Turkey; therefore, the results cannot be generalized. Yet, in order to provide a diverse range of cases and to obtain diverse responses, the participants have

been selected according to the maximum variation / heterogeneous purposive sampling. However, as the study was planned as an exploratory research, the results instead of being generalizable should provide insights for future generalizable studies.

B. Findings

When the responses for the seven research questions (Table II to Table VIII) are evaluated in overall it is seen that 186 (76.8%) responses display a negative sentiment and only 56 (23.2%) of them embody a positive sentiment, an indication that software developers approach and perceive this employment method negatively. Next, we performed word cloud visualization and quantitative readability analysis to see the big picture of the given responses. After removing articles, prepositions and common words that were in the question text (e.g. public, institution, employees), we used *wordclouds.com* to generate a word cloud showing the focuses of onsite working software developers based on their responses. Figure 2 depicts the word cloud and Table IX lists the 30 most common terms.



Fig.2. Onsite outsourcing in public institutions

Based on the results from data analysis with respect to the research questions, it is observed that software developers do not prefer to work on the premises of the client, but are forced to work on the client side following the successful bid and winning of the tender by the public institution. With respect to this, the majority of the participants have mentioned that they do experience conflicts with the management of the software organization that they work at and that are currently in search of a new job. However, as they have significant concerns regarding unemployment, when they are confronted with the obligation of having to work on client side in a public institution, they prefer to accept such position offerings even if they do not prefer this work model at all. On the other hand, when asked about the advantageous aspects of client side

working, the participants after evaluating the overall situation of the software industry, have stated that almost all software companies have contractual agreements with the public institutions.

TABLE IX
THIRTY MOST COMMON TERMS IN THE RESPONSES OF THE STUDY PARTICIPANTS

	N		N		N
projects	37	manager	11	complete	10
client	24	wars	11	contents	10
development	20	ego	11	getting	10
change	19	possibilities	10	respect	10
requests	17	professional	10	career	10
organization	16	capabilities	10	team	10
requirements	13	spectacular	10	knowledge	9
processes	12	excessive	10	bureaucracies	8
finished	12	deadlines	10	environment	8
time	12	attention	10	management	8

It is observed that the software developers have different views with respect to their productivity when working on client side projects. The participants have stated that works are assigned to them frequently and that assigned works have short completion times, thus they focus only on finishing the assigned work item on time, without having the opportunity to develop new ideas or approaches with respect to the problem at hand. Especially the participants working as project managers on projects developed on the client side have mentioned that the processes on the public institutions are slow, they have to exert too much effort and allocate too much time in order to obtain the required permissions and authorizations, thus they believe that they are becoming less productive.

On the other hand, when the results regarding the responses of the software developers on project quality are analyzed, it can be seen that there are similarities with the answers given for the productivity aspects in client side working. They have stated that the quality of a software project depends on the amount of time spent for reasoning, contemplation and rigorous planning; however, they generally do not have time for such an approach and they need to finish numerous work items in a very short time. This in return, as they have mentioned, leads to the project team members forsaking quality for the sake of finishing tasks in time. Moreover, it has been identified as a recurring observation that the staff at the public institution does not control the quality of the project but only focuses on the completion of work packages and tasks. The participants believe that the high turn-over rate of public servants and the apprehension they feel regarding the rapid changing processes is the major reason for this behavior.

It has been observed that for the most part of their work experiences the motivation of the participants is low. They have mentioned different explanations for this; however, the intense work schedule and the expectation of completing urgent and pressing tasks are the predominant reasons. The participants consider working on the client side partially as being unsuccessful and as having to keep up continuously with deadlines, findings that are consistent with the facts in literature [4,5,29]. Moreover, the interpersonal relationships

and team dynamics in the client side are generally different from the ones that the participants are accustomed to at the software development organization. The participants have stated that these new interpersonal relationships, blended with strict and structured hierarchical rules and bureaucracies, hinder the relaxed, resilient and high-motivated work of the software developers.

Almost all participants to this study have clearly stated that they would prefer to work on the premises of the software company instead of the client. Most of them have mentioned that their professional skills and competencies deteriorate when working on the client side, the possibility that the work tasks would get standardized is higher and that they generally do not add or implement new technologies but only solve problems by developing new algorithms. However, they mention that when working on the software company they have the opportunities to work with new technologies, testing and implementing them in projects; something which is not possible on the client side in public institutions as the software generally is already a being used system. The majority of the participants have stated that they would accept to work on the client side if the contents of the project are challenging and if the project is about an incoming software. Furthermore, some of the participants have stated that the hierarchical structure and the bureaucracy in the public institutions affect them negatively, and that they would prefer to work in a friendlier environment with coworkers and supervisors who approach them positively and supportively.

The participants have stated that they experience a number of issues regarding project planning, duration assignment to tasks and management of the projects when working at the client side. According to them the main reason for this is that the supervisors and employees in the public institutions generally comprehend and use project management concepts differently and almost always want their requests to be fulfilled immediately. Moreover, it has been mentioned during the interviews that there is a high turn-over rate in the public institutions and that managers change in an unpredictable way, with new managers bringing conflicting requirements and the necessity to plan the processes from scratch. Another important finding was that the participants perceive the technical knowledge and skills of the public institution employees to be inadequate, commenting that most of the time they are required to explain and train them regarding new technologies, hardware and programming languages.

Lastly, it has been observed from the interview texts that the assignment of the tasks from a single authority point has a positive effect on the participants. As working on the client side is already difficult because software developers are in an environment different from what they are used at, conflicts between the public institution employees result to changes in project plans that hinder the work. According to the participants, working under two different hierarchies results to being assigned work by both the public institution and the software company managers, whereas ideally the software developer should only be assigned work by the direct supervisor.

VI. CONCLUSION

As a result of this exploratory study it can be argued that the participants prefer to work with qualified, having positive attitude and favorable coworkers and managers, in a relaxed work environment isolated from hierarchical and bureaucratic structures, on projects that employ cutting edge technologies and follow the widely accepted project management standards.

With respect to this and based on the findings of the research questions, a future study employing structured tools like questionnaires has been planned to analyze in depth the software developers and software teams working on the client side, but also the characteristics of software projects developed with this employment model. Moreover, to better understand the dynamics and structure of the onsite outsourcing arrangement in public institutions for software development, not only employees affiliated with the developer side but also employees and managers from the client side are considered to be participants of the planned future study.

REFERENCES

- [1] B. W. Boehm, P. N. Papaccio, "Understanding and controlling software costs". IEEE transactions on software engineering, 14 (10), 1988, pp. 1462-1477.
- [2] A. Cockburn, J. Highsmith, "Agile software development, the people factor". Computer, 34(11), 2001, pp. 131-133.
- [3] S. de Barros, et al., "A review of productivity factors and strategies on software development", In 2010 Fifth International Conference on Software Engineering Advances (ICSEA), August 2010, pp. 196-204.
- [4] R. Ilies, T. A. Judge, "Understanding the dynamic relationships among personality, mood, and job satisfaction: A field experience sampling study". Organizational behavior and human decision processes, 89(2), 2002, pp. 1119-1139.
- [5] D. Graziotin, X. Wang, P. Abrahamsson, "Happy software developers solve problems better: psychological measurements in empirical software engineering". PeerJ, 2014, p. e289
- [6] P. Mohagheghi, M. Jørgensen, "What Contributes to the Success of IT Projects? An Empirical Study of IT Projects in the Norwegian Public Sector". Journal of Software, 12(9), 2017, pp. 751-759.
- [7] J. H. Lim, V. J. Richardson, R. W. Zmud, "Value implications of IT outsourcing contextual characteristics". Unpublished article, 2007.
- [8] T. Kremic, O. Icmeli Tukul, W. O. Rom. "Outsourcing decision support: a survey of benefits, risks, and decision factors" Supply Chain Management: An International Journal 11.6, 2006, pp. 467-482.
- [9] P. Mayring, *Qualitative content analysis: theoretical foundation, basic procedures and software solution*, 2014.
- [10] S. Betz, A. Oberweis, R. Stephan. "Knowledge transfer in offshore outsourcing software development projects: An analysis of the challenges and solutions from German clients." Expert Systems 31.3, 2014, pp. 282-297.
- [11] J. Kotlarsky, H. Scarbrough, I. Oshri. "Coordinating Expertise Across Knowledge Boundaries in Offshore-Outsourcing Projects: The Role of Codification." MIS Quarterly 38.2, 2014, pp. 607-627.
- [12] J. Gopal, et al. "Towards identifying the knowledge codification effects on the factors affecting knowledge transfer effectiveness in the context of GSD project outcome." Emerging ICT for Bridging the Future- Proceedings of the 49th Annual Convention of the Computer Society of India (CSI) Volume 1. Springer, Cham, 2015.
- [13] S.A. Kumar, A. Kumar Thangavelu. "Factors affecting the outcome of Global Software Development projects: An empirical study." 2013 International Conference on Computer Communication and Informatics, IEEE, 2013.
- [14] D. Mishra, B. Mahanty. "Business knowledge requirements and onsite offshore work division in Indian software outsourcing projects." Strategic Outsourcing: An International Journal 8.1, 2015, pp. 76-101.
- [15] P. Vlaar, P.C. van Fenema, V. Tiwari. "Cocreating understanding and value in distributed work: How members of onsite and offshore vendor teams give, make, demand, and break sense." MIS quarterly 32.2, 2008, pp 227-255.

- [16] A. Mishra, K. K. Sinha. "Work design and integration glitches in globally distributed technology projects." Production and Operations Management 25.2, 2016, pp. 347-369.
- [17] D. Mishra, B. Mahanty. "Study of maintenance project manpower dynamics in Indian software outsourcing industry." Journal of Global Operations and Strategic Sourcing 12.1, 2019, pp. 62-81.
- [18] D. Mishra, B. Mahanty. "The effect of onsite-offshore work division on project cost, schedule, and quality for re-engineering projects in Indian outsourcing software industry." Strategic Outsourcing: An International Journal 7.3, 2014, pp. 198-225.
- [19] D. Mishra, B. Mahanty. "A study of software development project cost, schedule and quality by outsourcing to low cost destination." Journal of Enterprise Information Management 29.3, 2016, pp. 454-478.
- [20] J.K. Winkler, J. Dibbern, A. Heinzl. "The impact of cultural differences in offshore outsourcing—Case study results from German–Indian application development projects." Information Systems Frontiers 10.2, 2008, pp. 243-258.
- [21] D. J. Bluhm, et al., "Qualitative Research in Management: A Decade of Progress", Journal of Management Studies, 2011 p. 48.
- [22] C. Robson, *Real World Research*. Oxford: Blackwell, 2002
- [23] P. L. Li, A. J. Ko, A. Begel. "Cross-disciplinary perspectives on collaborations with software engineers". In 2017 IEEE/ACM 10th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE), 2017, pp. 2-8
- [24] P. Runeson, M. Höst, "Guidelines for Conducting and Reporting Case Study Research in Software Engineering", Empirical software engineering, 14(2), 2008
- [25] L. Haggarty, "What is content analysis?." Medical Teacher 18.2, 1996, pp. 99-101.
- [26] V. Eatough, J. A. Smith. "Interpretative phenomenological analysis.", The Sage handbook of qualitative research in psychology, 2008.
- [27] R. H. Hycner, "Some guidelines for the phenomenological analysis of interview data." Human studies 8.3, 1985, pp. 279-303.
- [28] Ü. Akyuz, "Assessment of European Union Funded Projects Finalised by the Ministry of National Education based on the Opinions of Administrators and Experts", TED Education and Science, 2015, pp. 199-201.
- [29] Y. Lin, W. McKeachie, Y. Kim, "College Student Intrinsic and/or Extrinsic Motivation and Learning". Learning and Individual Differences, 2003, pp. 251-258.

BIOGRAPHIES



DOĞANÇAN KÖSEOĞLU was born in Ankara, Turkey, in 1992. He received his B.S. degree in Computer Engineering from Selçuk University, Konya, Turkey, in 2015 and his M.S. degree in Engineering Management from Hacettepe University, Ankara, Turkey in 2018. Currently he is working as a software engineer at Ekinoks Software, Ankara. His research interests include processes of software lifecycle management and software project planning.



OUMOUT CHOUSEINOGLU received his B.S. degree in Business Administration and M.S. and Ph.D. degrees in Information Systems from Middle East Technical University (METU) in 2000, 2004 and 2012 respectively. Since 2013, he has been with the Department of Industrial Engineering, Hacettepe University, Turkey. His current research and teaching interests include software engineering, software project management, cloud computing, data communications and networking, and management information systems.

Comparison of Contourlet and Time-Invariant Contourlet Transform Performance for Different Types of Noises and Images

M. F. ASLAN, K. SABANCI and A. DURDU


Abstract— A noiseless image is desirable for many applications. However, this is not possible. Generally, wavelet-based methods are used to noise reduction. However, due to insufficient performance of wavelet transforms (WT) on images, different multi-resolution analysis methods have been proposed. In this study, one of them is Contourlet Transform (CT) and the Translation-Invariant Contourlet Transform (TICT) which is an improved version of CT is compared using different noises. The fundus images are taken from the DRIVE dataset and benchmark images are used. Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), Mean Structural Similarity (MSSIM) and Feature Similarity Index (FSIM) are used as comparison criteria. The results showed that TICT is better in Gaussian noisy images.

Index Terms—Contourlet Transform, Image Denoising, Time Invariant Contourlet Transform


I. INTRODUCTION

IN RECENT years, many algorithms have been developed for image processing [1-3]. Therefore, digital image processing has been available in some areas, such as physics, defense industry, medicine, industrial applications, robotics, intelligent transportation systems, etc [4]. Considering the application fields, it is understood that it is used for important and sensitive tasks. However, in the process of obtaining an image, noise occurs on the image. This may be caused by the quality of the camera or the environment conditions. According to Patil [5], there is no noiseless signal.


MUHAMMET FATIH ASLAN, is with Department of Electrical and Electronic Engineering, Karamanoglu Mehmetbey University, Karaman, Turkey, (e-mail: mfatihaslan@kmu.edu.tr).

 <https://orcid.org/0000-0001-7549-0137>

KADIR SABANCI, is with Department of Electrical and Electronic Engineering, Karamanoglu Mehmetbey University, Karaman, Turkey, (e-mail: kadirsabanci@kmu.edu.tr).

 <https://orcid.org/0000-0003-0238-9606>

AKIF DURDU, is with Department of Electrical and Electronic Engineering, Konya Technical University, Konya, Turkey, (e-mail: adurdu@kton.edu.tr).

 <https://orcid.org/0000-0002-5611-2322>

Manuscript received June 03, 2019; accepted September 16, 2019.
DOI: [10.17694/bajece.573583](https://doi.org/10.17694/bajece.573583)

An image can also be considered as a 2D signal. Therefore, no image or video obtained can be noiseless. Therefore, it is necessary to remove the noise in order to obtain results that are more accurate.

There are many kinds of noise that can cause the image distortion. In general, noise types are Gaussian noise, Random noise, Salt and Pepper noise, Poisson noise, and Speckle noise [6]. Particularly in the case of remote sensing applications, the majority of the problem is caused by Speckle noise. Speckle noise directly reduces the quality of the image [7].

Denoising or noise reduction includes applications for removing noise that occurs after the image has been acquired. Different statistical methods have been developed to image enhancement [8, 9]. However, these methods blur the edges with a low pass filter for spatial filtering problems. It also strengthens the background with high-pass filter [10]. Linear techniques are also used in denoising. However, in the case of impulsive noise, such filters are insufficient [6]. Fourier Transform (FT) is also an alternative for denoising. But, while FT provides frequency resolution, it does not provide a time resolution. Therefore, the spatial location of the frequency change due to noise at one point cannot be determined. This problem can be solved by Short Time Fourier Transform (STFT), but in STFT method the window width is constant. A large window provides good frequency resolution but causes poor time resolution. Likewise, a narrow window provides good time resolution, but the frequency resolution is poor [11]. Therefore, the window width should not be constant and the window width should be changed depending on the frequency. This situation can be achieved by the wavelet transform which also includes the scale variable.

Wavelet transform (WT) is a time-scale analysis method used in image compression [12], edge detection [13] and deconvolution [14], besides image denoising [6]. There is an inverse relationship between the scale and the frequency. Wavelet algorithms process data at different scales or resolutions. By comparing the signal and wavelet in different scales and positions, a two-variable function is obtained. A smaller scale factor means more compression of the wavelet. Thanks to scaling, high frequency behaviors at low scales and low frequency behaviors at high scales are better analyzed. This is very important for non-stationary signals.

The WT of a continuous signal is called the Continuous Wavelet Transform (CWT). Due to the CWT requires an infinite number of inputs, it is not suitable for the computer. It also has a disadvantage in terms of speed. Therefore, the Discrete Wavelet Transform (DWT) is used for computer-based systems. While the scale (s) and position (τ) parameters are real numbers in the CWT, they get an integer value in DWT. DWT occurs by sampling the CWT.

The signal consisting of a discrete time series in DWT is divided into different frequency ranges. Because of this, the original signal is passed through the high-pass filters (HPF) and the low-pass filters (LPF), and the image or signal is divided into subbands. Then, time series are divided into low frequency (Approximation (A)) and high frequency (Detail (D)) components. As a result, approximate and detail coefficients are obtained. The coefficient A represents the low-frequency values in the time series, and the D coefficient represents the high-frequency values of the time series. This process may continue iteratively. Thus, the multi-resolution analysis (MRA) of the signal/image in the frequency domain is obtained [15, 16].

Wavelet-based transforms analyze the signal in the frequency-time domain. This corresponds to edge detection in the images. Edge information is a feature that best describes an image. One-dimensional transforms such as Fourier and WT are commonly used in capturing edges. In an image, one-dimensional edges, such as scan lines, are well decomposed by wavelets. However, the edges in natural images are not limited to this. The points of discontinuity can lie along the curve depending on an object in the image. Most of the natural images contain intrinsic geometric structure. Due to the WT is a one-dimensional transform, images are applied to the row and column for WT. For 2D data, the wavelets are well decomposed to the discontinuities at the edge points, but cannot smooth and continuously represent the edge points along the curve. In addition, WT has limited directional information [17]. WT is not given good results in speckle noise reduction [18]. Different MRA methods have been developed to solve such problems. For example, in the Ridgelet transform [19], angular windows are used, so that unlike WT, data is processed in different directions. In curvelet transform [20], windows are applied along second order curves. For Ripplet transform [21], the window is applied along the higher order curve. In the Tetrolet transform [22], the image is divided into 4x4 blocks. The most appropriate tetromin is selected for each block and applied WT to this region [23]. The Contourlet Transform (CT) [17] used in this study analyzes the smoothness along the contours better than WT by using the multiresolution and direction filter.

Denosing has always been one of the main problems in image processing. It is always desirable to protect edges, corners and other important features during the image denosing process. Image denosing is a highly needed method, especially in biomedical applications. Saha, et al. [24] introduced two mathematical transforms, wavelet and curvelet, in the field of biomedical imaging. Applications of the two transforms were compared using biomedical images. Jain and Tyagi [25]

proposed a new edge preserving image denosing method based on adaptive thresholding method and Tetrolet transform. The noisy image is decomposed into the tetrolet coefficients via a tetrolet transform. Using the locally adaptive threshold, the tetrolet coefficients are thresholded to effectively reduce noise while preserving the necessary features of the image. Huang, et al. [26] presented a novel multiscale approximation method called adaptive digital ridgelet (ADR) transform. Unlike conventional transform methods, this algorithm can adaptively handle line and curve information in an image, taking into account its infrastructure. Using a new curve part detection method, the curve parts in an image are detected. When this method was applied experimentally, successful results were obtained in image denosing application.

In this study, image denosing was performed. The CT and the TICT methods developed by Eslami and Radha [27] were used. The denosing application was made using the fundus and benchmark images. The performance of both methods was compared by using different noise and different noise ratios. In [27], TICT and STICT have been proposed as an alternative to CT and comparisons have been made. However, comparisons were made using Gaussian white noise. Noise types vary depending on the image used and the application field. Therefore, it would be appropriate to determine the transform method according to the noise distribution. For example, the noise distribution in traditional magnitude MR images is Rician [28]. Therefore, the aim of this article is to observe the results of TICT and CT at different noise distributions.

II. CONTOURLET TRANSFORM

Fourier and WT are 1D transform developed to capture discontinuous points. Therefore, the contours of the internal geometric shapes found in an image are determined locally by these transforms. Thus, the geometric smoothness of the contours cannot be achieved. To achieve smoothness along the contour in multi-dimensional signals, CT has been developed. Do and Vetterli [17] indicated the difference between WT and CT with Fig. 1. While only point discontinuities can be captured with wavelets according to Fig. 1, the series of linear points can be captured with CT. Thus, the image is represented with less coefficient.

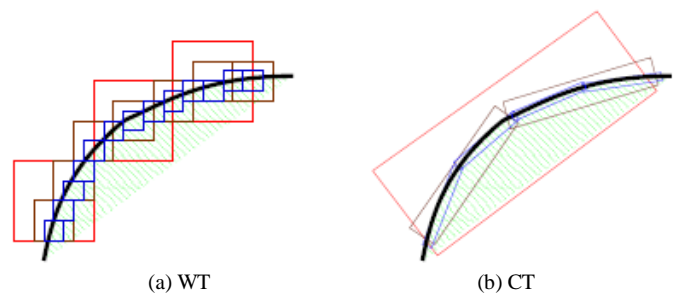


Fig. 1. Edge catching strategies using WT and CT

The double filter structure is used in CT. Firstly, Laplacian pyramid (LP) [29] is used to capture discontinuities in the image. The discontinuous points have been achieved as a result

of this. After that, the Directional Filter Bank (DFB) [30] is applied to transform discontinuous points into smooth geometric shape contours (see Fig. 2). In fact, the contour components obtained by decomposition are combined with the DFB. Combined contours can be in different direction and scale. In this way, more continuous edge points are achieved than WT. This can be easily understood from Fig. 3.

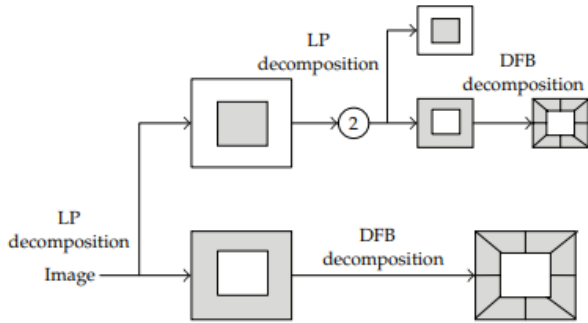


Fig. 2. The structure of CT [31]



(a) WT (b) CT
Fig. 3. Transforms for the same image [17]

III. TRANSLATION-INVARIANT COUROLET TRANSFORM

The energy calculated by the analysis of a wave using DWT is different from the energy calculated by the shift of the same wave. This is called a time-variant. Since LP and DFB resemble WT in terms of sub-sampling, CT is time-variant. To improve performance in image denoising, transform should be translation-invariant (time-invariant). Therefore, Translation-Invariant Contourlet Transform (TICT) by Eslami and Radha [27] was proposed. The big benefit of time invariant is that the performance of denoising studies is significantly improved when a time invariant scheme of subsampled transform is used [27]. Since the CT sub-sampled LP and DFB included, the TI method was applied at both process to generate a TICT.

IV. APPLICATION AND RESULTS

In this study, five benchmark images and 40 fundus images taken from the DRIVE dataset [32] were used (see Fig. 4.). The used benchmark images and some fundus images are shown in Fig. 4. Firstly, Random, Gaussian and Rician noises were added to these images respectively (sigma = 5, 10, 15 for Random noise; signal-to-ratio (snr) = 3, 5, 10 for Gaussian and Rician noise). Equations of these noise distributions are shown below respectively.

$$D1 = I + \text{sigma} \times \text{rand}(\text{size}(I)) \quad (1)$$

$$D2 = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2)$$

$$D3 = \frac{M}{\sigma^2} e^{-\frac{M^2+A^2}{2\sigma^2}} I_0\left(\frac{A \times M}{\sigma^2}\right) \quad (3)$$

In Equation 1 above, I represents the source image. In Equation 2, X is a random variable. It is usually shown as follows.

$$X \sim N((\mu, \sigma^2)) \quad (4)$$

In Equation 4, μ is the mean value of the Gaussian distribution. σ is the standard deviation value. In Equation 3, I_0 is the modified zeroth-order of Bessel function of the first kind. This is called as Rice density. M is observed noisy intensity and A is true signal intensity [33].

After the noise is added, these distorted images were analyzed using CT and TICT. The threshold were applied to the components obtained and then reconstruction was performed. As a result, the result image was compared to the original image. MSE, PSNR, MSSIM and FSIM metrics were used as comparison criteria. The mathematical expression of these metrics is shown below. $I(x, y)$ in Equation 5 represents the source image. M and N are image sizes. C_1, C_2, C_3 in Equation 7 are constant values. μ and σ are mean values and standard deviation, respectively. The FSIM in Equation 9 is a metric based on phase congruency (PC) and gradient magnitude (GM). S_{PC} is a local similar map of PC between PC_x and PC_y . S_{GM} is a local similar map of PC between GM_x and GM_y . T is a constant value [34]. The results obtained using the metrics described below are shown in Table I and Table II.

$$MSE = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [I(x, y) - I'(x, y)]^2 \quad (5)$$

$$PSNR = 20 \log_{10} \frac{255}{\sqrt{MSE}} \quad (6)$$

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)(\sigma_{xy} + C_3)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)(\sigma_x\sigma_y + C_3)} \quad (7)$$

$$MSSIM = \frac{1}{M} \sum_{i=1}^M SSIM(x_i, y_i) \quad (8)$$

$$FSIM = \frac{\sum_{x \in \Omega} S_{PC} x S_{GM} x PC_m}{GM_x^2 + GM_y^2 + T} \quad (9)$$

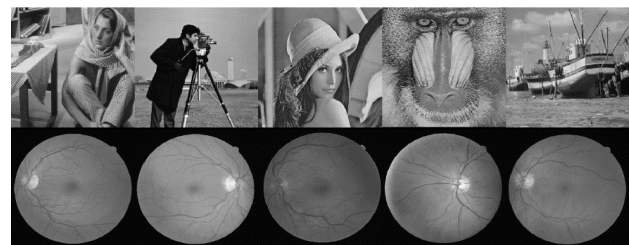


Fig. 4. Images used in the application

TABLE I
DENOISING PERFORMANCE RESULTS OF FUNDUS IMAGES

Type of Noise	Noise Ratio	Evaluation Criteria	CT	TICT
Random	Sigma=5	PSNR	40.1699	40.1634
		MSE	6.2531	6.2624
		MSSIM	0.1773	0.0605
		FSIM	0.8268	0.9143
	Sigma=10	PSNR	34.1521	34.1429
		MSE	24.9962	25.0492
		MSSIM	0.0914	0.0132
		FSIM	0.8061	0.9014
	Sigma=15	PSNR	30.6288	30.6219
		MSE	56.2608	56.3497
		MSSIM	0.0573	0.0048
		FSIM	0.7963	0.8964
Gaussian	Snr=3	PSNR	23.5333	27.4376
		MSE	301.4683	120.1661
		MSSIM	0.7279	0.6507
		FSIM	0.8684	0.9735
	Snr=5	PSNR	23.9313	29.4056
		MSE	275.8284	76.3219
		MSSIM	0.7513	0.7233
		FSIM	0.8694	0.9769
	Snr=10	PSNR	24.3999	34.0896
		MSE	248.3114	25.7974
		MSSIM	0.7969	0.8520
		FSIM	0.8696	0.9819
Rician	Snr=3	PSNR	37.4514	37.4403
		MSE	11.6957	11.7256
		MSSIM	0.1514	0.0235
		FSIM	0.7878	0.9044
	Snr=5	PSNR	32.6911	32.6777
		MSE	34.9946	35.1029
		MSSIM	0.0856	0.0065
		FSIM	0.7863	0.9013
	Snr=10	PSNR	26.4217	26.4145
		MSE	148.2254	148.4694
		MSSIM	0.0327	0.0010
		FSIM	0.7820	0.8945

V. CONCLUSION

In this study, CT which is an important MRA method, and TICT image denoising performance were compared. Both methods have been used to remove the noise of different types and different rates added to the benchmark and fundus images. If comparison is made in terms of noise types, although there was no significant difference between two transforms in Random and Rician noises, TICT was performed much better performance than CT in Gaussian noise. If comparison is made in terms of image types, fundus images were better denoised, especially in Gaussian noise. In random noise, the results from fundus images are partially better. However, in Rician noise, benchmark images have also been partially denoise better.

The main aim of this study is to compare TICT and CT techniques in terms of different image types and different noise types. In this way, it is emphasized that a transform method should be selected depending on the noise and image type.

In this study, CT and TICT were preferred because they are very successful methods. The same operations can be carried out with bandelet, tetrolet, brushlets, wedgelets, etc. transforms.

This study showed that denoising performances change depending on the image. Therefore, in the future studies, the dataset with more image types will be used and thus suitable transforms for different image types will be determined. For this, different transform methods will be discussed.

TABLE II
DENOISING PERFORMANCE RESULTS OF BENCHMARK IMAGES

Type of Noise	Noise Ratio	Evaluation Criteria	CT	TICT
Random	Sigma=5	PSNR	40.1491	40.1389
		MSE	6.2830	6.2979
		MSSIM	0.0996	0.0489
		FSIM	0.0964	0.1117
	Sigma=10	PSNR	34.1500	34.1369
		MSE	25.0080	25.0836
		MSSIM	0.0542	0.0150
		FSIM	0.0958	0.1104
	Sigma=15	PSNR	30.6139	30.6105
		MSE	56.4526	56.4974
		MSSIM	0.0351	0.0081
		FSIM	0.0949	0.1093
Gaussian	Snr=3	PSNR	15.4055	20.6055
		MSE	2575.5545	1779.0742
		MSSIM	0.3203	0.4806
		FSIM	0.0980	0.1181
	Snr=5	PSNR	15.4231	21.4481
		MSE	2570.4336	1713.6258
		MSSIM	0.3205	0.5145
		FSIM	0.0979	0.1186
	Snr=10	PSNR	15.4399	23.5224
		MSE	2565.4484	1600.8019
		MSSIM	0.3176	0.5807
		FSIM	0.0979	0.1197
Rician	Snr=3	PSNR	37.6979	37.6976
		MSE	11.0487	11.0495
		MSSIM	0.0778	0.0153
		FSIM	0.0942	0.1103
	Snr=5	PSNR	32.8443	32.8279
		MSE	33.7799	33.9076
		MSSIM	0.0439	0.0055
		FSIM	0.0940	0.1102
	Snr=10	PSNR	26.5021	26.4918
		MSE	145.5028	145.8472
		MSSIM	0.0181	0.0003
		FSIM	0.0935	0.1096

REFERENCES

- [1] K. Panetta, S. Aghaian, J.-C. Pinoli, and Y. Zhou, Image processing algorithms and measures for the analysis of biomedical imaging systems applications, *International journal of biomedical imaging*, vol. 2015, 2015.
- [2] K. Leung, A. Cunha, A. W. Toga, and D. S. Parker, Developing image processing meta-algorithms with data mining of multiple metrics, *Computational and mathematical methods in medicine*, vol. 2014, 2014.
- [3] E. Abele, L. Holland, and A. Nehrbass, Image acquisition and image processing algorithms for movement analysis of bearing cages, *Journal of Tribology*, vol. 138, no. 2, p. 021105, 2016.
- [4] S. Ya-Lin and B. Chen-Xi, Research and analysis of image processing technologies based on dotnet framework, *Physics Procedia*, vol. 25, pp. 2131-2137, 2012.
- [5] R. Patil, Noise reduction using wavelet transform and singular vector decomposition, *Procedia Computer Science*, vol. 54, pp. 849-853, 2015.
- [6] A. Boyat and B. K. Joshi, Image denoising using wavelet transform and median filtering, in *Engineering (NUiCONE), 2013 Nirma University International Conference on*, 2013, pp. 1-6: IEEE.
- [7] R. Sivakumar, G. Balaji, R. S. J. Ravikiran, R. Karikalan, and S. S. Janaki, Image Denoising using Contourlet Transform, in *2009 Second International Conference on Computer and Electrical Engineering*, 2009, vol. 1, pp. 22-25: IEEE.
- [8] D. Bhonsle, V. Chandra, and G. Sinha, Medical image denoising using bilateral filter, *International Journal of Image, Graphics and Signal Processing*, vol. 4, no. 6, p. 36, 2012.
- [9] A. B. Hamza and H. Krim, Image denoising: A nonlinear robust statistical approach, *IEEE transactions on signal processing*, vol. 49, no. 12, pp. 3045-3054, 2001.
- [10] A. Dixit and P. Sharma, A Comparative Study of Wavelet Thresholding for Image Denoising, *IJ Image, Graphics and Signal Processing*, vol. 12, pp. 39-46, 2014.

- [11] N. Mehala and R. Dahiya, A comparative study of FFT, STFT and wavelet techniques for induction machine fault diagnostic analysis, in *Proceedings of the 7th WSEAS international conference on computational intelligence, man-machine systems and cybernetics, Cairo, Egypt, 2008*, vol. 2931.
- [12] T. Bernas, E. K. Asem, J. P. Robinson, and B. Rajwa, Compression of fluorescence microscopy images based on the signal-to-noise estimation, *Microscopy research and technique*, vol. 69, no. 1, pp. 1-9, 2006.
- [13] R. M. Willett and R. D. Nowak, Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging, *IEEE Transactions on Medical Imaging*, vol. 22, no. 3, pp. 332-350, 2003.
- [14] J. B. De Monvel, S. Le Calvez, and M. Ulfendahl, Image restoration for confocal microscopy: improving the limits of deconvolution, with application to the visualization of the mammalian hearing organ, *Biophysical Journal*, vol. 80, no. 5, pp. 2455-2470, 2001.
- [15] P. V. Lavanya, C. V. Narasimhulu, and K. S. Prasad, Transformations analysis for image denoising using complex wavelet transform, in *Innovations in Information, Embedded and Communication Systems (ICIECS), 2017 International Conference on*, 2017, pp. 1-7: IEEE.
- [16] P. Rakheja and R. Vig, Image Denoising using Various Wavelet Transforms: A Survey, *Indian Journal of Science and Technology*, vol. 9, no. 48, 2017.
- [17] M. N. Do and M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, *IEEE Transactions on image processing*, vol. 14, no. 12, pp. 2091-2106, 2005.
- [18] P. Hiremath, P. T. Akkasaligar, and S. Badiger, Performance comparison of wavelet transform and contourlet transform based methods for despeckling medical ultrasound images, *International Journal of Computer Applications*, vol. 26, no. 9, pp. 34-41, 2011.
- [19] E. J. Candès and D. L. Donoho, Ridgelets: A key to higher-dimensional intermittency?, *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, pp. 2495-2509, 1999.
- [20] E. J. Candès and D. L. Donoho, Curvelets: A surprisingly effective nonadaptive representation for objects with edges, Stanford Univ Ca Dept of Statistics 2000.
- [21] J. Xu, L. Yang, and D. Wu, Ripplet: A new transform for image processing, *Journal of Visual Communication and Image Representation*, vol. 21, no. 7, pp. 627-639, 2010.
- [22] J. Krommweh, Tetrolet transform: A new adaptive Haar wavelet algorithm for sparse image representation, *Journal of Visual Communication and Image Representation*, vol. 21, no. 4, pp. 364-374, 2010.
- [23] M. Ceylan and A. E. Canbilin, Performance Comparison of Tetrolet Transform and Wavelet-Based Transforms for Medical Image Denoising, *International Journal of Intelligent Systems and Applications in Engineering*, vol. 5, no. 4, pp. 222-231, 2017.
- [24] M. Saha, M. K. Naskar, and B. Chatterji, Wavelet and Curvelet Transforms for Biomedical Image Processing, in *Handbook of Research on Information Security in Biomedical Signal Processing*: IGI Global, 2018, pp. 95-129.
- [25] P. Jain and V. Tyagi, An adaptive edge-preserving image denoising technique using tetrolet transforms, *The Visual Computer*, vol. 31, no. 5, pp. 657-674, 2015.
- [26] Q. Huang, B. Hao, and S. Chang, Adaptive digital ridgelet transform and its application in image denoising, *Digital Signal Processing*, vol. 52, pp. 45-54, 2016.
- [27] R. Eslami and H. Radha, Translation-invariant contourlet transform and its application to image denoising, *IEEE Transactions on image processing*, vol. 15, no. 11, pp. 3362-3374, 2006.
- [28] M. Bouhrara *et al.*, Incorporation of Rician noise in the analysis of biexponential transverse relaxation in cartilage using a multiple gradient echo sequence at 3 and 7 Tesla, *Magnetic resonance in medicine*, vol. 73, no. 1, pp. 352-366, 2015.
- [29] P. J. Burt and E. H. Adelson, The Laplacian pyramid as a compact image code, in *Readings in Computer Vision*: Elsevier, 1987, pp. 671-679.
- [30] R. H. Bamberg and M. J. Smith, A filter bank for the directional decomposition of images: Theory and design, *IEEE transactions on signal processing*, vol. 40, no. 4, pp. 882-893, 1992.
- [31] P. Han and J. Du, Spatial images feature extraction based on bayesian nonlocal means filter and improved contourlet transform, *Journal of Applied Mathematics*, vol. 2012, 2012.
- [32] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, Ridge-based vessel segmentation in color images of the retina, *IEEE transactions on medical imaging*, vol. 23, no. 4, pp. 501-509, 2004.
- [33] V. N. P. Raj and T. Venkateswarlu, Denoising of Poisson and Rician Noise from Medical Images using Variance Stabilization and Multiscale Transforms, *International Journal of Computer Applications*, vol. 57, no. 21, 2012.
- [34] Z. Wang, Z. Li, W. Lin, and C. Liu, Image quality assessment based on improved feature similarity metric, *work*, vol. 22, p. 24, 2011.

BIOGRAPHIES



MUHAMMET FATIH ASLAN is a research assistant at Karamanoglu Mehmetbey University (KMU), Karaman, Turkey. After completing his BSc with a high degree in Selçuk University (SU), Konya, Turkey in 2016, he started to work in Karamanoglu Mehmetbey University in 2017. In 2018, he completed his master's degree at Selçuk University. He is currently a PhD student in Electrical and Electronic Engineering at Konya Technical University. His research interests include robotics, image processing, machine learning and object tracking.



KADIR SABANCI was born in 1978. He received his B.S. and M.S. degrees in Electrical and Electronics Engineering (EEE) from Selçuk University, Turkey, in 2001, 2005 respectively. In 2013, he received his Ph.D. degree in Agricultural Machineries from Selçuk University, Turkey. He has been working as Assistant Professor in the Department of EEE at Karamanoglu Mehmetbey University. His current research interests include image processing, data mining, artificial intelligent, embedded systems and precision agricultural technology.



AKIF DURDU has been an assistant professor with the Electrical-Electronics Engineering Department at the Konya Technical University since 2013. He earned a PhD degree in Electrical-Electronics engineering from the Middle East Technical University (METU), in 2012. He received his B.Sc. degree in Electrical-Electronics Engineering in 2001 at the Selçuk University. His research interests include mechatronic design, search & rescue robotics, robot manipulators, human-robot interaction, multi-robots networks and sensor networks. Dr. Durdu is teaching courses in control engineering, robotic and mechatronic systems.

An alignment-free method for bulk comparison of protein sequences from different species

B. DOĞAN


Abstract—The available number of protein sequences rapidly increased with the development of new sequencing techniques. This in turn led to an urgent need for the development of new computational methods utilizing these data for the solution of different biological problems. One of these problems is the comparison of protein sequences from different species to reveal their evolutionary relationship. Recently, several alignment-free methods have been proposed for this purpose. Here in this study, we also proposed an alignment-free method for the same purpose. Different from the existing methods, the proposed method not only allows for a pairwise comparison of two protein sequences, but also it allows for a bulk comparison of multiple protein sequences simultaneously. Computational results performed on gold-standard datasets showed that, bulk comparison of multiple sequences is much faster than its pairwise counterpart and the proposed method achieves a performance which is quite competitive with the state-of-the-art alignment-based method, ClustalW.

Index Terms— ClustalW, ND5 proteins, Phylogenetic analysis, Protein sequence similarity, Sequence comparison.

I. INTRODUCTION

DEVELOPMENT OF next-generation sequencing technologies led to a dramatic increase in the number of available DNA and protein sequences. It is quite crucial to effectively extract the biological information provided with this huge number of sequences which has given rise to new research challenges in computational biology and bioinformatics. Similarity analysis of protein sequences from different species is one of these research challenges. With the development of new tools for protein sequence similarity analysis, scientist will be able to elucidate the function of unknown proteins which may shed light on identification of potential drug targets and to gain insights on underlying molecular mechanisms of diseases [1].

BERAT DOĞAN, is with Department of Biomedical Engineering, Inonu University, Malatya, Turkey. (e-mail: berat.dogan@inonu.edu.tr)

 <https://orcid.org/0000-0003-4810-1970>

Manuscript received March 16, 2019; accepted Sep 23, 2019.
DOI: [10.17694/bajece.540873](https://doi.org/10.17694/bajece.540873)

In literature, methods proposed for the sequence similarity analysis are usually investigated under two different groups: alignment-based methods and alignment-free methods. In alignment-based methods, a score function is used to represent insertion, deletion, and substitution of nucleotides or amino acids in the compared biological sequences. The overall objective of these methods is to align the sequences with the highest scores [2-7]. Although, alignment-based methods are successfully utilized for sequence similarity analysis, they are generally time consuming and memory demanding. Therefore, alignment-free methods are proposed as computationally inexpensive alternatives to alignment-based methods.

On the other hand, the main difficulty of alignment-free methods is the need for an effective method to map a protein sequence into a numerical format (either a vector or a matrix) that could be used in subsequent analyses. In literature, physicochemical properties of amino acids are usually utilized for this purpose. In [8], based on two physicochemical properties of amino acids, authors first converted a protein sequence into a three-letter sequence. Then, based on this three-letter sequence, they obtained a graph without loops and multiple edges and its geometric line adjacency matrix. Next, to characterize a protein sequence numerically, they constructed a generalized PseAAC (pseudo amino acid composition) model. By using the obtained numerical representation of protein sequences, similarity analysis among β -globin proteins of seventeen species and seventy-two spike proteins of coronaviruses were made. They showed that the resulting clusters agreed well with the established taxonomic groups. In [9], authors proposed a novel position-feature-based model for protein sequences by employing physicochemical properties of amino-acids and the measure of graph energy. The proposed method puts emphasis on sequence order information and describes local dynamic distributions of sequences, from which one can get characteristic B-vector. Afterwards, they applied the relative entropy to the sequences representing B-vectors to measure their similarity. They showed that the proposed method competes with the widely utilized alignment-based method, ClustalW [10]. In [11], authors proposed a method to analyze the similarity of proteins by Discrete Fourier Transform (DFT) and Dynamic Time Warping (DTW). They first converted the protein sequences into numerical sequences according to their physicochemical properties. Next, they obtained the power spectra of sequences from DFT and extended the spectra to the

same length to calculate the distance between different sequences by DTW. They tested their method on different datasets and the results were compared with the existing methods. They showed that the proposed method overperformed the existing methods. In [12], based on the three important physicochemical properties of amino acids: the hydrophathy index, polar requirement and chemical composition of the side chain, authors proposed a 24-dimensional feature vector describing the composition of amino acids in protein sequences. The results on beta-globin, mammals, and three virus datasets showed that the proposed method is fast and accurate for classifying proteins and inferring the phylogeny of organisms. In [13], based on eight physicochemical properties of amino acids, authors proposed a 40-dimensional feature vector for numerical characterization of each protein sequence. They used the cosine distances of feature vectors to measure the similarity of proteins. Analysis results performed over two real datasets demonstrated that the new scheme is effective in similarity research and phylogenetic analysis. In [14], based on three physicochemical properties of amino acids, authors reduced a protein primary sequence into a six-letter sequence, and then they extracted a set of elements which reflect the global and local sequence-order information. Combining these elements with the frequencies of 20 native amino acids, they obtained a numerical vector to characterize each protein sequence. The utility of the proposed approach was illustrated by phylogenetic analysis and identification of DNA-binding proteins. In [15], based on two physicochemical properties of amino acids, a 2D graphical representation of protein sequences is proposed. The proposed graphical representation of proteins is then used to obtain a numerical vector for each protein which is then used to compute the similarity of different proteins. Experiments performed on ND5 proteins of nine species showed that their method is simple and effective. In [16], based on 12 major physicochemical properties of amino acids and the principal component analysis (PCA) method, authors proposed a simple and intuitive 2D graphical mapping method for protein sequences. Next, they extracted a 20D vector from the graphical mapping which is used to characterize a protein sequence. To validate the proposed method, they first gave a comparison of protein sequences, which consists of nine ND6 proteins. They showed that the similarity/dissimilarity matrix for nine ND6 proteins correctly reveals their evolutionary relationship. Next, they performed the cluster analysis of HA genes of influenza A (H1N1) isolates, results of which is shown to be consistent with the known evolution fact of the H1N1 virus. In [17], authors proposed a new protein map which is based on ten physicochemical properties of amino acids. The proposed method both considers phylogenetic factors arising from amino acid mutations and provides computational efficiency for the huge amount of data. Authors showed that the proposed model is easier and quicker in handling protein sequences than multiple alignment methods and gives protein classification greater evolutionary significance at the amino

acid sequence level. In [18], author presented a 2D spectrum-like graphical representation of protein sequences based on the hydrophobicity scale of amino acids. The frequencies of amplitudes of 4-subsequences were adopted to characterize a spectrum-like graph, and a 17D vector was used for numerical representation of a protein sequence. By using protein sequences from the mitochondrion genome of 20 different species, they compared their method to ClustalW to show the utility of their method. In [19], by using nine physicochemical properties of amino acids and PCA, authors proposed a novel graphical representation of protein sequences called ADLD (Alignment Diagonal Line Diagram). Experiments performed on 16 different ND5 proteins and the 29 different spike proteins showed that their method is not only visual, intuitional, and effective in the similarity/dissimilarity analysis of protein sequences but also quite simple, since there are no high dimensional matrices required to be constructed. In some other studies different from the physicochemical properties of amino acids, authors also used codon information [20-21], pseudo-Markov transition probability vector among the 20 amino acids [22], the distributions of each kind of adjacent amino acid (AAA) within sequence [23], K-string dictionary [24], and a set of point masses representing a sequence in a 20D space [25] for alignment-free similarity analysis of protein sequences.

Each of the above studies may have several advantages when compared to one another. However, to the best of our knowledge, none of the above studies allow for a bulk comparison of protein sequences for alignment-free sequence similarity analysis. Here, different from the above studies, in this paper we propose a method which allows for not only pairwise comparison of protein sequences but also a bulk comparison of several protein sequences for similarity analysis. Experiments showed that, bulk comparison of protein sequences with the proposed approach is much faster than the pairwise comparison of the same proteins using the same similarity measure. Moreover, there is no degradation on the performance in terms of the accuracy of the obtained results. It is also shown that, the proposed method achieves a clustering performance which is comparable to the state-of-the-art ClustalW method.

The rest of this paper is organized as follows: in section 2, details of the proposed method is introduced. Section 3 covers a definition of the datasets used in experiments along with the computational results obtained with the proposed and the existing methods. Finally, section 4 concludes the study.

II. METHODOLOGY

In literature, for comparison of different methods a gold-standard dataset which includes sequences of ND5 proteins from nine different species is widely utilized. In this study, we will also use this dataset along with the others for comparison of our method with the existing methods. However, let us first continue with the description of our method by using this gold-standard dataset which we believe will be more informative to the readers.

TABLE I
 TWELVE MAJOR PHYSICOCHEMICAL CHARACTERISTIC VALUES OF 20 AMINO ACIDS. (P1: CHEMICAL COMPOSITION OF THE SIDE CHAIN; P2: POLAR REQUIREMENT; P3: HYDROPATHY INDEX; P4: ISOELECTRIC POINT; P5: MOLECULAR VOLUME; P6: POLARITY; P7: AROMATICITY; P8: ALIPHATICITY; P9: HYDROGENATION; P10: HYDROXYTHIOLATION; P11: PK1(-COOH); P12: PK2(-NH3+))

Amino acids	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
A (Ala)	0	7.0	1.8	6.00	31	8.1	-0.11	0.239	0.33	-0.062	2.34	9.69
C (Cys)	2.75	4.8	2.5	5.07	55	5.5	-0.184	0.22	0.074	0.38	1.71	10.78
D (Asp)	1.38	13.0	-3.5	2.77	54	13.0	-0.285	0.171	-0.371	-0.079	2.09	9.82
E (Glu)	0.92	12.5	-3.5	3.22	83	12.3	-0.067	0.187	-0.254	-0.184	2.19	9.67
F (Phe)	0	5.0	2.8	5.48	132	5.2	0.438	0.234	0.011	0.074	1.83	9.13
G (Gly)	0.74	7.9	-0.4	5.97	3	9.0	-0.073	0.16	0.37	-0.017	2.34	9.60
H (His)	0.58	8.4	-3.2	7.59	96	10.4	0.32	0.205	-0.078	0.056	1.82	9.17
I (Ile)	0	4.9	4.5	6.02	111	5.2	0.001	0.273	0.149	-0.309	2.36	9.68
K (Lys)	0.33	10.1	-3.9	9.74	119	11.3	0.049	0.228	-0.075	-0.371	2.18	8.95
L (Leu)	0	4.9	3.8	5.98	111	4.9	-0.008	0.281	0.129	-0.264	2.36	9.60
M (Met)	0	5.3	1.9	5.74	105	5.7	-0.041	0.253	-0.092	0.077	2.28	9.21
N (Asn)	1.33	10.0	-3.5	5.41	56	11.6	-0.136	0.249	-0.233	0.166	2.02	8.80
P (Pro)	0.39	6.6	-1.6	6.30	32.5	8.0	-0.016	0.165	0.37	-0.036	1.99	10.60
Q (Gln)	0.89	8.6	-3.5	5.65	85	10.5	-0.246	0.26	-0.409	-0.025	2.17	9.13
R (Arg)	0.65	9.1	-4.5	10.76	124	10.5	0.079	0.211	-0.176	-0.167	2.17	9.04
S (Ser)	1.42	7.5	-0.8	5.68	32	9.2	-0.153	0.236	0.022	0.47	2.21	9.15
T (Thr)	0.71	6.6	-0.7	6.16	61	8.6	-0.208	0.213	0.136	0.348	2.63	10.43
V (Val)	0	5.6	4.2	5.96	84	5.9	-0.155	0.255	0.245	0.212	2.32	9.62
W (Trp)	0.13	5.2	-0.9	5.89	170	5.4	0.493	0.183	0.011	0.05	2.38	9.39
Y (Tyr)	0.20	5.4	-1.3	5.66	136	6.2	0.381	0.193	-0.138	0.22	2.20	9.11

The ND5 dataset (Table A1 of appendix) consist of sequences from nine different species; human, gorilla, pigmy chimpanzee (pchimp), common chimpanzee (chimp), fin whale (fwhale), blue whale (bwhale), mouse, rat and opossum.

several masks representing the sequences from S_{ND5} . For the S_{ND5} , the set of masks are found in the following manner.

Let $S_C = S_1 \circ S_2 \circ \dots \circ S_9$ be a sequence obtained from the concatenation of the sequences from S_{ND5} , where \circ represents the concatenation operator. The obtained S_C includes the necessary information to form the required mask set. To create the set of masks, one should first find the all 2-mers (or dimers) exist in the sequence S_C along with their frequencies. Here, considering the frequency of each 2-mer, we select a threshold to extract the required set of masks from the set of all 2-mers found. The final mask set should ideally represent each sequence from the ND5 dataset. Therefore, selection of the threshold is of critical importance. As shown in Fig.1, the threshold is selected as the median of the frequencies found for each 2-mer. 2-mers those have a frequency below the defined threshold are selected to form the mask set. For the ND5 dataset, the threshold is found to be 9 and the total number of 2-mers in the mask set M is found to be 148, some of which are explicitly shown in Fig.1.

Having formed the required set of masks, we need to define a measure to compute the affinity of sequence $S_i \in S_{ND5}$, $i = 1, 2, \dots, 9$ to mask $m_j \in M$, $j = 1, 2, \dots, 148$. For this purpose, we utilized the twelve physicochemical properties of amino acids listed in Table 1 [16]. By using the PCA, Table 1 is first projected into a two-dimensional space (Table 2) in which each amino acid could be represented as a two-dimensional vector. A vectoral representation of each amino acid is shown in Fig.2A. Now, a given 2-mer, i.e., "CS", could be represented by concatenating two vectors head-to-tail. As shown in Fig.2B, once we concatenate the two vectors, we will have three different points in two-dimensional space. Because the point $p_0 = (0,0)$ is common for the all 2-

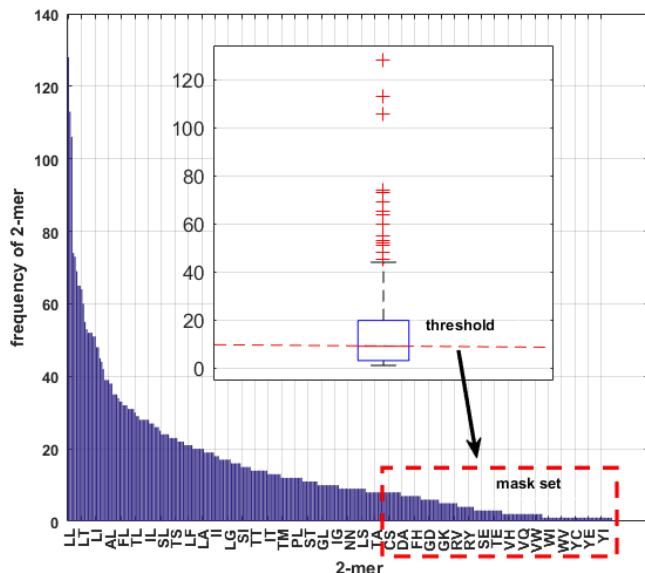


Fig.1. Frequency distribution of 2-mers obtained from the concatenated sequence S_C . 2-mers those have a frequency value below the threshold (median of the distribution) are selected to form the required set of masks.

Let $S = s_1 s_2 \dots s_n$ be a protein sequence, where n represents the sequence length. Then, we could formally describe a sequence in the ND5 dataset as $S_i \in S_{ND5}$, $i = 1, 2, \dots, 9$. As mentioned before, the proposed method aims for a bulk comparison of the sequences. This is achieved by using

mers, we could simply ignore it. Thus, there remains two different points (p_1, p_2) to represent a 2-mer in two-dimensional space. These two points are used to form a four-dimensional feature vector $\vec{f} = [p_1, p_2]$ for a given 2-mer. Considering the 2-mer "CS" the resulting feature vector is $\vec{f}_{CS} = [-28.93, 5.51, -80.96, 5.87]$ as shown in Fig.2B. Each mask similarly could easily be represented as four-dimensional feature vectors $\vec{f}_{m_j} \in R^4, j = 1, 2, \dots, 148$.

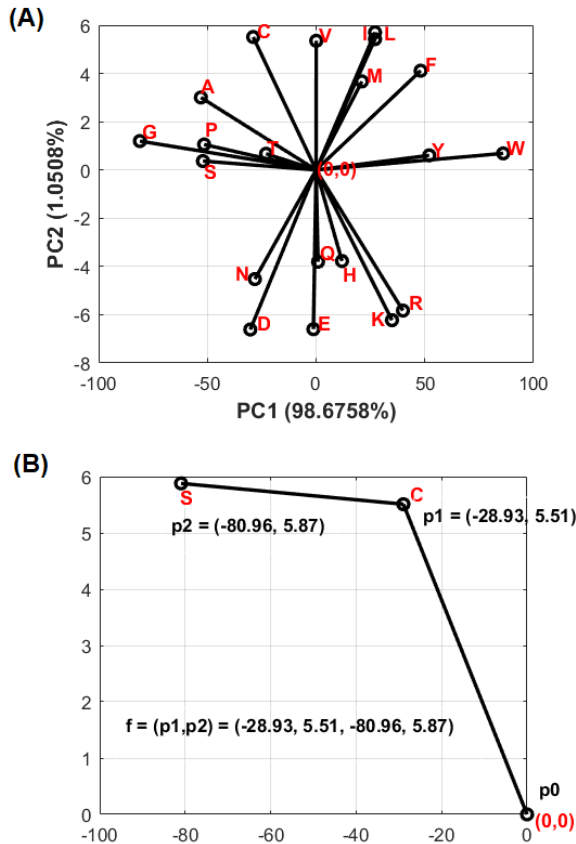


Fig.2. (A) Physicochemical properties of amino acids given in Table 1 are projected into two-dimensional space with the help of PCA. (B) A given 2-mer i.e. "CS" could be represented by a feature vector f which is formed by two points (p_1, p_2) obtained after concatenating vectors for amino acids "C" and "S" head-to-tail.

To compute the affinity of a sequence to a mask, we used a sliding window approach. In this approach, a 2-length window is slid through the sequence $S_i, i = 1, 2, \dots, 9$ and the affinity of a sequence to a mask is computed in the following manner. First, sequences of each windowed 2-mer $S_{W_1}^i, S_{W_2}^i, \dots, S_{W_n}^i$ is found. Here, n represents the total number of windowed 2-mers obtained from S_i after sliding operation. Next, corresponding feature vectors of each windowed 2-mer are found as shown in Fig.2B. Feature vectors for each windowed 2-mer is represented as $\vec{f}_{W_1}^i, \vec{f}_{W_2}^i, \dots, \vec{f}_{W_n}^i$. Euclidian distances between each feature vector of the windowed 2-mers and the feature vector of a mask m_j , which is \vec{f}_{m_j} , are then computed as in Equation (1).

$$d_k^i = \text{dist}(\vec{f}_{W_k}^i, \vec{f}_{m_j}), \quad (1)$$

$$i = 1, 2, \dots, 9, \quad k = 1, 2, \dots, n, \quad j = 1, 2, \dots, 148$$

$\text{dist}(\vec{f}_{W_k}^i, \vec{f}_{m_j})$ represents the Euclidian distance from $\vec{f}_{W_k}^i$ to \vec{f}_{m_j} . Euclidian distances obtained between the windowed 2-mers of sequence S_i and mask pairs are then stored in a vector $\vec{d}_i = (d_1^i, d_2^i, \dots, d_n^i)$. Affinity of a sequence S_i to mask m_j is then computed as in Equation (2).

$$a_{ij} = \min(\vec{d}_i) \quad (2)$$

Computed affinity values $a_{ij}, i = 1, 2, \dots, 9, j = 1, 2, \dots, 148$ could be represented as an $N \times M$ (for ND5 dataset $N = 9$ and $M = 148$) affinity matrix A . This matrix is then used to compute the similarity between sequences.

TABLE II
FIRST TWO PRINCIPAL COMPONENTS OF 20 AMINO ACIDS
OBTAINED AFTER PCA PROJECTION OF 12 MAJOR
PHYSICO-CHEMICAL CHARACTERISTIC VALUES GIVEN IN TABLE I

Amino acids	PC1 (98.68%)	PC2 (1.05%)
A (Ala)	-52.9779	3.0042
C (Cys)	-28.9349	5.5100
D (Asp)	-30.2566	-6.6135
E (Glu)	-1.2373	-6.6004
F (Phe)	48.0692	4.1135
G (Gly)	-81.0102	1.1981
H (His)	11.9222	-3.7753
I (Ile)	27.0888	5.7064
K (Lys)	34.8912	-6.2279
L (Leu)	27.0930	5.4234
M (Met)	21.0652	3.6697
N (Asn)	-28.1420	-4.5234
P (Pro)	-51.4851	1.0570
Q (Gln)	0.8943	-3.8033
R (Arg)	39.9295	-5.8285
S (Ser)	-52.0304	0.3699
T (Thr)	-23.0132	0.6717
V (Val)	0.0729	5.3547
W (Trp)	86.0342	0.6879
Y (Tyr)	52.0269	0.6058

III. COMPUTATIONAL RESULTS

A. Datasets used in the experiments

In this study, along with the ND5 dataset (Table A1 of appendix), three other datasets are used to evaluate the performance of the proposed method. The second dataset consists of thirty-five Coronavirus Spike Proteins which were derived from the NCBI. The information and accession numbers [20] of proteins are listed in Table A2 of appendix. The third dataset consists of twenty-four transferrin and lactoferrin proteins from fish, amphibians and mammals of twenty-four vertebrates. Taxonomic information and accession numbers of these proteins are provided in Table A3 of appendix [20]. The fourth dataset consists of twenty-seven antifreeze protein sequences (AFPs) from spruce budworm (*Choristoneura fumiferana*, CF), yellow mealworm (*Tenebrio*

molitor, TM), Hypogastrura harveyi (HH), Dorcus curvidens binodulosus (DCB), Microdera dzhungarica punctipennis (MDP) and Dendroides canadensis (DC) for which the detailed information could be found in [26].

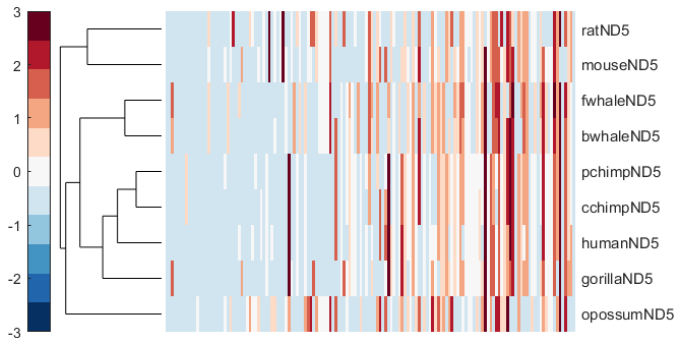


Fig.3. ND5 dataset clustered with the proposed method correctly clusters proteins into their correct groups.

B. Similarity analysis of ND5 proteins

Let A_{ND5} be the affinity matrix found for the ND5 proteins. From the previous section we know that A_{ND5} is a $N \times M = 9 \times 148$ matrix where N represents the number of sequences and M represents the total number of masks. A_{ND5} could directly be used to cluster the sequences for which the clustering result is shown in Fig.3. From Fig.3 and Fig.4, it is shown that, clusters found by the proposed method is in good agreement with the clusters found by the ClustalW.

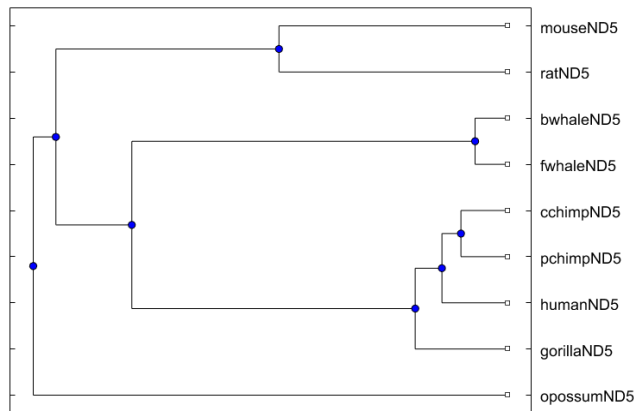


Fig.4. ClustalW result for the ND5 dataset.

To measure the performance of the proposed method, we also provide the correlation coefficients calculated between the distance matrix found by the proposed method and the one found by the ClustalW. The distance matrix for the proposed method is found as in Equation (3).

$$D_{ij} = \text{dist}(A_{ND5}^i, A_{ND5}^j) \quad (3)$$

$$i = 1, 2, \dots, 9, \quad j = 1, 2, \dots, 9$$

The calculated distance matrix D is shown in Table 3 and the distance matrix found by the ClustalW is provided in Table 4 [20]. In Table 5, Pearson's correlation coefficients calculated between Table 3 and Table 4 are provided along with some other previously published methods. From Table 5, it can be shown that the proposed method is quite competitive and provides relatively higher correlation coefficients when compared to other methods.

C. Pairwise comparison of protein sequences with the proposed method

Although the proposed method mainly proposed for bulk comparison of several protein sequences, it also allows for a pairwise comparison of two protein sequences. To achieve this, all we need to do is to concatenate two sequences $S_C = S_1 \circ S_2$ that we want to compare. Once we obtained the concatenated sequence S_C , the same procedure given in section 2 is followed. Thereby, one could compute the similarity of two proteins by simply calculating the distances between the affinity vectors \vec{a}_1 and \vec{a}_2 found for each sequence.

In Fig.5, correlation coefficients obtained by the pairwise comparison are compared to those obtained with the bulk comparison method. From Fig.5, it is shown that both the pairwise and bulk comparisons with the proposed method provide quite similar results. However, bulk comparison of the proteins from ND5 dataset is much faster (21.69 sec.) when compared to the pairwise comparison of all sequences (305.64 sec.) from the same dataset.

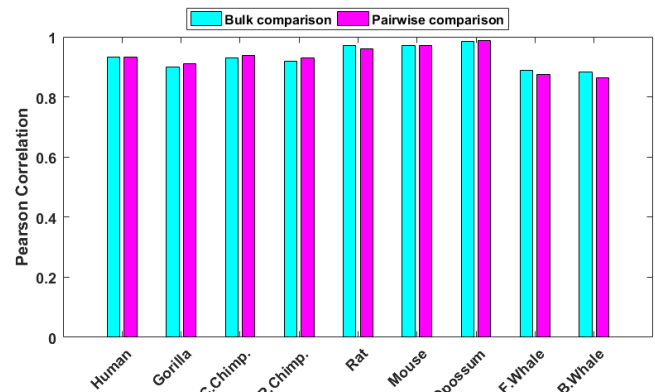


Fig.5. Bulk comparison and pairwise comparison of the sequences from ND5 dataset with the proposed method provides similar results when they are compared to ClustalW.

D. Similarity analysis of coronavirus spike proteins

Coronaviruses are important human and animal pathogens which are associated with upper respiratory tract infections in adults and probably also play a role in severe respiratory infections in both children and adults. Coronaviruses are traditionally classified into three groups, where the first group and the second group come from mammalian, and the third group comes from poultry (chicken and turkey). Apart from these three groups, SARS-CoV forming a fourth group, for which the phylogenetic position and origins remain a matter of some controversy. In this study, we utilized the proposed method to classify SARS-CoV spike proteins along with the proteins from other three groups.

In Fig.6, clustering result with the proposed method is shown. From this figure, it is shown that all the four groups of proteins successfully separated by the proposed method. The SARS-CoV spike proteins (group IV) remain closer to the group II. Previous studies [32, 33] also showed that, the closest relatives of SARS-CoVs are the murine, bovine and human coronaviruses from group II which is consistent with the obtained results. On the other hand, ClustalW also correctly separates the proteins into their correct groups (Fig.7).

TABLE III
DISTANCE MATRIX FOUND BY THE PROPOSED METHOD

	Human	Gorilla	C.Chimp.	P.Chimp.	Rat	Mouse	Opossum	F.Whale	B.Whale
Human	0	40.42	29.95	34.30	79.38	71.52	63.52	59.04	59.19
Gorilla		0	44.23	41.94	74.88	72.51	69.43	57.81	60.47
C.Chimp.			0	18.83	78.52	69.27	65.05	58.98	54.91
P.Chimp.				0	76.50	69.23	64.91	55.15	50.78
Rat					0	53.42	71.99	71.68	70.73
Mouse						0	68.74	64.30	63.56
Opossum							0	61.87	59.86
F.Whale								0	24.80
B.Whale									0

TABLE IV
DISTANCE MATRIX FOUND BY THE CLUSTALW [20]

	Human	Gorilla	C.Chimp.	P.Chimp.	Rat	Mouse	Opossum	F.Whale	B.Whale
Human	0	0.104	0.067	0.069	0.456	0.443	0.464	0.375	0.377
Gorilla		0	0.096	0.093	0.469	0.453	0.494	0.39	0.387
C.Chimp.			0	0.048	0.461	0.448	0.472	0.37	0.37
P.Chimp.				0	0.453	0.443	0.459	0.368	0.368
Rat					0	0.241	0.494	0.41	0.407
Mouse						0	0.469	0.422	0.415
Opossum							0	0.486	0.486
F.Whale								0	0.034
B.Whale									0

TABLE V
CORRELATION COEFFICIENTS CALCULATED BETWEEN DISTANCE MATRICES OF SOME STATE-OF-THE-ART METHODS AND CLUSTALW FOR NDS PROTEINS

	This work	Wu et al.[20]	Yao et al.[18]	Ellakani and Mahran [27]	Zhang et al. [15]	Mu et al. [28]	Liu et al. [29]	Wu et al. [30]	Huang and Hu [31]
Human	0.93	0.96	0.93	-0.09	0.91	0.93	0.94	0.93	0.89
Gorilla	0.90	0.93	0.88	-0.03	0.92	0.93	0.93	0.91	0.93
C.Chimp.	0.93	0.96	0.94	-0.11	0.93	0.91	0.94	0.91	0.95
P.Chimp.	0.92	0.95	0.95	-0.11	0.91	0.93	0.93	0.76	0.91
Rat	0.97	0.96	0.95	0.72	0.92	0.93	0.84	0.63	0.93
Mouse	0.97	0.96	0.98	0.75	0.87	0.97	1.00	0.66	0.86
Opossum	0.99	0.99	0.94	0.99	0.99	0.93	0.89	0.52	0.92
F.Whale	0.89	0.85	0.91	0.16	0.92	0.93	0.89	0.53	0.92
B.Whale	0.88	0.85	0.93	0.15	0.92	0.96	0.87	0.69	0.93

E. Similarity analysis of transferrin and lactoferrin proteins

Iron is an essential element for almost all living organisms as it participates in a wide variety of metabolic processes, including oxygen transport, deoxyribonucleic acid (DNA) synthesis, and electron transport [34]. Iron is transported in the blood by transferrin (TF) proteins. Lactoferrin (LF) is also an iron binding protein which is structurally similar to transferrin. Previous studies have demonstrated the phylogenetic relation between the transferrin and lactoferrin [35, 36]. In this study, we analyzed the similarity of transferrin and lactoferrin

proteins with the proposed method. From Fig.8., it is shown that, lactoferrin and transferrin proteins mostly clustered into their corresponding groups. Some of the transferrin proteins (the ones from mammals) remain closer to the lactoferrin proteins which is also consistent with the results reported in [20]. For this dataset, ClustalW achieves a better clustering performance (Fig.9). However, in ClustalW clustering result again the TF proteins from mammals remain closer to the LF proteins which is in good agreement with the results found by the proposed method.

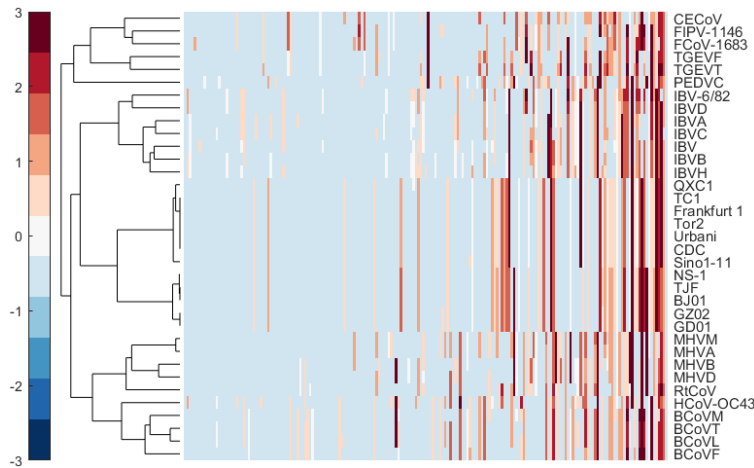


Fig.6. Coronavirus spike proteins clustered with the proposed method. All the four groups of proteins (see Table A2) successfully separated by the proposed method

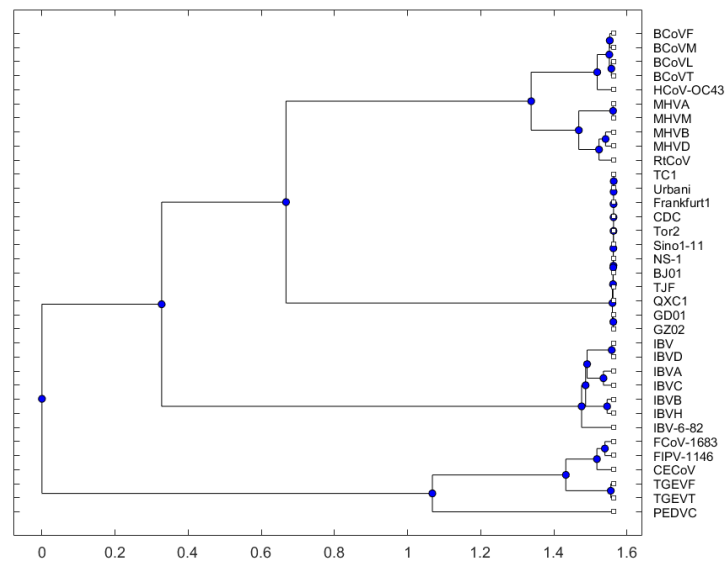


Fig.7. ClustalW clustering result for coronavirus spike proteins is in good agreement with the proposed method.

F. Similarity analysis of antifreeze proteins

Antifreeze proteins (AFPs) are biological antifreezes with unique properties, including thermal hysteresis (TH), ice recrystallization inhibition (IRI), and interaction with membranes and/or membrane proteins. These properties have been utilized in the preservation of biological samples at low temperatures [37]. These unique properties give AFPs to have potential in frozen food industry avoiding the damage in the structure of animal or vegetal foods. In this paper, by utilizing the proposed method, we analyzed the similarity of twenty-seven antifreeze proteins from spruce budworm (*Choristoneura fumiferana*, CF), yellow mealworm (*Tenebrio molitor*, TM), *Hypogastrura harveyi* (HH), *Dorcus curvidens binodulosus* (DCB), *Microdera dzhungarica punctipennis* (MDP) and *Dendroides canadensis* (DC). The proposed method (Fig.10) achieves a better clustering performance in comparison to the ClustalW. In ClustalW result (Fig.11), TM60593179 forms a separate branch which is far away from its own group. On the other hand, other proteins are clustered into their correct groups.

IV. CONCLUSION

Protein sequences similarity analysis is one of the major topics in bioinformatics. It allows researchers to find out evolutionary relationships of different species. Within this context, this paper presented a new method for the similarity analysis of proteins from different species. Different from the existing studies, the proposed method not only provides a pairwise comparison of two proteins, but it also allows for a bulk comparison of multiple proteins.

The idea behind the proposed method is quite simple and effective. To achieve a bulk comparison of multiple sequences, we used several masks which are selected from the corresponding sequences that are wanted to be compared. Simply, if a sequence A is similar to a mask C and a sequence B is again similar to mask C, then sequence A and sequence B must also be similar in some degree.

The performance of the proposed method was evaluated on four different datasets. The first dataset is the gold-standard ND5 dataset. Experiments performed on this dataset showed that the proposed method performs quite well, and the

obtained results are superior to most of the other previously published methods. The second dataset consists of thirty-five coronavirus spike proteins. Experiments performed on this dataset showed that, the proposed method successfully clusters the proteins into their correct groups and the obtained results are in good agreement with the ClustalW results. The third dataset consists of twenty-four transferrin and lactoferrin proteins. For this dataset, the results obtained by ClustalW is superior to the proposed method. However, the phylogenetic tree obtained with the proposed method is mostly in good agreement with the one obtained by the ClustalW. The last

dataset consists of twenty-seven antifreeze proteins from six different species. For this dataset, the proposed method exhibited a superior performance when compared to the ClustalW.

The concept behind the proposed method could also be utilized to cluster nucleotide sequences which is also a challenging problem in bioinformatics. Future studies will mostly cover the application of this concept in nucleotide sequence clustering problems. However, it could also be utilized in many other areas of pattern recognition.

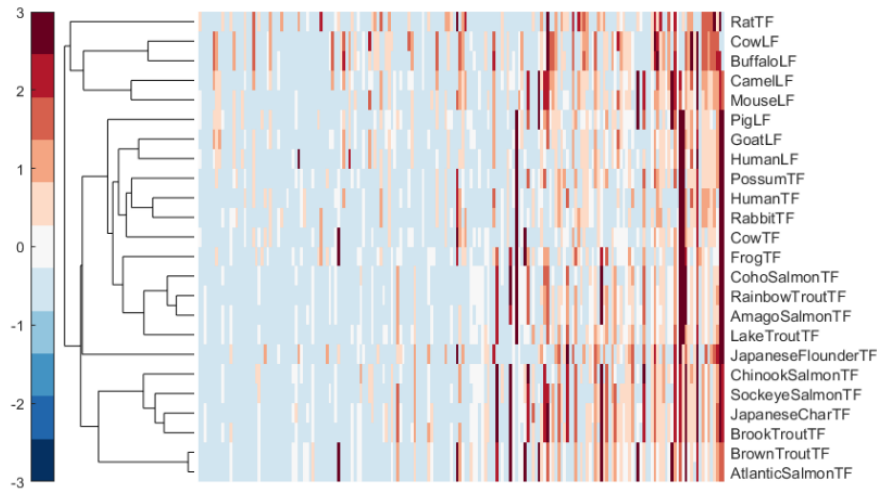


Fig.8. Transferrin and lactoferrin proteins clustered with the proposed method. Some of the LF proteins are grouped with TF proteins. However, they remain close to each other.

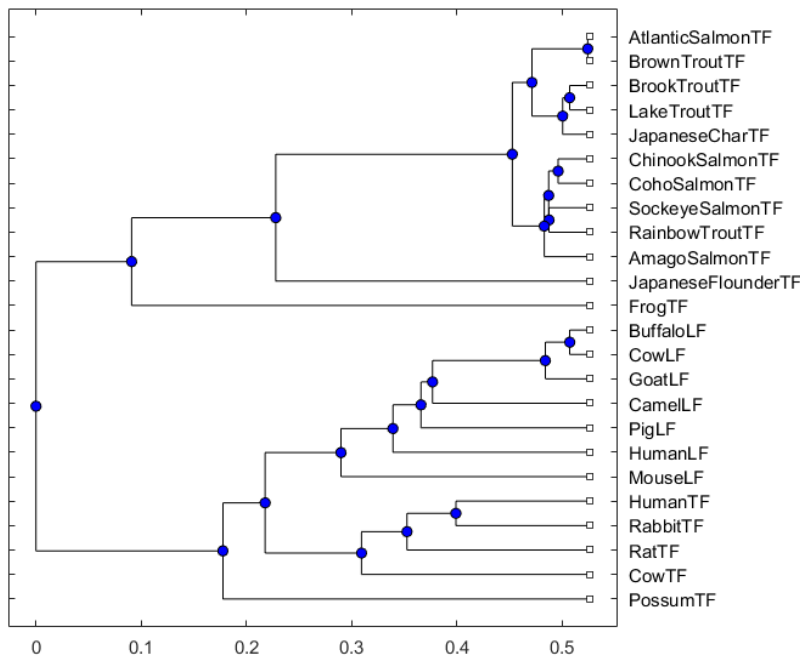


Fig.9. ClustalW clustering result for transferrin and lactoferrin proteins. For this dataset, ClustalW performs better and correctly clusters the TF and LF proteins into their corresponding groups.

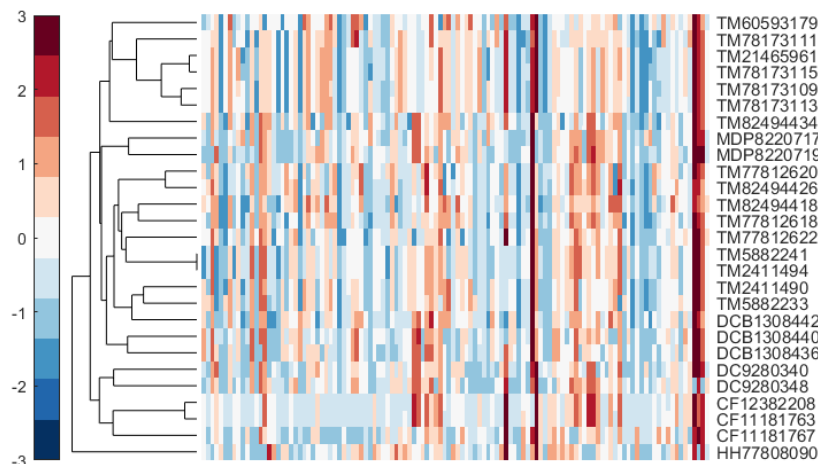


Fig.10. Antifreeze proteins clustered with the proposed method. The proposed method correctly clusters the proteins into their correct groups.

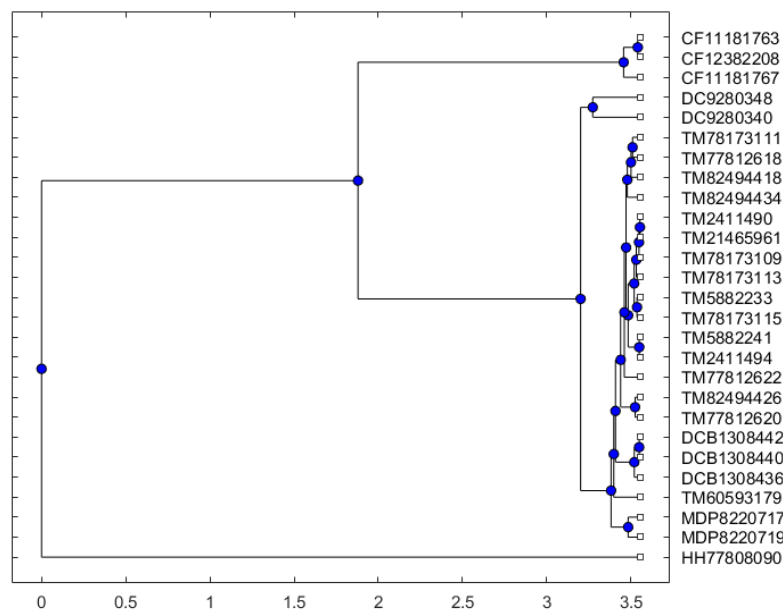


Fig.11. ClustalW clustering result for antifreeze proteins. TM60593179 forms a separate branch which is far away from its own group.

REFERENCES

- [1] Z. Jiang and Z. Yanhong, "Using bioinformatics for drug target identification from the genome." *American Journal of Pharmacogenomics* 5.6 (2005): 387-396.
- [2] M.S. Waterman, "Identification of common molecular subsequence." *Mol. Biol* 147 (1981): 195-197.
- [3] S. F. Altschul, et al., "Basic local alignment search tool." *Journal of molecular biology* 215.3 (1990): 403-410.
- [4] J. Yang and L. Zhang, "Run probabilities of seed-like patterns and identifying good transition seeds." *Journal of Computational Biology* 15.10 (2008): 1295-1313.
- [5] A. Chakraborty and B. Sanghamitra, "FOGSAA: Fast optimal global sequence alignment algorithm." *Scientific reports* 3 (2013): 1746.
- [6] O. Gotoh, "An improved algorithm for matching biological sequences." *Journal of molecular biology* 162.3 (1982): 705-708.
- [7] X. Liu, et al., "Number of distinct sequence alignments with k-match and match sections." *Computers in biology and medicine* 63 (2015): 287-292.
- [8] C. Li, et al., "Protein Sequence Comparison and DNA-binding Protein Identification with Generalized PseAAC and Graphical Representation." *Combinatorial chemistry & high throughput screening* 21.2 (2018): 100-110.
- [9] L. Yu, et al., "Protein sequence comparison based on physicochemical properties and the position-feature energy matrix." *Scientific Reports* 7 (2017): 46237.
- [10] J.D. Thompson, G.H. Desmond and J.G. Toby, "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." *Nucleic acids research* 22.22 (1994): 4673-4680.
- [11] W. Hou, et al., "A new method to analyze protein sequence similarity using Dynamic Time Warping." *Genomics* 109.2 (2017): 123-130.
- [12] L. He, et al. "A novel alignment-free vector method to cluster protein sequences." *Journal of theoretical biology* 427 (2017): 41-52.
- [13] Z. Qi, and J. Meng-Zhe, "An Intuitive Graphical Method for Visualizing Protein Sequences Based on Linear Regression and Physicochemical Properties." *MATCH-Communications in Mathematical and in Computer Chemistry* 75.2 (2016): 463-480.
- [14] C. Li, L. Xueqin and L. Yan-Xia, "Numerical characterization of protein sequences based on the generalized Chou's pseudo amino acid composition." *Applied Sciences* 6.12 (2016): 406.
- [15] Y. Zhang, et al., "Novel numerical characterization of protein sequences based on individual amino acid and its application." *BioMed research international* 2015 (2015).
- [16] Z. Qi, et al., "A protein mapping method based on physicochemical properties and dimension reduction." *Computers in biology and medicine* 57 (2015): 1-7.

- [17] C. Yu, et al., "Protein map: an alignment-free sequence comparison method based on various properties of amino acids." *Gene* 486.1 (2011): 110-118.
- [18] Y. Yao, et al., "Similarity/dissimilarity analysis of protein sequences based on a new spectrum-like graphical representation." *Evolutionary Bioinformatics* 10 (2014): EBO-S14713.
- [19] L. Wang, P. Hui and Z. Jinhua, "ADLD: a novel graphical representation of protein sequences and its application." *Computational and mathematical methods in medicine 2014* (2014).
- [20] C. Wu, et al., "A novel model for protein sequence similarity analysis based on spectral radius." *Journal of theoretical biology* 446 (2018): 61-70.
- [21] N. Jafarzadeh and A. Iranmanesh, "A new measure for pairwise comparison of protein sequences." *MATCH: Communications in Mathematical and in Computer Chemistry* 74 (2015): 563-574.
- [22] Y. Li, et al., "An alignment-free algorithm in comparing the similarity of protein sequences based on pseudo-Markov transition probabilities among amino acids." *PloS one* 11.12 (2016): e0167430.
- [23] H.J. Yu and H. De-Shuang, "Normalized feature vectors: a novel alignment-free sequence comparison method based on the numbers of adjacent amino acids." *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)* 10.2 (2013): 457-467.
- [24] C. Yu, L.He. Rong and SS. Yau, "Protein sequence comparison based on K-string dictionary." *Gene* 529.2 (2013): 250-256.
- [25] A. Czerniecka, et al., "20D-dynamic representation of protein sequences." *Genomics* 107.1 (2016): 16-23.
- [26] Y. Zhang, "A new model of amino acids evolution, evolution index of amino acids and its application in graphical representation of protein sequences." *Chemical Physics Letters* 497.4-6 (2010): 223-228.
- [27] A. El-Lakkani, and H. Mahran, "An efficient numerical method for protein sequences similarity analysis based on a new two-dimensional graphical representation." *SAR and QSAR in Environmental Research* 26.2 (2015): 125-137.
- [28] Z. Mu, et al., "3D-PAF Curve: A Novel Graphical Representation of Protein Sequences for Similarity Analysis." *MATCH: Communications in Mathematical and in Computer Chemistry* 75.2 (2016): 447-462.
- [29] Y. X. Liu, et al, "P-H curve, a graphical representation of protein sequences for similarities analysis." *MATCH Communications in Mathematical and in Computer Chemistry* 70.1 (2013): 451-466.
- [30] ZC. Wu, X. Xuan and C. Kuo-Chen, "2D-MH: A web-server for generating graphic representation of protein sequences based on the physicochemical properties of their constituent amino acids." *Journal of theoretical biology* 267.1 (2010): 29-34.
- [31] G. Huang, and J. Hu., "Similarity/Dissimilarity Analysis of Protein Sequences by a New Graphical Representation." *Current Bioinformatics* 8.5 (2013): 539-544.
- [32] K.V. Holmes, "SARS coronavirus: a new challenge for prevention and therapy." *The Journal of clinical investigation* 111.11 (2003): 1605-1609.
- [33] E.J. Snijder, et al., "Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage." *Journal of molecular biology* 331.5 (2003): 991-1004.
- [34] N. Abbaspour, R. Hurrell and R. Kelishadi, "Review on iron and its importance for human health." *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences* 19.2 (2014): 164.
- [35] M.J. Ford, "Molecular evolution of transferrin: evidence for positive selection in salmonids." *Molecular biology and evolution* 18.4 (2001): 639-647.
- [36] G. Chang, and W. Tianming, "Phylogenetic analysis of protein sequences based on distribution of length about common substring." *The protein journal* 30.3 (2011): 167-172.
- [37] H. Kim, et al., "Marine antifreeze proteins: structure, function, and application to cryopreservation as a potential cryoprotectant." *Marine drugs* 15.2 (2017): 27.

BIOGRAPHIES



BERAT DOĞAN was born in Malatya, Turkey. He received the B.S. degree in electronics engineering from Erciyes University, Turkey in 2006, M.S. degree in biomedical engineering from Istanbul Technical University, Turkey in 2009 and the Ph.D. degree in electronics engineering from Istanbul Technical University, Turkey, in 2015.

From 2008 to 2009 he was a software engineering at Nortel Networks Netas Telecommunication Inc. From 2009 to 2015 he worked as a research assistant at Department of Electronics and Communication Engineering, Istanbul Technical University, Turkey. In 2015, he started to work as an assistant professor at Department of Biomedical Engineering, Inonu University, Turkey. Between 2017-2018 he was a postdoc researcher at Department of Human Genetics, McGill University, Canada. His research interests include, bioinformatics, biomedical signal and image processing and metaheuristics.

APPENDIX

TABLE AI
INFORMATION FOR NINE ND5 PROTEINS

Number	Species	ID (NCBI)	Length
1	Human	AP_000649	603
2	Gorilla	NP_008222	603
3	Common chimpanzee	NP_008196	603
4	Pigmy chimpanzee	NP_008209	603
5	Fin whale	NP_006899	606
6	Blue whale	NP_007066	606
7	Rat	AP_004902	610
8	Mouse	NP_904338	607
9	Opossum	NP_007105	602

TABLE AII
INFORMATION OF THIRTY-FIVE CORONAVIRUS SPIKE PROTEINS

ID (NCBI)	Abbreviation	Name	Group
P10033	FIPV-1146	Feline infectious peritonitis virus strain 79-1146	I
Q66928	FCoV-1683	Feline coronavirus strain 79-1683	I
Q91AV1	PEDVC	Porcine epidemic diarrhea virus strain CV777	I
Q9DY22	TGEVT	Transmissible gastroenteritis virus strain TO14	I
P18450	TGEVF	Porcine transmissible gastroenteritis coronavirus strain FS772/70	I
P36300	CECoV	Canine enteric coronavirus strain INSAVC-1	I
Q9J3E7	MHVM	Murine hepatitis virus strain ML-10	II
Q83331	MHVB	Murine hepatitis virus strain Berkeley	II
P11224	MHVA	Murine hepatitis virus strain A59	II
O55253	MHVD	Murine hepatitis virus strain DVIM	II
Q9IKD1	RtCoV	Rat coronavirus strain 681	II
P25190	BCoVF	Bovine coronavirus strain F15	II
P15777	BCoVM	Bovine coronavirus strain Mebus	II
Q9QAR5	BCoVL	Bovine coronavirus strain LSU-94LSS-051	II
Q91A26	BCoVT	Bovine enteric coronavirus 98TXSF-110-ENT	II
P36334	HCoV-OC43	Human coronavirus strain OC43	II
Q82666	IBV	Infectious bronchitis virus	III
P05135	IBV-6/82	Avian infectious bronchitis virus strain 6/82	III
P12722	IBVD	Avian infectious bronchitis virus strain D274	III
Q64930	IBVC	Infectious bronchitis virus strain CU-T2	III
Q82624	IBVA	Infectious bronchitis virus strain Ark99	III
P11223	IBVB	Avian infectious bronchitis virus strain Beaudette	III
Q98Y27	IBVH	Infectious bronchitis virus strain H52	III
AAP41037	Tor2	SARS coronavirus Tor2	IV
AAP30030	BJ01	SARS coronavirus BJ01	IV
AAR91586	NS-1	SARS coronavirus NS-1	IV
AAP51227	GD01	SARS coronavirus GD01	IV
AAP33697	Frankfurt 1	SARS coronavirus Frankfurt 1	IV
AAP13441	Urbani	SARS coronavirus Urbani	IV
AAQ01597	TC1	SARS coronavirus Taiwan TC1	IV
AAU81608	CDC	SARS Coronavirus CDC 200301157	IV
AAS00003	GZ02	SARS coronavirus GZ02	IV
AAR86788	QXC1	SARS coronavirus ShanghaiQXC1	IV
AAR23250	Sino1-11	SARS coronavirus Sino1-11	IV
AAT76147	TJF	SARS coronavirus TJF	IV

TABLE AIII
THE CONCISE INFORMATION FOR TWENTY-FOUR TRANSFERRIN (TF) AND LACTOFERRIN (LF) PROTEIN SEQUENCES

Sequence name	Species	Accession no	Length
Human TF	Homo sapiens	S95936	698
Rabbit TF	Oryctolagus coniculus	X58533	695
Rat TF	Rattus norvegicus	D38380	698
Cow TF	Bos Taurus	U02564	704
Buffalo LF	Bubahts amee	AJ005203	708
Cow LF	Bos Taurus	X57084	708
Goat LF	Copra hircus	X78902	708
Camel LF	Camehts dromedaries	AJ131674	708
Pig LF	Sus scrofa	M92089	704
Human LF	H. sapiens	NM 002343	710
Mouse LF	Mus musculus	NM 008522	707
Possum TF	Trichosurus vulpecula	AF092510	711
Frog TF	Xenopus laevis	X54530	702
Japanese flounder TF	Pctralichthys olivaceiis	D88801	685
Atlantic salmon TF	Salmo salar	L20313	690
Brown trout TF	Salmo trutta	D89091	691
Lake trout TF	Salvelimts namaycush	D89090	691
Brook trout TF	Sahelinus fontinalis	D89089	691
Japanese char TF	Sahelinus phius	D89088	691
Chinook salmon TF	Oncorhynchus tshawytscha	AH008271	677
Coho salmon TF	Oncorhynchus kisuich	D89084	691
Sockeye salmon TF	Oncorhynchus nerka	D89085	691
Rainbow trout TF	Oncorhynchus mykiss	D89083	691
Amago salmon TF	Oncorhynchus masou	D89086	691

An Active Genomic Data Recovery Attack

M. AKGÜN


Abstract— With the decreasing cost and availability of human genome sequencing, genomic privacy becomes an important issue. Several methods have been proposed in the literature to overcome these problems including cryptographic and privacy-preserving data mining methods: homomorphic encryption, cryptographic hardware. In recent work, Barman et. al studied privacy threats and practical solutions considering an SNP based scenario. The authors introduced a new protocol where a malicious medical center processes an active attack in order to retrieve genomic data of a given patient. The authors have mentioned that this protocol provides a trade-off between privacy and practicality. In this paper, we first give an overview of the system for SNP based risk calculation. We provide the definitions of privacy threats and briefly Barman et al.'s protocol and solution. The authors proposed to use a weighted sum of SNP coefficients for calculating disease tendency. They argue that the specific choice of the bases would prevent unique identification of SNPs. Our main observation is that this is not true. Contrary to the security claim, SNP combinations can be identified uniquely in many different scenarios. Our method exploits a pre-computed look-up table for retrieving SNPs' values from the test result. An attacker can obtain all SNP values of a given patient by using the pre-computed look-up table. We provide practical examples of weights and pre-computed tables. We also mention that even in the case where the table is large and the attacker can not handle it at one time, he can still gather information using multi queries. Our work shows that more realistic attack scenarios must be considered in the design of genetic security systems.

Index Terms—Genomic privacy, secure computation.

I. INTRODUCTION

RECENT DEVELOPMENTS in high throughput sequencing technologies led to a decrease in the cost of genomic sequencing. As a result of this, next-generation sequencing is deployed more and more in clinical diagnosis and treatment. Large scale genomic projects are announced which aim to sequence thousands of individuals (Genomics England [1]). Since genomic data includes sensitive information for individuals and their relatives, efficient use of this data with privacy-preserving techniques becomes an important issue.

METE AKGÜN, is with the Institute for Translational Bioinformatics at the University Hospital Tübingen, Tübingen, Germany, (e-mail: mete.akguen@uni-tuebingen.de).

 <https://orcid.org/0000-0003-4088-2784>

Manuscript received March 22, 2019; accepted August 16, 2019.

DOI: [10.17694/bajece.543555](https://doi.org/10.17694/bajece.543555)

In hospitals, there is a lack of expertise in protecting the genomic data of their patients. Due to the size of the data and the limited resources, it is often difficult for hospitals to safely store, process and maintain the genomic data of patients. The prevention of cyber-attacks by hospitals may not be possible due to insufficient high skilled workers and technology. The solution to this problem is the storage and processing of genomic data in a privacy-protected manner in a third-party service provider. In this case, service providers must process them without seeing the content of the data.

The human genome consists of four different nucleotides (A,C,G,T). These nucleotides form about 20.000 - 25.000 genes responsible for producing various types of proteins which are assigned inside the cells during whole life processes. About %99.5 of the genome is common in the human population where the remaining portion makes up the genetic variance. Most genetic variants in an individual are Single Nucleotide Polymorphisms (SNPs). A single nucleotide poly-morphism (SNP) can be defined as a variation occurring with some probability in a population where a single nucleotide differs from the reference genome. As a result of Genome-Wide Association Studies (GWAS), SNPs provide probabilistic information about the susceptibility of a disease. Generally, a few SNP combinations are evaluated together to calculate the overall inclination to a syndrome such as cardiovascular disease or Alzheimer. Since SNPs form the nonredundant part of the genome and contains minimalistic information, it makes sense to consider privacy-preserving protocols in terms of SNP's.

There are many different types of threats and security models where genomic privacy is a concern: querying on private genomic data, secure querying on public data, secure sequence alignment in public clouds [2]. Several methods have been proposed in the literature to overcome these problems including cryptographic and privacy-preserving data mining methods: homomorphic encryption [3], cryptographic hardware [4],[5].

A. Related Work

Ayday et al. [6] proposed a system based on homomorphic encryption to protect individual's privacy in disease risk tests. This work also proposes to use storage and processing unit to store sensitive data in encrypted form and disease risk tests are performed by authorized institutions using homomorphic encryption technique and secure integer comparison. In this solution, a storage and processing unit (SPU) stores all the

SNPs (approximately 40 million) of the patient. Ayday et. al solved the storage problem in [13] without sacrificing privacy. They classify SNPs as real SNPs and potential SNPs, where real SNPs are set of SNPs observed in the patient. SPU stores

the real SNPs instead of storing all SNPs. However, this constitutes a problem for privacy as SPU stores positions of

TABLE I
COMPARISON OF PREVIOUS SOLUTIONS

Work	Privacy			Authorization	Efficiency	Weighted Av
	Test Inference	Passive SNP	Active SNP			
Ayday et al. [6]	✓	X	X	X	X	✓
Ayday et al. [7]	✓	✓	X	X	X	✓
Danezis and De Cristofaro [8]	✓	X	X	X	✓	✓
Djatkiko et al. [9]	✓	X	X	X	X	✓
Zhang et al. [10]	✓	X	X	X	✓	✓
Fan and Mohanty [11]	✓	X	X	X	✓	✓
Perillo and De Cristofaro [12]	✓	X	X	✓	✓	✓

the real SNPs in plain text. Ayday et. al solved this problem by storing the real SNPs along with some redundant content from the set of potential SNPs.

In order to improve Ayday et al.'s scheme [6], Danezis and De Cristofaro [8] used Additively Homomorphic Elliptic Curve based El-Gamal (AH-ECC) [14] instead of the Paillier cryptosystem in order to decrease the computational overhead. The patient has a smartcard that participates in the protocol execution. The lost of the smartcard can cause privacy violation. Furthermore, the cloud provider knows the number of SNPs of each patient. This is also a data leakage. Perillo and De Cristofaro [12] proposed a cryptographic protocol for running different types of tests on individuals' genetic data. Their scheme is also based on the use of AH-ECC [14]. Differently it provides authorization which means SNP wights and locations are verified by central authority such as the FDA.

Djatkiko et al. [9] proposed a privacy-preserving algorithm to compute genomic tests that need the linear combination of SNP values. They applied partially homomorphic Paillier encryption and private information retrieval techniques to protect patients' privacy. The computational overhead of their solution is very high when compared to that of Ayday et al.'s solution.

Zhang et al. [10] proposed a framework for disease risk calculation using SNP values. Their framework reduces the storage overhead of previous solutions significantly by using bloom filters. It also reduces communication cost by indexing the encrypted genomic data.

Fan and Mohanty [11] proposed a solution for privacy preserving calculation of the susceptibility of a patient to a particular disease. The proposed scheme is based on Shamir's (1, n) secret sharing [15] which allows the computation of a certain number of multiplications and unlimited additions. It is more efficient than Ayday et al.'s solution [6] in terms of storage and computation time.

Readers are recommended to read surveys in [16], [2], [17], [18], [19] and [20] for more information on genomic privacy.

B. Our Contributions

All existing works provide security under semi-honest model in which the involving parties are not able to deviate the protocol description. It is very easy to provide security under this model with lower communication and computation complexities because adversaries are not allowed to change their inputs and to collude with other parties. This shows that all previous works are vulnerable to active SNP retrieval attacks in which an attacker can modify SNP weights in order to learn SNP values. The comparison of previous solutions is given in Table I.

Barman et al. [21] proposed a solution that makes all existing works secure to active SNP retrieval attacks. They studied privacy threats and practical solutions considering an SNP based disease risk calculation scenario. The authors introduced a new protocol where a malicious medical center processes an active attack in order to get SNP values of an individual. The authors mentioned that this protocol provides a tradeoff between privacy and practicality. In this study, we show that the solution offered by Barman et. al [21] does not prevent the leakage of SNP values. We show that SNP combinations can be uniquely identified in many different scenarios. Our method uses a pre-calculated lookup table to retrieve the values of the SNPs from the test result. The attacker can obtain all SNP values of a particular patient using the previously calculated lookup table. We present practical examples of weights and pre-calculated tables. We also observe that even if the lookup table is very large to handle, and the attacker can infer SNP values with multiple queries Our study shows that more realistic attack scenarios should be considered in the design of genetic security systems.

This paper is organized as follows. In Section II, we give an overview of the system model for genetic risk test calculation. In Section III, we give the definitions of privacy threats. In Section IV, we briefly define Barman et al.'s protocol [21] and their privacy solution. In Section V, we explain our observation that in fact, the solution is redundant. In Section

VI, we explain possible and existing countermeasures in order to eliminate active SNP retrieval attacks. Finally, in Section VII, we conclude the paper.

II. SYSTEM MODEL

In this section, we give the overview of the generic model described in the literature [6], [21] before. This model is constructed in order to calculate genetic risk test in a privacy-preserving way. In brief, a patient (P) sends his sample to the certificated institution (CI) for sequencing. The CI extracts genomic variants (SNPs) of the patient and encrypts SNPs. Then, the CI sends encrypted genomic data to the data center (DC). The CI is also responsible to distribute encryption keys to the related parties. The DC stores the encrypted genomic data. Medical center(s) (MC) communicate with the DC in order to compute genetic risk test in a privacy-preserving way. The system model is summarized in Figure 1.

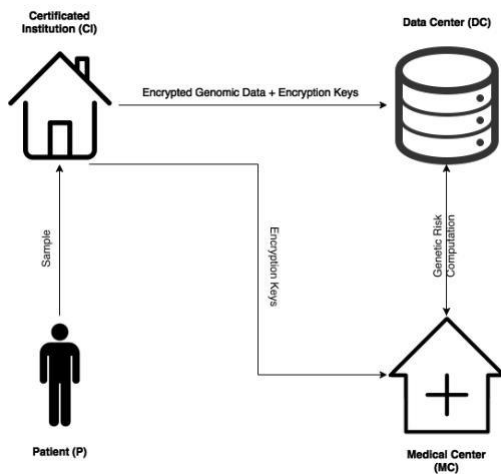


Fig.1. System model architecture

The genetic risk (G) is usually computed as a weighted sum of SNPs' values (Equation 1). W_i is the contribution (weight) of SNP_i . This computation can be done in a privacy-preserving way using secure multiparty computation or smart cards [22]. At the end of the test, the MC learns only the test result, but not the SNPs' values. Furthermore, DC does not learn the SNPs' weights.

$$\sum_{i=0}^n (W_i * SNP_i)$$

III. PRIVACY THREATS

Barman et al. [21] investigated privacy threats for system model architecture described in Section II. In the literature, P and CI are considered as honest parties and the MC and the DC are considered as honest-but-curious parties. Barman et al. [21] extended possible privacy threats by considering the MC and the DC as honest, semi-honest (passive) and dishonest (active). They describe three main attacks.

A. Test Inference Attacks

The semi-honest DC can learn which SNPs are used and how often they are used from test queries. Therefore, DC can infer the disease which a corresponding patient is suffering from. If the DC can re-identify P, this violates the privacy of the patient P. Danezis et al. [8] proposed to use all SNP values of a given patient in the genetic risk computation in order to prevent test inference attack. Another solution [23] proposed to use oblivious RAM to prevent the DC from learning access patterns of the MC.

B. Passive SNP Retrieval Attack

The MC can learn SNPs' values from the test result because the risk calculation is a linear equation and the MC knows some parameters used in this equation such as SNPs' weights. As the number of queries increases for a given patient P, P's privacy decreases. Ayday et al. [7] proposed to deliver test result as a range in order to prevent this attack.

C. Active SNP Retrieval Attack

In active SNP retrieval attack, the dishonest MC can manipulate SNPs' weights in order to retrieve SNPs' values from test results easily. For example, the MC sets all SNP weights to 0 except $W_j = 1$. The MC can retrieve the value of SNP_j which is equal to the test result. In another version of active SNP retrieval attack, the MC sets SNPs' weights as consecutive powers of a number. Consider a test with three SNPs, the MC sets SNPs' weights as the following: $W_0 = 4^0$, $W_1 = 4^1$ and $W_2 = 4^2$. The test result G is $(36)_{10} = (210)_4$. An attacker can retrieve SNPs' values from $G = (210)_4$ as the following: $SNP_2 = 2$, $SNP_1 = 1$ and $SNP_0 = 0$.

IV. BARMAN ET AL.'S PROTOCOL

Barman et al. [21] offer a solution to overcome the active SNP retrieval attack. According to the authors' definition: active SNP retrieval attack can be practiced by the MC by setting new SNP weights for a given test to retrieve the SNPs' raw values without being detected. Their solution is to force the MC to iteratively utter some SNP weights to the DC until the DC assures that the current test is legitimate. As the authors' mention, this solution weakens the MC's privacy while giving more power to the DC. Learning the test parameters might allow the DC to practice the test inference attack and also the test parameters might be private to the MC. So, the MC can abort the protocol if it thinks that it has to give too much information about the test parameters. The authors assume that only the MC can get the mapping information from the CI and the SNPs are stored as shuffled at the DC. The suggested protocol based on the described system model is as follows:

1. The MC wants to compute a genetic risk test based on R SNPs of a given patient but it adds D dummy SNPs with zero weight to its query. By adding D dummy SNPs, it prevents the DC from a test

inference attack and by adding zero weights to the dummy SNPs, it convinces the DC that the test is legitimate. Dummy SNPs have no effect on the test result. The authors call $N = R+D$ as the total length of the query.

- The MC sends a request of N SNPs and a commitment for each SNP weight W_i , to the DC. They call $C_i = \text{Commit}(W_i) \forall i \in [0, N-1]$ as the commitment.

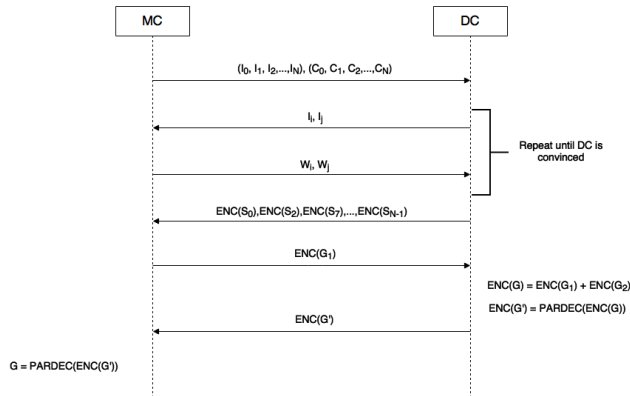


Fig.2. Barman et al.'s Protocol

- The DC asks for the weights of random two indices $j, k \in [0, N-1]$. The MC responds with the relevant W_j and W_k to the DC.
- The DC controls both the commitments C_j and C_k , and the weights W_j and W_k . If both weights are non-zero, and not different powers of the same number, the DC assures that the test is not an active SNP retrieval attack. Steps 3 and 4 are repeated until the DC is convinced or the MC aborts the protocol. For each iteration after the first, the DC can ask for only one new weight.
- After believing that the test is not an active SNP retrieval attack, the DC sends the S ($S = N-2$ at most) encrypted SNPs corresponding to the weights not seen during the previous steps.
- The MC homomorphically computes and sends the encryption of the first part of the test result, $\text{ENC}(G_1)$, according to the S SNPs.
- The DC computes the encryption of the second part of the test result $\text{ENC}(G_2)$, according to the other encrypted SNPs whose weights are known from steps 3 and 4. The two partial results are homomorphically added into $\text{ENC}(G) = \text{ENC}(G_1) + \text{ENC}(G_2)$. $\text{ENC}(G')$, partial decryption of $\text{ENC}(G)$ is sent to the MC.
- The MC decrypts the $\text{ENC}(G')$ and obtains G . The protocol ends

The authors declare that once the DC makes sure that the test is legitimate, it computes the encryption of the partial test result, $\text{ENC}(G_2)$. This guarantees that an active SNP retrieval attack cannot be performed, independent from the weights used for $\text{ENC}(G_1)$.

V. PROPOSED ACTIVE SNP RETRIEVAL ATTACK

In this section, we present an active SNP retrieval attack. We apply our attack to Barman et al.'s protocol. The authors suggest using specific bases for preventing unique identification of SNPs. We discover that this is redundant. An SNP combination can be identified uniquely in many different scenarios, we provide some examples. Our attack uses a pre-computed look-up table for retrieving SNPs' values from the test result. The attack can be described as follows.

- The dishonest MC chooses R prime numbers as SNPs weights ($W_0, W_1, W_2, \dots, W_R$).
- The dishonest MC calculates test results for all possible SNPs' values and stores SNP's values and tests results in the table T . The sample look-up table with $R = 3$, $W_0 = 3$, $W_1 = 11$ and $W_2 = 23$ is shown in Table II.
- The dishonest MC calculates commitments of SNPs weights.
- The dishonest MC creates a query and adds D dummy SNPs with zero weights to its query.
- The dishonest MC sends a request of $N = (R + D)$ SNPs and a commitment for each SNP weight W_i , to the DC.
- The DC asks for the weights of random two indices $j, k \in [0, N-1]$. The MC responds with the relevant W_j and W_k to the DC.
- The DC controls both the commitments C_j and C_k , and the weights W_j and W_k . If both weights are non-zero, and not different powers of the same number, the DC assures the test is not an active SNP retrieval attack. Steps 3 and 4 are repeated until the DC is convinced or the MC aborts the protocol. For each iteration after the first, the DC can ask for only one new weight.
- After believing that the test is not an attack, the DC sends the S ($S = N-2$ at most) encrypted SNPs corresponding to the weights not seen during the previous steps. In our attack, the DC is convinced eventually because we choose at least two non-zero weights and SNPs' weights are guaranteed not to be different powers of the same number.
- The dishonest MC homomorphically computes and sends the encryption of the first part of the test result, $\text{ENC}(G_1)$, according to the S SNPs.

10. The DC computes the encryption of the second part of the test result $ENC(G_2)$, according to the other encrypted SNPs whose weights are known from steps 3 and 4. The two partial results are homomorphically added into $ENC(G) = ENC(G_1) + ENC(G_2)$. $ENC(G')$, partial decryption of $ENC(G)$ is sent to the dishonest MC.
11. The dishonest MC decrypts the $ENC(G')$ and obtains G . Then, the dishonest MC retrieves SNPs' values by using the look-up table T.

The predetermined weight values give unique test results each SNPs' values as shown in Table II. Therefore, the success probability of our attack is 1. An attacker can retrieve all SNPs' values of a given patient by using a look-up table like Table II. In the protocol design, Barman et al. [21] considered two special attack scenarios. Our attack shows that more realistic attack scenarios must be considered when designing a security solution for the genetic system model shown in Figure 1.

TABLE II
ATTACK TABLE WITH SNP WEIGHTS ($W_0 = 3, W_1 = 11, W_2 = 23$)

SNP ₀	SNP ₁	SNP ₂	Weighted Sum
0	0	0	0
0	0	1	23
0	0	2	46
0	1	0	11
0	1	1	34
0	1	2	57
0	2	0	22
0	2	1	45
0	2	2	68
1	0	0	3
1	0	1	26
1	0	2	49
1	1	0	14
1	1	1	37
1	1	2	60
1	2	0	25
1	2	1	48
1	2	2	71
2	0	0	6
2	0	1	29
2	0	2	52
2	1	0	17
2	1	1	40
2	1	2	63
2	2	0	28
2	2	1	51
2	2	2	74

Although a small size look-up table is sufficient to retrieve all SNPs of a given patient, an attacker may want to create a large size look-up table. The difficulty of creating unique look-up tables increases, as the number of SNPs queried increases. When the dishonest MC cannot create a unique table, it can create multiple partial unique tables in order to create unique test result for all possible SNPs values. When the MC calculates the test result G which is not unique as shown in Table III, the MC sends another query for the same SNPs by using weights of another table in order to calculate the unique test result. For example, an attacker queries SNP_i ,

SNP_j and SNP_k with SNP Weights ($W_0 = 3, W_1 = 5, W_2 = 7$) given in Table III. If the calculated test result is 10, the attacker cannot determine the SNPs' values because 10 is not unique. Then, the attacker queries SNP_i, SNP_j and SNP_k with SNP Weights ($W_0 = 3, W_1 = 5, W_2 = 11$) given in Table III. The calculated test result is 14. The attacker can retrieve SNPs' values ($SNP_i = 1, SNP_j = 0$ and $SNP_k = 1$). As a result, the attacker retrieves the SNPs' values by using two partial unique look-up tables.

VI. COUNTERMEASURES

Many studies on genomic privacy have focused on the semi-honest model as an adversarial model. There are a few studies evaluating the dishonest model in the literature. The attack described in this article is carried out by malicious attackers playing on the inputs.

TABLE III
ATTACK TABLE WITH SNP WEIGHTS ($W_0 = 3, W_1 = 5, W_2 = 7$)
AND SNP WEIGHTS ($W_0 = 3, W_1 = 5, W_2 = 11$)

SNP ₀	SNP ₁	SNP ₂	Weighted Sum 1	Weighted Sum 2
0	0	0	0	0
0	0	1	7	11
0	0	2	14	22
0	1	0	5	5
0	1	1	12	16
0	1	2	19	27
0	2	0	10	10
0	2	1	17	21
0	2	2	24	32
1	0	0	3	3
1	0	1	10	14
1	0	2	17	25
1	1	0	8	8
1	1	1	15	19
1	1	2	22	30
1	2	0	13	13
1	2	1	20	24
1	2	2	27	35
2	0	0	6	6
2	0	1	13	17
2	0	2	20	28
2	1	0	11	11
2	1	1	18	22
2	1	2	25	33
2	2	0	16	16
2	2	1	23	27
2	2	2	30	38

Clinicians want to see the result of the risk calculation in order to give the right treatment or prevention. The clinician (adversary) knows the weight values used in the risk calculations and can select them as desired. The risk is usually calculated as the weighted sum of the SNP values. In this case, the clinician will obtain SNP values from the calculated risk value. It is possible to make it difficult for an attacker to obtain SNP values by controlling weight values as Barman et al [21]. But this is not the definitive solution. Appropriate weight values will always be selected to obtain SNP values.

Ayday et al. [7] proposed to give risk value as a range. In this solution, as the range value increases, patient privacy increases but the consistency of the test decreases.

As another solution, weight values can be stored encrypted and clinicians do not know these values. Thus, it becomes

impossible to obtain SNP values for calculations where more than one SNP value is used. In this solution, it is difficult to keep the weight values secretly in a central system. In real life, this solution is very difficult to implement.

Another solution is to give the test result to the patient privately. To do this, the test result must be given to the patient in an encrypted form and the patient must be able to decrypt it. The patient can share the test result with the clinician if he wishes. A method for transferring, storing and decrypting the data must be specified. These operations can be done safely using smart card technology. The partially decrypted test result is transferred to the smart card of the user. The user can read the test result privately using a reader and software on the personal computer. Smart cards are capable of decryption. Since the smart cards are tamper-proof devices, the test result can be reliably stored.

None of the proposed solutions can provide a definitive solution. The clinician learning the exact result of the test can always infer the SNP values.

VII. CONCLUSION

A recent study [21] proposed a new threat model where malicious medical center tries to retrieve genomic data of a given patient. The authors also proposed a solution to this type of attack. They claim that their solution may be vulnerable to more sophisticated attacks involving multiple queries. We show that in fact there is a simpler type of attack. The attacker can learn genomic data using a simple pre-computed look-up table. It remains for future work to develop a security solution to prevent our active SNP retrieval attack. The researchers have to make a trade-off between privacy and efficiency in order to reduce the effects of active SNP retrieval attacks.

REFERENCES

- [1] "Genomics England — 100,000 Genomes Project," accessed: 2015-07-05. [Online]. Available: <http://www.genomicsengland.co.uk/>
- [2] M. Akgun, A. O. Bayrak, B. Ozer, and M. S. Sagiroglu, "Privacy preserving processing of genomic data: A survey," *Journal of Biomedical Informatics*, vol. 56, no. 0, pp. 103–111, 2015.
- [3] M. Goodrich, "The mastermind attack on genomic data," in *Security and Privacy, 2009 30th IEEE Symposium on*, May 2009, pp. 204–218.
- [4] M. Canim, M. Kantarcioglu, and B. Malin, "Secure management of biomedical data with cryptographic hardware," *Trans. Info. Tech. Biomed.*, vol. 16, no. 1, pp. 166–175, Jan. 2012.
- [5] C. Uhler, A. B. Slavkovic, and S. E. Fienberg, "Privacy-preserving data sharing for genome-wide association studies," *Journal of Privacy and Confidentiality*, vol. 5, no. 1, pp. 137–166, 2013.
- [6] E. Ayday, J. L. Raisaro, P. J. McLaren, J. Fellay, and J.-P. Hubaux, "Privacy-preserving computation of disease risk by using genomic, clinical, and environmental data," in *Proceedings of the 2013 USENIX Conference on Safety, Security, Privacy and Interoperability of Health Information Technologies*, ser. *HealthTech'13*. Berkeley, CA, USA: USENIX Association, 2013, pp. 1–1.
- [7] E. Ayday, J. L. Raisaro, J. Hubaux, and J. Rougemont, "Protecting and evaluating genomic privacy in medical tests and personalized medicine," in *Proceedings of the 12th annual ACM Workshop on Privacy in the Electronic Society*, WPES 2013, Berlin, Germany, November 4, 2013, 2013, pp. 95–106.

- [8] G. Danezis and E. D. Cristofaro, "Fast and private genomic testing for disease susceptibility," in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, WPES 2014, Scottsdale, AZ, USA, November 3, 2014, 2014, pp. 31–34. [Online]. Available: <http://doi.acm.org/10.1145/2665943.2665952>
- [9] M. Djatmiko, A. Friedman, R. Boreli, F. Lawrence, B. Thorne, and S. Hardy, "Secure evaluation protocol for personalized medicine," in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, ser. WPES '14. New York, NY, USA: ACM, 2014, pp. 159–162. [Online]. Available: <http://doi.acm.org/10.1145/2665943.2665967>
- [10] J. Zhang, L. Zhang, M. He, and S. Yiu, "Privacy-preserving disease risk test based on bloom filters," in *Information and Communications Security - 19th International Conference, ICICS 2017, Beijing, China, December 6-8, 2017, Proceedings*, 2017, pp. 472–486. [Online]. Available: https://doi.org/10.1007/978-3-319-89500-0_41
- [11] G. Fan and M. Mohanty, "Privacy-preserving disease susceptibility test with shamir's secret sharing," in *Proceedings of the 14th International Joint Conference on e-Business and Telecommunications (ICETE 2017) - Volume 4: SECUREPT*, Madrid, Spain, July 24–26, 2017., 2017, pp. 525–533.
- [12] M. Perillo and E. D. Cristofaro, "PAPEETE: private, authorized, and fast personal genomic testing," in *Proceedings of the 15th International Joint Conference on e-Business and Telecommunications, ICETE 2018 - Volume 2: SECUREPT*, Porto, Portugal, July 26–28, 2018., 2018, pp. 650–655.
- [13] E. Ayday, J. L. Raisaro, and J.-P. Hubaux, "Personal Use of the Genomic Data: Privacy vs. storage Cost," in *IEEE Global Communications Conference, Exhibition and Industry Forum – GLOBECOM*, 2013.
- [14] T. E. Gamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," *IEEE Trans. Information Theory*, vol. 31, no. 4, pp. 469–472, 1985.
- [15] A. Shamir, "How to share a secret," vol. 22, no. 11, pp. 612–613, 1979.
- [16] Y. Erlich and A. Narayanan, "Routes for breaching and protecting genetic privacy," *Nature Reviews Genetics*, vol. 15, no. 6, pp. 409–421, 2014.
- [17] M. Naveed, E. Ayday, E. W. Clayton, J. Fellay, C. A. Gunter, J.-P. Hubaux, B. A. Malin, and X. Wang, "Privacy in the genomic era," *ACM Computing Surveys*, vol. 48, no. 1, pp. 1–44, 2015. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2808687.2767007>
- [18] M. Z. Hasan, M. S. R. Mahdi, and N. Mohammed, "Secure count query on encrypted genomic data: A survey," *IEEE Internet Computing*, vol. 22, no. 2, pp. 71–82, 2018. [Online]. Available: <https://doi.org/10.1109/MIC.2018.112102323>
- [19] M. M. A. Aziz, M. N. Sadat, D. Alhadidi, S. Wang, X. Jiang, C. L. Brown, and N. Mohammed, "Privacy-preserving techniques of genomic data - a survey," *Briefings in Bioinformatics*, vol. 20, no. 3, pp. 887–895, 2019.
- [20] A. Mittos, B. Malin, and E. D. Cristofaro, "Systematizing genome privacy research: A privacy-enhancing technologies perspective," *PopET's*, vol. 2019, no. 1, pp. 87–107, 2019. [Online]. Available: <https://doi.org/10.2478/popets-2019-0006>
- [21] L. Barman, M. T. Elraini, J. L. Raisaro, J. Hubaux, and E. Ayday, "Privacy threats and practical solutions for genetic risk tests," in *2015 IEEE Symposium on Security and Privacy Workshops, SPW 2015, San Jose, CA, USA, May 21–22, 2015, 2015*, pp. 27–31. [Online]. Available: <https://doi.org/10.1109/SPW.2015.12>
- [22] M. Akgun, B. Erguner, A. O. Bayrak, and M. S. Sagiroglu, "Human genome in a smart card," in *HEALTHINF 2014 - Proceedings of the International Conference on Health Informatics, ESEO, Angers, Loire Valley, France, 3–6 March, 2014, 2014*, pp. 310–316. [Online]. Available: <http://dx.doi.org/10.5220/0004799903100316>
- [23] N. P. Karvelas, A. Peter, S. Katzenbeisser, E. Tews, and K. Hamacher, "Privacy-preserving whole genome sequence processing through proxy-aided ORAM," in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, WPES 2014, Scottsdale, AZ, USA, November 3, 2014, 2014, pp. 1–10. [Online]. Available: <http://doi.acm.org/10.1145/2665943.2665962>

BIOGRAPHIES



METE AKGÜN was born in Giresun, Turkey, in 1980. He received his B.Sc. degree in electrical engineering (with high honor) from Bahçeşehir University, Istanbul, in 2005, the M.Sc. and the Ph.D. degrees in computer engineering from Boğaziçi University, İstanbul, in 2009 and 2015, respectively. He worked as a research engineer in TÜBİTAK between 2006 and 2019. Since 2018, he is a postdoctoral researcher at University of Tübingen, Germany. His research interests include security, data privacy and bioinformatics.

Volt / VAR Regulation in Energy Transmission Systems Using SVC and STATCOM Devices

H.F. CARLAK and E. KAYAR


Abstract— To ensure energy continuity in the interconnected power systems, production and consumption amounts should be in balance, and that can be possible by providing constant voltage and frequency. Otherwise, undesirable large dynamic oscillations, voltage collapse may occur, and the quality of the electrical energy deteriorates. In this study, the power system of the Denizli Region in the Western Mediterranean Region is modeled using realistic national data. The simulation study is carried out to obtain the voltages, active and reactive powers of buses in the case of no-fault and fault occurrence. Using FACTS (Flexible Alternating Current Transmission Systems) technology such as SVC and STATCOM, the power systems can be controlled, and the carrying capacities can be improved within specified limits. Facts devices are capable of producing and consuming reactive power depending on immediate needs, safe and operating flexibility, and having a high reaction time in the simulation studies. The maximum load limits are increased, and the control of the power system is facilitated. STATCOM provides 31.09 % more achievement with respect to SVC device at energy production and consumption rates for the modeled pilot region. The reliability of the power system has been increased against a voltage collapse, thanks to the SVC and STATCOM controllers. Furthermore, the voltage stability of the generator is also raised, and correspondingly the capacity of the power system enhances. The results are compared to the existing system, and the obtained improvements may be assessed for the enhancement of the interconnected network grid of Turkey.

Index Terms—Power System Analysis, Statcom, Svc, Facts, Voltage Regulation, Reactive Power Regulation


I. INTRODUCTION

POWER ELECTRONICS-BASED Flexible Alternating Current Transmission Systems (FACTS) are used for active power, reactive power, impedance, and voltage control in power systems.

HAMZA FEZA CARLAK, is with Department of Electrical and Electronics Engineering Akdeniz University, Antalya, Turkey, (e-mail: fezacarlak@akdeniz.edu.tr).

 <https://orcid.org/0000-0002-8561-4591>

ERGIN KAYAR, is with Department of Electrical and Electronics Engineering Akdeniz University, Antalya, Turkey, (e-mail: erginkayar07@gmail.com).

 <https://orcid.org/0000-0002-7356-2165>

Manuscript received July 23, 2019; accepted October 16, 2019.

DOI: [10.17694/bajece.595761](https://doi.org/10.17694/bajece.595761)

In the recent years, flexible AC transmission system (FACTS) technology has been used in efficient energy utilization, demand control, power quality enhancement, harmonic mitigation, voltage regulation, reactive power compensation, transient and steady state voltage stability enhancement, power loss reduction, power conditioning and quality improvement [1, 2, 3, 4]. The emerging use of renewable and distributed generation has accelerated and expanded the role of power electronic devices for efficient electrical utilization and enhanced security and reliability of the electric utility grid [5]. Also, new applications have emerged for standalone microgrids with regards to renewable energy utilization using solar photovoltaic (PV) systems, micro-hydroelectric systems, the wind, biomass, waste-to-energy and hybrid ac-dc sources with battery energy storage for remote villages [6]. Renewable energy sources (RESs) are utilized at an accelerating rate and connected to both transmission and distribution/utilization systems using power electronic converters. This results in increased harmonics and deterioration of power quality at the point of common coupling. Power quality issues and mitigation have emerged as serious challenges and issues facing electric utilities and industrial/commercial/residential users [7].

Various FACTS devices and control strategies can help to mitigate power quality problems. For efficient use of power system resources, the concept of FACTS has been introduced in the late 1980's. The basic concept of FACTS devices has been based upon the use of high-voltage power electronics to control real and reactive power flow and voltage in the transmission system [8]. Extensive research has focused on new topologies and architectures of voltage-source converters (VSCs) to increase the performance of FACTS devices in transmission and distribution systems and consequently enhance power system security [9,10]. FACTS devices and smart control strategies have been gaining a more prominent role in energy generation from renewable sources [11]. The results of the implementation of FACTS devices in smart grids with renewable systems are encouraging [12,13].

Determining the location and size of FACTS devices which provide important benefits in power flow control, increasing transmission transfer capacity, voltage stability, reactive power control is of great importance both technically and economically. The voltage regulation is provided by the help of FACTS devices to prevent the system from operating under

unstable conditions that may cause the partial or complete collapse. The study is implemented by modelling 10 bus prototype model of the energy transmission system of Denizli Region in the Western Mediterranean Region of the dynamic model SVC (Static Var Compensator), and STATCOM (Static Synchronous Compensators) devices using realistic data. The analyses are implemented to determine the voltage stability of the pilot electrical power system region. Optimum location and performance of FACTS devices in terms of bus voltage fluctuations and line capacities are tested. Simulation study is carried out using national data of the realistic pilot region and the results are compared to the current system and the benefits may be evaluated for the improvement of the Turkish interconnected system. The aim of the study is minimizing losses and providing voltage and reactive power control by using FACTS technology instead of the existed capacitor groups or reactors which have pretty long reaction times.

The usage and development of FACTS in power transmission systems brings many applications to improve the stability of power systems [14]. They may be used to increase the stability of the system and control the power flow. The most significant advantages of such devices are their flexibility and controllability [15]. Their applications are generally concentrated on issues such as increasing voltage stability, damping oscillations, voltage control in power systems and improving stability of power systems. These applications can be made by checking the voltage value and the phase angle [16].

Reactive compensation using STATCOM and DGM-STATCOM (Pulse Width Modulation Static Synchronous Compensator) devices were performed in Coteli's study to compare the controller performances of both devices. The simulation study shows that STATCOM responds very quickly to unexpected sudden voltage variations [17]. A methodology for introducing FACTS models into the energy transmission network using DIG-SILENT Power Factory program was disclosed in Cepeda et al's study [18]. The application methodology SVC, TCSC, SSSC, and STATCOM for the application of DIG-SILENT Power Factory to the stability studies of electrical power systems involves four stages. The SVC, TCSC (Thyristor Controlled Series Capacitor), SSSC (Static Synchronous Series Compensator), and STATCOM models are properly applied into the system and the contribution of FACTS devices in improving the stability of the power system were denoted.

In the literature, the studies on FACTS technology of power systems have been mostly carried out by classifying the FACTS controllers. However, feasibility studies have not been already carried out by using FACTS technology for the realistic power system model, which is constructed with realistic values by taking into account all parameters for the energy transmission system which is the novelty of this study.

II. METHOD AND MODELING

A. Load Flow Analysis

In the feasibility studies of the modelled region, the power system elements such as transformers, capacitors, reactors and FACTS devices used in electrical power systems, which play

an active role in electrical networks, are analyzed by performing load flow algorithms.

The effects of generators, non-linear loads, and other devices connected to the grid nodes are reflected in the node current. Constant impedance loads (linear) are also included in the node admittance matrix. In the formation of node equations, the equation will be linear if I current inputs are known in non-linear power flow equations. The current inputs depend on the P, Q and V parameters in any k node:

$$\begin{bmatrix} \Delta P_2^{(k)} \\ \vdots \\ \Delta P_n^{(k)} \\ \Delta Q_2^{(k)} \\ \vdots \\ \Delta Q_n^{(k)} \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial P_2}{\partial \delta_2}\right)^{(k)} & \dots & \left(\frac{\partial P_2}{\partial \delta_n}\right)^{(k)} & \left(\frac{\partial P_2}{\partial |V|_2}\right)^{(k)} & \dots & \left(\frac{\partial P_2}{\partial |V|_n}\right)^{(k)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \left(\frac{\partial P_n}{\partial \delta_2}\right)^{(k)} & \dots & \left(\frac{\partial P_n}{\partial \delta_n}\right)^{(k)} & \left(\frac{\partial P_n}{\partial |V|_2}\right)^{(k)} & \dots & \left(\frac{\partial P_n}{\partial |V|_n}\right)^{(k)} \\ \left(\frac{\partial Q_2}{\partial \delta_2}\right)^{(k)} & \dots & \left(\frac{\partial Q_2}{\partial \delta_n}\right)^{(k)} & \left(\frac{\partial Q_2}{\partial |V|_2}\right)^{(k)} & \dots & \left(\frac{\partial Q_2}{\partial |V|_n}\right)^{(k)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \left(\frac{\partial Q_n}{\partial \delta_2}\right)^{(k)} & \dots & \left(\frac{\partial Q_n}{\partial \delta_n}\right)^{(k)} & \left(\frac{\partial Q_n}{\partial |V|_2}\right)^{(k)} & \dots & \left(\frac{\partial Q_n}{\partial |V|_n}\right)^{(k)} \end{bmatrix} \begin{bmatrix} \Delta \delta_2^{(k)} \\ \vdots \\ \Delta \delta_n^{(k)} \\ \Delta |V|_2^{(k)} \\ \vdots \\ \Delta |V|_n^{(k)} \end{bmatrix}$$

Where, P and Q are specified for load bus bars and P and V are defined for voltage-controlled bus bars. For other node types, the relations between P, Q, V and I are defined by the characteristics of the devices connected to those nodes. The boundary condition problems brought about by various types of nodes is turned into a nonlinear problem and solved iteratively using Fast-Decoupled Newton-Raphson Method which is an accelerated version of the Newton-Raphson Method [19].

In order to further reduce the calculation time of the Fast-Decoupled Method, the Jacobian matrix is created according to the initial conditions and the Fast-Decoupled method with constant Jacobian is going to be applied during the calculation.

$$\begin{aligned} \frac{\partial P_i}{\partial V_j} &= 2V_i G_{ii} + \sum_{\substack{k=1 \\ \neq i}}^n V_k Y_{ik} \cos(\theta_i - \theta_k - \alpha_{ik}) \quad (2) \\ &= 2V_i G_{ii} + \sum_{\substack{k=1 \\ \neq i}}^n V_k Y_{ik} [\cos(\theta_i - \theta_k) \cos \alpha_{ik} + \sin(\theta_i - \theta_k) \sin \alpha_{ik}] \\ &= 2V_i G_{ii} + \sum_{\substack{k=1 \\ \neq i}}^n V_k [G_{ik} \cos(\theta_i - \theta_k) + B_{ik} \sin(\theta_i - \theta_k)]; \quad j = i \quad (3) \end{aligned}$$

$$\begin{aligned} \frac{\partial P_i}{\partial V_j} &= V_i Y_{ij} \cos(\theta_i - \theta_j - \alpha_{ij}) \quad (4) \\ &= V_i Y_{ij} [\cos(\theta_i - \theta_j) \cos \alpha_{ij} + \sin(\theta_i - \theta_j) \sin \alpha_{ij}] \\ &= V_i [G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j)]; \quad j \neq i \quad (5) \end{aligned}$$

$$\frac{\partial P_i}{\partial V_i} \approx 0 \quad ve \quad \frac{\partial P_i}{\partial V_j} \approx 0 \quad \Rightarrow \quad J_2 \approx 0 \quad (6)$$

$$\frac{\partial Q_i}{\partial \theta_j} = \sum_{k=1, k \neq i}^n V_i V_k [G_{ik} \cos(\theta_i - \theta_k) + B_{ik} \sin(\theta_i - \theta_k)]; j = i \quad (7)$$

$$\frac{\partial Q_i}{\partial \theta_j} = -V_i V_j [G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j)]; j \neq i \quad (8)$$

$$\frac{\partial Q_i}{\partial \theta_i} \approx 0 \quad ve \quad \frac{\partial Q_i}{\partial \theta_j} \approx 0 \quad \Rightarrow \quad J_3 \approx 0 \quad (9)$$

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} J_1 & 0 \\ 0 & J_4 \end{bmatrix} \begin{bmatrix} \Delta \theta \\ \Delta V \end{bmatrix} \quad (10)$$

In this study, dynamic load flow analysis is performed under discontinuously distributed generation, and variable power demand situations and hourly changes of electrical parameters of busbars are calculated. Thus, the effects of distributed generation, which is formed from sources showing production discontinuity, on the voltage and power factor stability of busbars can be analyzed in the face of changing power demands.

B. Theoretical Background

Compensation is made by the help of static controllers and power electronics elements to improve the controllability of the power transferred by the power transmission lines and to provide the reactive power demand of the system quickly. Shunt reactive compensator devices can be designed with switching type converters. FACTS devices can produce and consume reactive power utilizing switched converter circuits without the need for capacitor or reactor groups in the compensation of transmission lines. Sine FACTS devices, which constitute modern compensation methods, react in a short time, the controllability of each phase separately, can compensate unbalanced loads, the use of these devices is an essential [20]. These devices increase the stability limits of the transmission lines when used properly. Today, many power flow controllers have been developed under the name FACTS. The most commonly used ones are; Static Var Compensator (SVC), Thyristor Controlled Series Capacitor (TCSC), Static Compensator (STATCOM), Combined Power Flow Controller (UPFC), Phase Shifter and Static Synchronous Serial Capacitor (SSSC).

The SVC is part of a family of FACTS devices that can be connected in parallel to the power system to generate or consume reactive power to control power system parameters, such as voltage. Its primary purpose is to produce and consume fast effective, precise and adjustable continuous reactive power to the system, have a high response time, operate in an unlimited range, are safe and have operational flexibility [21]. SVC is also used for dynamic power factor correction when the demand for reactive power in large industrial plants is high, SVC increases the power factor of the plant, minimizes voltage fluctuations at the input of the plant and reduces the operating costs of the plant. It's mainly used in SVC power systems for the voltage control and system stability improvement.

The operating principle of the SVC element is based on obtaining shunt impedance of variable value by inserting and removing the capacitors and reactors into the network depending on the calculated trigger angles. Reactive power can be adjusted in a wide range from maximum capacitive reactive

power value to maximum inductive reactive power value in the busbar to which static VAR generator is connected with an appropriate triggering [22]. The inductance determines the capacitive or inductive operation of the device. The value of inductance is determined by the following equation [23].

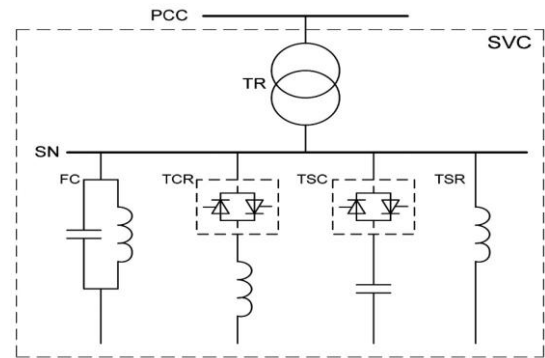


Fig.1. SVC Connection Structure [24]

STATCOM, known as Advanced Static Var Compensator (ASVC), is a FACTS controller, which is controlled to draw reactive current from the power system and connected to an inverter between a dc energy storage element and a three-phase system. A shunt is connected to the STATCOM transmission line, regulating the voltage of the transmission line at the connection point by drawing a controlled reactive current from the STATCOM transmission line [25]. A STATCOM controller consists of a connection transformer, voltage source inverter, and DC energy storage element (Fig.2.). Since the energy storage element is a tiny capacitor, it can only exchange reactive power.

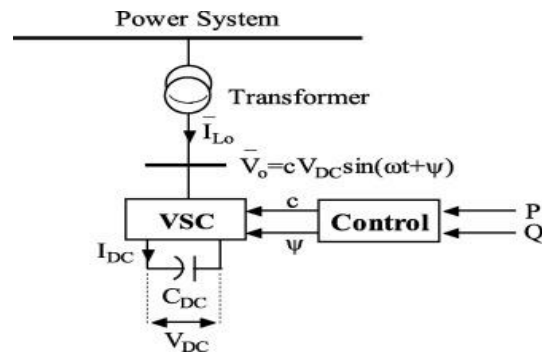


Fig.2. STATCOM circuit diagram [26]

The controller used in the simulation is the AC voltage control mode. The power is divided into two main parts. One is for an angle order, and another is for the rule of modulation index. The angle is ordered in such a way that the net real power absorbed from the line by this shunt FACTS device is equal to the losses of the converters and the transformer only. The remaining capacity of this shunt converter can be used to exchange reactive power with the line so to provide VAR compensation at the connection point. The reactive power is electronically provided by the shunt converter and the active power is transmitted to the DC terminals. The shunt converter reactive current is automatically regulated to maintain the transmission line voltage at the point of connection to a reference value. The line voltage and DC link voltage across

capacitor are measured to calculate the amount of reactive power to regulate the line voltage, and consequently, the modulation index is varied in such a way as to calculate reactive power can be injected at the point of connection, and thus the shunt FACTS device acts as a voltage regulator [27].

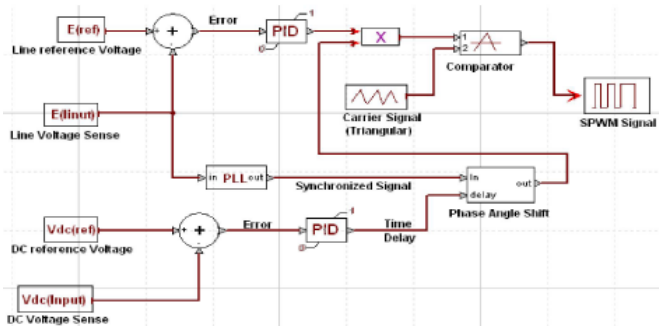


Fig. 2a. Block diagram of STATCOM Controller [27]

Varying the amplitude of the 3-phase output voltage of the voltage-driven inverter, STATCOM can be controlled whether to generate or draw reactive power. If the output voltage (V_0) of the inverter is greater than the ac system voltage (V_{ac}), then the ac current (I_{ac}) flows from the inverter to the ac system and generating reactive power via the transformer reactance. In this case, the inverter generates capacitive current for the ac system at an angle beyond its voltage. If the amplitude of the inverter output voltage is smaller than the ac system voltage, the ac current flows from the ac system to the voltage source inverter. And hence, the inverter draws an inductive current at an angle behind the voltage and consumes an inductive reactive power. If the output voltage of the inverter and the amplitude of the ac system voltages are equal, there will be no ac current flow from the inverter to the ac system. In short, the inverter will not produce or consume reactive power [28].

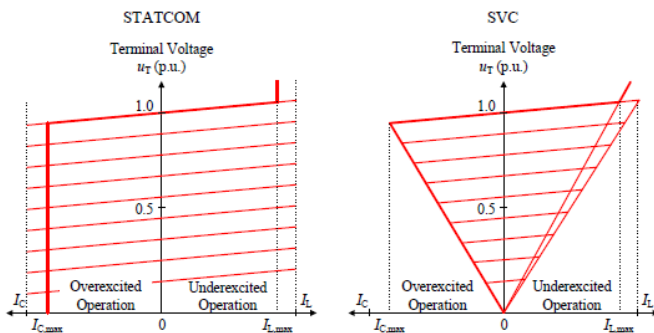


Fig.3. STATCOM, SVC, V-I characteristics [29]

Fig. 3. shows the SVC and STATCOM V-I characteristics. At low voltages, STATCOM current supply capacity is much better than SVC. STATCOM can provide either full capacitive or full inductive output current at any system voltage. The amount of reactive power compensation provided by STATCOM is higher than the amount of reactive power provided by SVC. Since at a low voltage level in the SVC, reactive power decreases proportionally with the square of the voltage, while STATCOM decreases linearly with the voltage.

This makes STATCOM's reactive power more controllable than the SVC [30].

III. SIMULATION STUDIES

A. Modeling

The realistic field data of the West Mediterranean Denizli energy transmission model is used to model the pilot power system region. The energy transmission model was created using the Dig-SILENT Power Factory program (Fig. 14). FACTS devices are connected to transformer substations and load flow analysis and voltage fluctuation controls are implemented. SVC and STATCOM which are FACTS controllers were connected to the power grid and analyses were carried out for various situations such as voltage drop and instantaneous power generation. The effects of the results on increasing the reliability and stability of the system and limiting the effects of faults and hardware defects were investigated. Moreover, increasing the stability of the voltage and power control of the Turkish interconnected system leads to more continuous, efficient, quality and economical way of providing electrical energy. Since the SVC and STATCOM devices are connected in parallel to the electrical power system, no auxiliary virtual bus is needed. In the modelling, lower and upper voltage limit values are defined as (0.9-1 pu) and (1-1.1 pu) for the 160 kV (1 pu) base voltage value. Capacitor banks, SVC and STATCOM devices are modelled separately in Denizli-1' bus bar. The simulation study was carried out with balanced 3-phase system voltages during 4-second steps over a total of 20-second operation duration in the case of one of the transmission lines (Denizli-2, Denizli-4 line-2a) of the Western Mediterranean region is out of order. The loading range of line and busloads of the energy transmission system varies between 15% and 20%, voltage is 1 pu, and the system production is 9,12% for G1 (Generator 1), 68,41% for G2 (Generator 2), and 74,39% for G3 (Generator 3). Voltage simulation values of Denizli-1, Denizli-4, and Acarsoy. In case of steady state or transient failures, the voltage fluctuations of the energy transmission system network of Denizli-1, Denizli-4, Acarsoy, and Mains bus bars have been analyzed.

B. Simulation Results

The effects of Capacitor Banks, STATCOM and SVC on the stability limits of the system against voltage collapses and also the elimination of voltage oscillations after the disturbing effects such as faults have been evaluated. The optimum reactive power value supported from capacitor banks are obtained as 17.2 MVar. Therefore, capacitor banks are adjusted to inject 18 MVar to the power network. The line voltage stabilities improve by the help of FACTS devices (Fig.4.). STATCOM provides more stability and also faster reaction time compared to SVC device (Fig.5.).

The voltage-loading curves for the most critical bus bars with voltage regulation problems have been evaluated in the modelled Western Mediterranean electricity transmission system. When SVC and STATCOM systems are inserted to the power system, their contribution to the reactive power has been examined and to what extent the system increases the loading limits in case of failure of the lines supplying Denizli-1 bus bar has been investigated (Fig. 6.). The load flow study has been

implemented at the fault case and Denizli-1 bus bar is obtained as the most critical bus bar. The voltage stability fluctuation has been shown when SVC or STATCOM is added to Denizli-1 bus bar in case of different disturbing effects in the system.

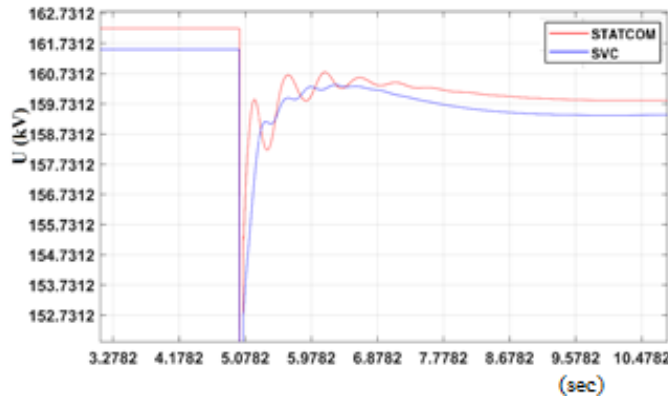


Fig. 4. Voltage fluctuations at the usage of STATCOM and SVC in case of line faults.

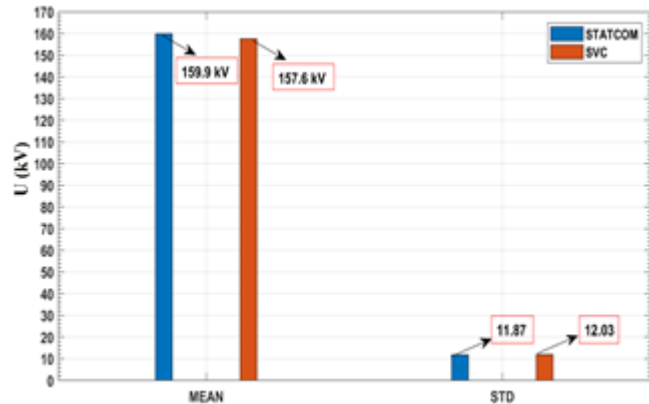


Fig.5. Mean and standard deviation values of bus voltages at the usage of STATCOM and SVC in case of faults

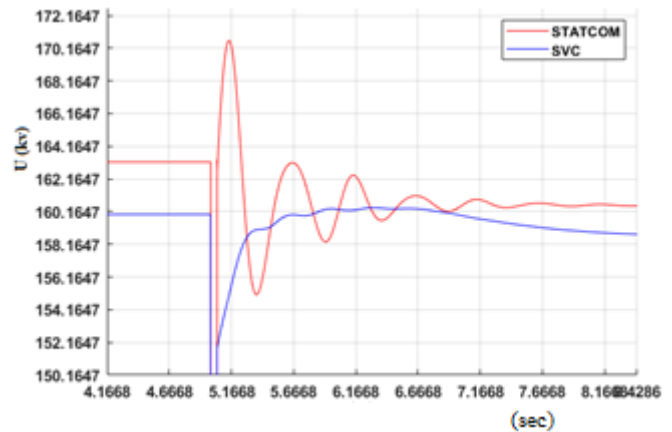
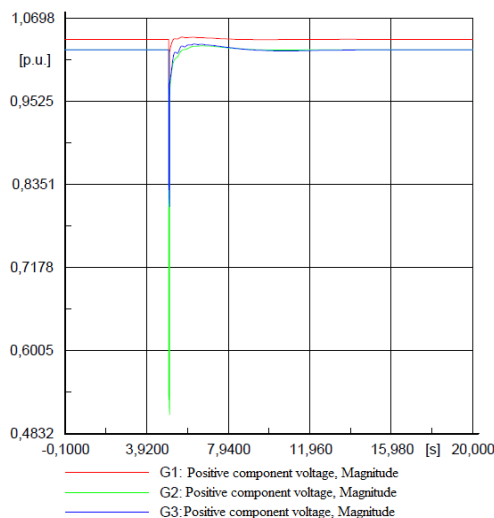


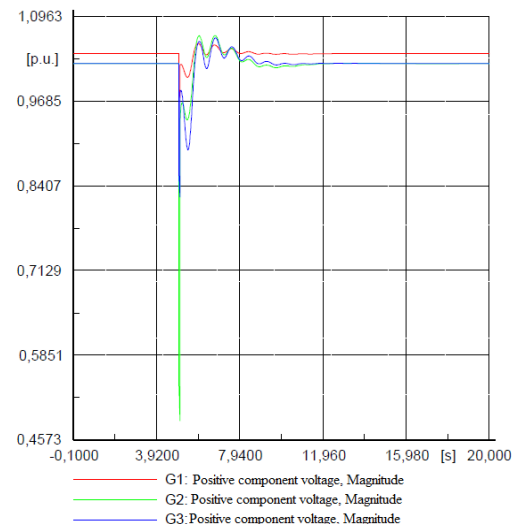
Fig. 6. Loading limits of the critical bus voltages at the usage of STATCOM and SVC in case of faults

As a result of load flow analysis, SVC and STATCOM were added in parallel to the required bars, and by providing voltage regulation, the points where the system suffered voltage collapse were identified and the effectiveness of FACTS devices at this point was evaluated. The SVC and STATCOM controllers have been significantly increased the reliability of the system against voltage collapses. STATCOM shows better performance and operates more stable in reactive power loading situations, and also provides the best power transfer in terms of active and reactive power values according to the simulation results. Required reactive power compensation is rapidly achieved with STATCOM, and despite sudden load changes and different load characters, it provides the desired reference voltage value of the bus bar.

When the rotor angle and rotor speed oscillations of the generators are examined, it is seen that their contribution to the oscillation suppression improves the generator voltage stability and helps to increase the reliability and capacity of the power system (Fig. 7. & Fig. 8a, 8b). With the use of STATCOM, inductive and capacitive energy is given to the pilot power system, making the system work much more stable.

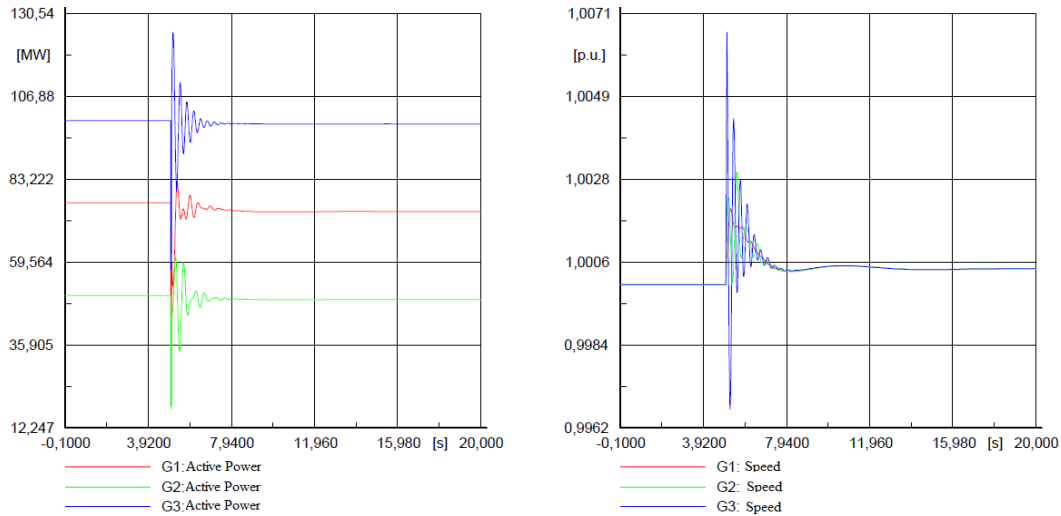


(STATCOM)

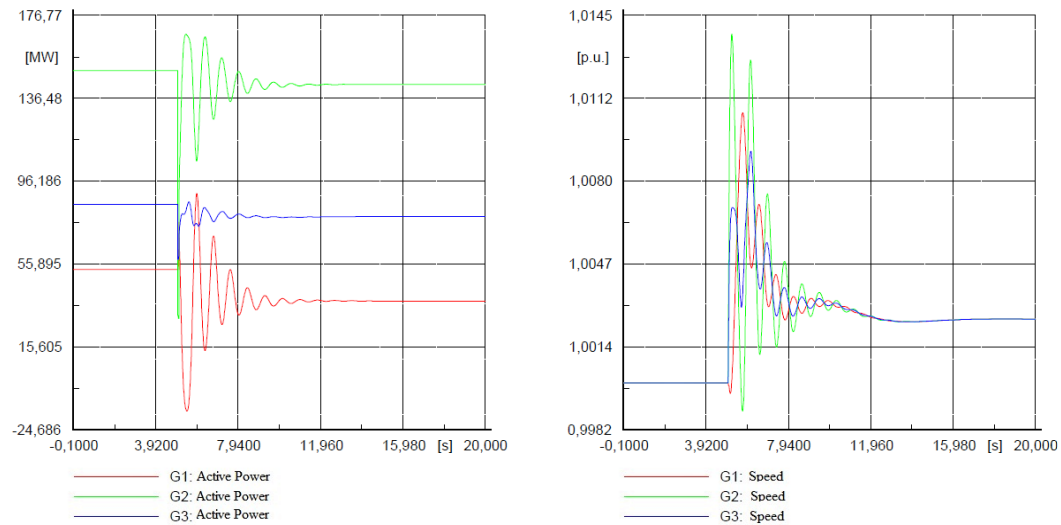


(SVC)

Fig.7. Generator voltages at the usage of STATCOM and SVC



(a) STATCOM
 Fig. 8.a. Rotor angle and rotor speed oscillations of generators at the usage of STATCOM



(b) SVC
 Fig. 8.b. Rotor angle and rotor speed oscillations of generators at the usage of SVC

C. Factors Causing Technical Losses in Electrical Grids

The characteristics of a power system change with time as the load grows and production increases. If the transmission facilities are not sufficiently developed, the power system becomes weaker against the steady state, and transient state problems and the stability limits are narrowed [31]. Transmission system limitations may include one or more characteristics such as steady state power transfer limit, voltage stability limit, dynamic voltage limit, transient stability limit, power system surge damping limit, loop flow limit, thermal limit, and short circuit current limit. Each transmission barrier or regional restriction has one or more level problem stages [30]. In each application, a FACTS controller is used to alleviate certain transmission limitations, to facilitate distribution of production, and to provide adequate system reliability [32]. Losses in electrical systems can be analyzed in

three main categories as production, transmission, and distribution losses. The majority of electrical losses occur in transmission and distribution systems. Approximately 6% of the generated electricity is due to internal loss and internal consumptions [33]. Therefore, the reduce of electrical losses is of great importance for the efficiency of electrical energy systems. Reasons for technical losses in electrical systems can be categorized as harmonic distortion, improper grounding of consumer loads, overloaded and long phase lines, voltage drops and use of non-standard equipment, reactive power drawn from the electrical grid [34]. Fig. 9 and 10 denote that STATCOM technology automatically activates and deactivates this process much faster and more stable than SVC devices, and hence minimizing losses and increasing the operating limits of the modelled transmission system. STATCOM, which has a flexible working structure, is more effective than SVC in reducing losses.

The effects of active and reactive power flow and the critical power system parameters such as total transmission losses on the transmission line are investigated when FACTS devices are inserted. In the usage of SVC and STATCOM, the differences of consumed and generated average active power were 0.9361 and 0.645 MW/h respectively (Fig.9. & 10.). STATCOM provides 31.09 % more achievement with respect to SVC device, and this corresponds to 0.2911 MW/h less generation and consumption loss.

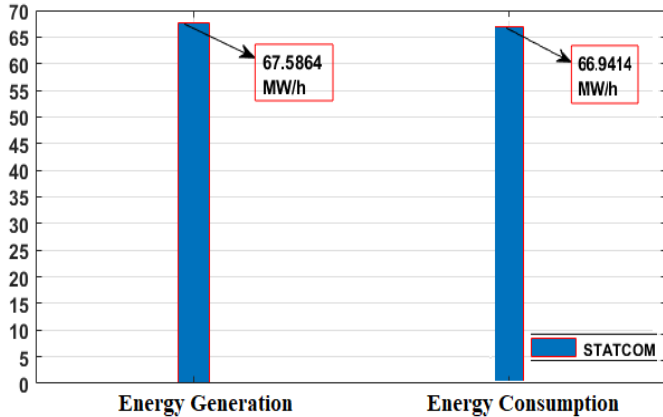


Fig.9. Energy production and consumptions at the usage of STATCOM

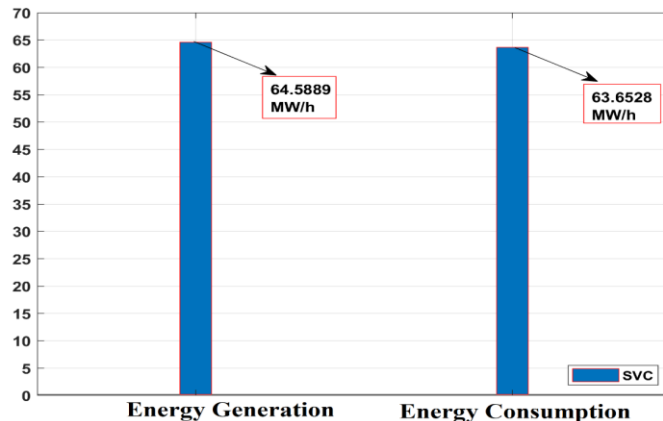


Fig.10. Energy production and consumptions at the usage of SVC

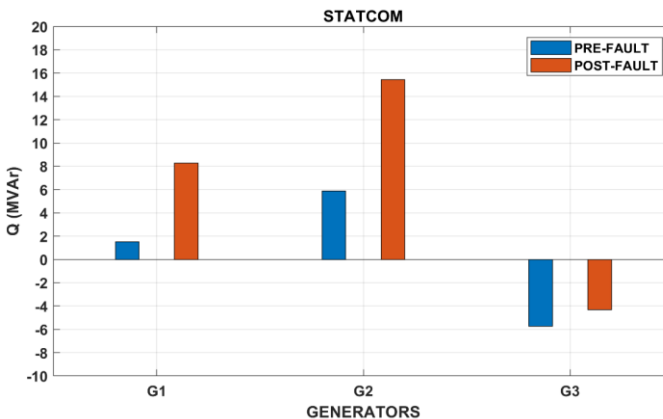


Fig.11. Generator Reactive Power Generations (STATCOM) Before/After Line Faults

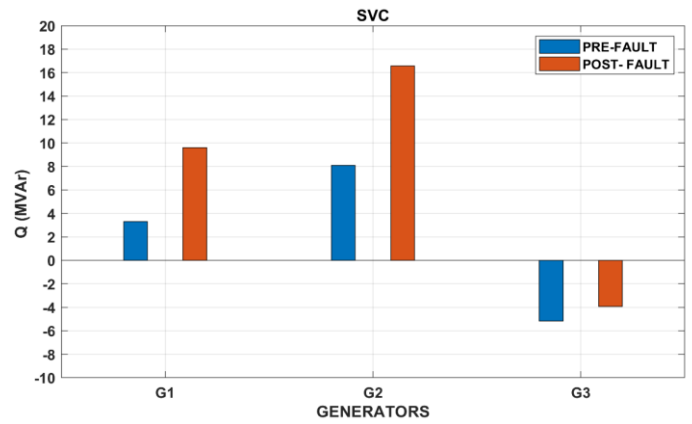


Fig.12. Generator Reactive Power Generation (SVC) Before/After Line Faults

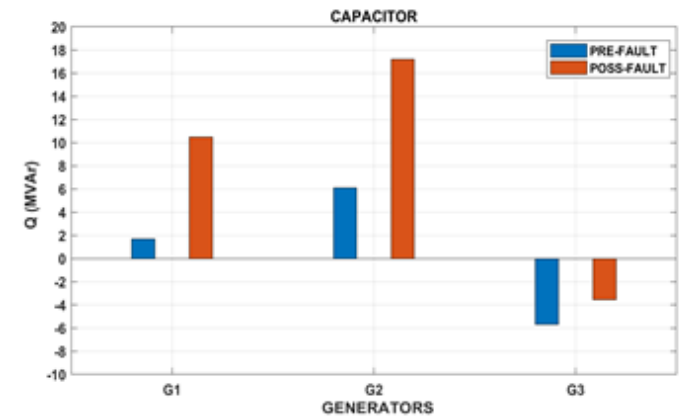


Fig.13. Generator Reactive Power Generation (Capacitor) Before/After Line Faults

In the simulation study, line faults resulted in critical increases in generator loadings. Reactive power loadings may be reduced by the usage of capacitor banks, SVC, and STATCOM. The total reactive power generation of three generators after the faults are 24.139, 22.249, and 19.389 MVar for capacitor banks, SVC and STATCOM, respectively. The results depict that the lowest reactive power necessity of the power network before/after the line faults occurs when STATCOM device is used (Fig. 11., 12. & 13.).

The comparison of reactive compensation devices such as capacitor banks, SVC and STATCOM on active and reactive power loadings and reactive power losses are shown by Table I. Active power loadings, capacitive reactive loadings, and reactive power losses of the transmission lines have been reduced by SVC/STATCOM devices 0.24%, 0.55%, 1% respectively. SVC and STATCOM devices decrease the loadings and also reactive power losses more than capacitor banks. Considering the all Turkish electrical power system, these reductions in losses are going to constitute a very substantial economic benefit. Voltage stability of the grid was provided as a result of STATCOM and SVC use of the power system model of the Western Mediterranean region electricity power transmission system of the Denizli region. Fluctuations of voltage stability in case of system overload or failure have been analyzed. Voltage and power stability values of the lines, and voltage values of load bus bars have been obtained for the generated power system model. Bus bar voltages are shown by

using Capacitor, SVC, and STATCOM devices before and after faults (Tables II and III). As can be seen from the results, the STATCOM device regulates the voltage much better than SVC

and capacitor banks. Bus bar voltages are hardly affected by the fault thanks to the use of STATCOM (Table III).

TABLE I
TRANSMISSION LINE STATUS (AT THE USAGE OF FACTS DEVICES) AFTER LINE FAULTS

Lines	ACTIVE POWER LOADING (%)		CAPACITIVE REACTIVE LOADING (MVar)		REACTIVE POWER LOSS (MVar)	
	CAPACITOR ENABLED	SVC / STATCOM ENABLED	CAPACITOR ENABLED	SVC / STATCOM ENABLED	CAPACITOR ENABLED	SVC / STATCOM ENABLED
Line	7.99699	7.99706	11.01927	10.99988	-10.17619	-10.15604
Line (1)	1.91172	1.91000	11.10064	11.08061	-11.10064	-11.08061
Line (2)	15.43685	15.4004	7.91846	7.87871	-4.49485	-4.45075
Line 1	30.22983	30.36543	8.97983	8.93827	3.32629	3.45375
Line 2a	9.30884	9.22839	1.54765	1.54355	-1.33054	-1.32984
Line 2b	9.13218	9.06786	13.81726	13.74428	-12.07093	-11.99212
Line 3	11.42976	11.36274	7.30717	7.29224	-6.06996	-6.06952
Line 4	25.05796	25.10335	10.30962	10.29939	-0.41096	-0.37028
Line 5	12.48466	12.4395	18.51381	18.49452	-15.12844	-15.12671
Line 6	19.07388	18.85184	15.77896	15.44046	-5.70820	-5.39699

TABLE II
BUS VOLTAGES WHEN FACTS DEVICES OPERATES (THERE IS NO FAULT)

Bus ID	CAPACITOR ENABLED	SVC ENABLED	STATCOM ENABLED
DENİZLİ-1 (kV) (pu)	161.9342 1.01208	160.0 1.0	160.0 1.0
DENİZLİ-2 (kV) (pu)	163.2398 1.02024	163.1085 1.01942	163.1085 1.01942
DENİZLİ-3 (kV) (pu)	158.4197 0.99012	158.3302 0.98956	158.3302 0.98956
DENİZLİ-4 (kV) (pu)	161.4339 1.008961	161.0137 1.00633	161.0137 1.00633
ACARSOY (kV) (pu)	165.177 1.03235	165.08 1.03175	165.08 1.03175
NETWORK (kV) (pu)	160.8582 1.005361	160.7506 1.00469	160.7506 1.00469

TABLE III
BUS VOLTAGES WHEN FACTS DEVICES OPERATES (AFTER LINE FAULT OCCURS)

Bus ID	CAPACITOR ENABLED	SVC ENABLED	STATCOM ENABLED
DENİZLİ-1 (kV) (pu)	157.8052 0.98628	160.0 1.0	160.0 1.0
DENİZLİ-2 (kV) (pu)	162.1797 1.01362	162.3029 1.01439	162.3029 1.01439
DENİZLİ-3 (kV) (pu)	157.8219 0.98638	157.9073 0.98692	157.9073 0.98692
DENİZLİ-4 (kV) (pu)	158.2907 0.98931	158.8408 0.99275	158.8408 0.99275
ACARSOY (kV) (pu)	164.3486 1.02717	164.4754 1.02797	164.4754 1.02797
NETWORK (kV) (pu)	159.9385 0.99961	160.0793 1.00049	160.0793 1.00049

IV. CONCLUSION

Implementing voltage and reactive power control by SVC/STATCOM controller in the power system, technical losses can be reduced, and electrical energy can be transmitted in a stable, controlled, and most economical way. Results denote that the sensitivity of voltage and reactive power control are improved and the production and consumption balance become

more stable by the usage of FACTS devices. When the modeled pilot region is evaluated in terms of oscillatory operation and voltage drop, FACTS devices can increase the maximum load point of the system against voltage collapse more than the continually operated capacitor banks. In case of failure of one of the lines in the network, it has been determined that FACTS devices can suppress oscillatory operation mode, which occurs at bus voltage, generator rotor speeds, and rotor angles. Results show that STATCOM and SVC systems increase the maximum load and limits, improve the stability limits in the case of a disturbing effect on the power system, and prevent the voltage collapse under overload conditions. The reactive power losses and correspondingly the electricity costs may be reduced by 1 % by SVC/STATCOM devices. *STATCOM controllers have %31.09 higher performance compared to SVC at assuring the energy production and consumption balance.*

ACKNOWLEDGMENT

This study has been supported by Akdeniz University Scientific Research Projects Coordination Department within the scope of the project no: FBA-2018-3792 for the possibility to complete a scientific research.

REFERENCES

- [1] Kalair A, Abas N, Kalair AR, Saleem Z, Khan N. "Review of harmonic analysis, modeling and mitigation techniques. *Renew Sustain Energy*" Rev 2017, 78 (10), pp. 1152–1187.
- [2] Martínez EB, Camacho CÀ. "Technical comparison of FACTS controllers in parallel connection" *J Appl Res Technol* 2017, 15(1), pp.36–44.
- [3] Jumaat SA, Musirin I, Baharun MM. "A voltage improvement of transmission system using static var compensator via matlab/Simulink" *Indones J Electr Eng Comput Sci* 2017, 6(2), pp. 1–17.
- [4] Fadaee M, Radzi MAM. "Multi-objective optimization of a stand-alone hybrid renewable energy system by using evolutionary algorithms: a review. *Renew Sustain Energy*" Rev 2012, 16(5), pp. 3364–2269.
- [5] Sadaipappan S, Renuga P, Kavitha D. "Modeling and simulation of series compensator to mitigate power quality problems" *Int J Eng Sci Technol* 2010, 2(12), pp. 7385–7394.

- [6] Liu L, Li H, Xue Y, Liu W. "Reactive power compensation and optimization strategy for grid-interactive cascaded photovoltaic systems" *IEEE Trans Power Electron* 2015, 30(1), pp. 188–202.
- [7] Darabian M, Jalilvand A. "A power control strategy to improve power system stability in the presence of wind farms using FACTS devices and predictive control" *Int J Electr Power Energy Syst* 2017, 85(2), pp. 50–66.
- [8] Sharaf AM, Gandoman FH. "A switched hybrid filter - DVS/green plug for smart grid nonlinear loads, in Smart Energy Grid Engineering" (SEGE), *IEEE International Conference on*; 17–19 Aug. 2015, pp. 1–6.
- [9] Sharaf AM, Khaki B. Novel "switched capacitor-filter compensator for smart gridelectric vehicle charging scheme" in *Proc. IEEE SGE*, Oshawa, Canada; Aug 2012, pp.1–6.
- [10] Abdelsalam AA, Desouki ME, Sharaf AM. "Power quality improvement using FACTS power filter compensation scheme" *J Electr Syst* 2013, 9(1), pp.86–96.
- [11] Mahela OP, Shaik AG, Gupta N. "A critical review of detection and classification of power quality events" *Renew Sustain Energy Rev* 2015, 41, pp.495–505.
- [12] Velamuri S, Sreejith S. "Power flow analysis incorporating renewable energy sources and FACTS devices" *Int J Renew Energy Res* 2017, 7(1), pp.452–458.
- [13] Crow ML. "Power quality enhancement using custom power devices" *IEEE Power Energy Mag* 2004, pp.2-50.
- [14] Hingorani N. G. "Flexible AC Transmission" *IEEE reprinted from IEEE Spectrum*, Vol. 30, No.4, 1993, pp. 40-45.
- [15] Cheng, H. In, I. and Chen S. "DC-Link Voltage Control and Performance Analysis of STATCOM", 2002.
- [16] Yang, Z. Shen, C. Zhang, L. Crow M. L. "Integration of a STATCOM and Battery Energy Storage", *IEEE Trans. on Power System*, Vol. 16, no. 2, May 2001, pp. 254-260.
- [17] Çötelî, "STATCOM ile Güç Akış Kontrolü", Yüksek Lisans Tezi Elektrik Eğitimi, Fırat Üniversitesi Fen Bilimleri Enstitüsü Elazığ, 2006.
- [18] Jaime Cepeda, Esteban Agüero, "FACTS models for stability studies in DIGSILENT Power Factory" *IEEE Transmission and Distribution Latin America*, 2014; DOI: 10.1109/TDC-LA.2014.6955182.
- [19] T. Sriyawong, P. Sriyanyong, P. Koseeyaporn and P. Kongsakom "A Modified Fast Decoupled Power Flow Algorithm" *International Energy Journal*: Vol. 6, No.1, Part 2, June 2005, pp. 2-95.
- [20] H. Feza Carlak, E Kayar, "TCR-TSC SVC Sistemleri Kullanılarak Enerji İletim Sistemlerinde Gerilim Regülasyonu Analizi", 4th International Mediterranean Science and Engineering Congress IMSEC Alanya, 2019, pp. 476-481.
- [21] Hingorani, Ng. Gyugyi, L. "Understanding FACTS: concepts and technology of flexible AC transmission systems" *IEEE Press*, New York, 2000.
- [22] U. Arifoğlu, "Güç Sistemlerinin Bilgisayar Destekli Analizi", Alfa Yayınları, İstanbul, 2002.
- [23] C.A. Canizares, T.F. Zeno, "Analysis of SVC and TCSC Controllers in Voltage Collapse", *IEEE Transactions on Power Systems*, Vol. 14, No. 1, February 1999.
- [24] R. Kowalak, "Resonant Conditions in a Node with an SVC Compensator," *Acta Energetica*, vol. 3, no.28, July 2016, pp. 70-75.
- [25] Schauder, C. and Mehta, H. "Vector Analysis and Control of Advanced Static VAR Compensators", *IEE Proceedings-C*, Vol. 140, No.4, 1993, pp.299–306.
- [26] S.M.Abd-Elazim, E.S.Ali, "Imperialist competitive algorithm for optimal STATCOM design in a multimachine power" *International Journal of Electrical Power & Energy Systems* Vol. 76, March 2016, pp. 136-146.
- [27] Md. Nazrul Islam, Md. Arifur Kabil, and Yashiro Kazushige "Design and Simulation of STATCOM to Improve Power Quality" *International Journal of Innovation and Applied Studies* ISSN 2028-9324 Vol. 3, No.3, July 2013, pp. 871-878.
- [28] Gyugyi, L. "Power Electronics in Electric Utilities: Static Var Compensators", *Proceedings of The IEEE*, vol. 76, no. 4, 1988, pp. 483-493.
- [29] Y. H. Song et al., *Flexible AC Transmission Systems (FACTS)*, London: IEEE, 1999.
- [30] Paserba J. "How FACTS Controllers Benefit AC Transmission System" *Trans. And Dist. Con. and Exp.*, *IEEE PES Volume 3*, 7-12 Sept. vol.3 2003, pp. 949 – 956.
- [31] Acha E., Fuerte-Esquivel C., Ambriz-Pe´rez R., Angeles-Camacho H.C., "FACTS Modelling and Simulation in Power Networks", John Wiley & Sons LTD, 2004.
- [32] Habur, K. D. O'Leary, "FACTS-Flexible Alternating Current Transmission Systems-For Cost Effective and Reliable Transmission of Electrical Energy", 2000.
- [33] Celal Yaşar, Yılmaz Aslan, Tarık Biçer, "Bir Dağıtım Tranformatörü Bölgesindeki Kayıpların İncelenmesi", *Dumlupınar Üniversitesi, Fen Bilimleri Enstitüsü Dergisi*, Sayı 22, Ağustos 2010.
- [34] J.P Navani, N.K Sharma, Sonal Sapra, "Technical and Non-Technical Losses in Power System and Its Economic Consequence in Indian Economy", *International Journal of Electronics and Computer Science Engineering*, ISSN: 2277-1956, 2012.

BIOGRAPHIES



HAMZA FEZA CARLAK received the B.S. degree in electrical engineering from the Istanbul Technical University, Istanbul, in 2000 and the integrated Ph.D. degree in electrical and electronics engineering from Middle East Technical University, Ankara, in 2012.

From 2003 to 2013, he was a Research Assistant in the Middle East Technical University. Since 2017, he has been an Assistant Professor with the Electrical and Electronics Engineering Department, Akdeniz University. He is the author of more than 30 articles, and has an invention on the medical imaging modality. His research interests include power systems, medical imaging, thermal imaging, image processing, renewable energy and artificial neural network algorithms.



ERGİN KAYAR received the B.S. degree in electrical and electronics engineering from the Akdeniz University, Antalya, in 2017 and he has been doing a master's degree at Akdeniz University, Department of Electrical and Electronics Engineering. He has been working in Turkey Electrical Communication Corporation (TEIAS) as an electrical engineer for 8 years. His research interests include power systems, renewable energy and energy transmission systems.

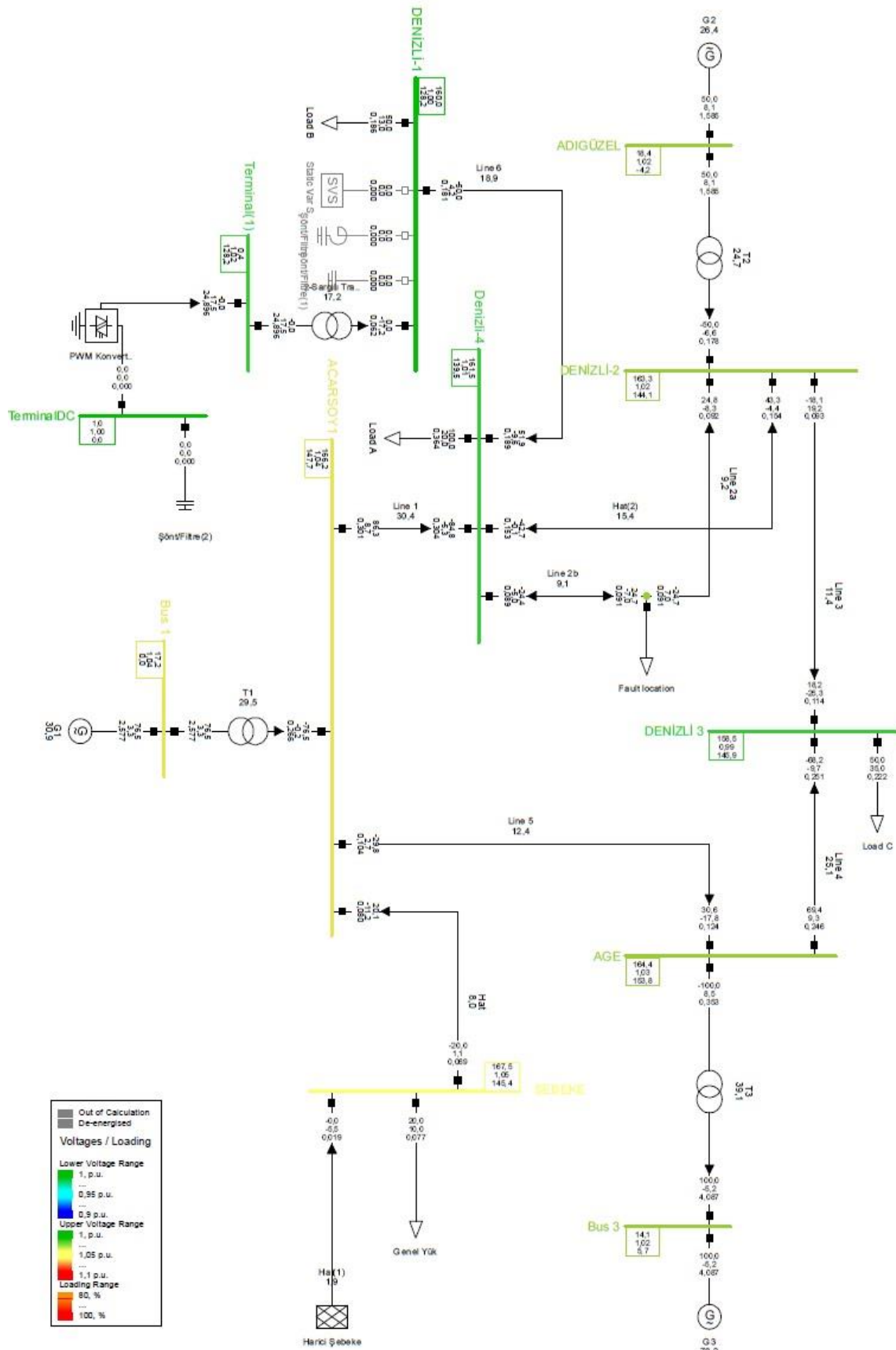


Fig.14. Western Mediterranean Energy Transmission System Network Model

Simultaneous Sensing of Dual Analyte with Photonic Crystal Fiber Based Liquid Sensor

H. ADEMGIL

Abstract— This article proposes three novel models of photonic crystal fiber (PCF) based dual liquid sensor. The performances of Z-Model, X-Model, and V-Model structures have been evaluated by applying the well-known full vectorial finite element method (FV-FEM) where boundaries are defined with perfectly matched layers (PMLs). The proposed models have been investigated with two types of analyte arrangements (A-Type and B-Type) in the core region. The birefringence and the sensitivity levels of proposed sensors have been studied where core holes are filled with water ($n=1.33$) and ethanol ($n=1.354$). Our numerical results have shown that X-Model and Z-Model of A-Type sensor is an ideal structure for water and ethanol sensing, respectively.

Index Terms— Birefringence, evanescent sensing, Optical Sensor, photonic crystal fiber

I. INTRODUCTION

OPTICAL FIBERS were initially established for telecommunication applications. However, their application areas are not limited with this sector. Progressively it can be employed in different areas competently. Photonic crystal fiber (PCF) is another class of optical fiber in light of the properties of photonic crystals [1]. They comprise of a silica core, encompassed by an occasional cluster of air-holes positioned alongside the full extent of the fiber structure. On account of its capacity to confine the light in hollow cores or with repression qualities unrealistic in traditional optical fiber, PCF is presently discovering applications in various sectors. These microstructured fibers have demonstrated virtues in various fields such as communication [1], nonlinear optics [2], high power technology [3] and sensing [4-5].

Features such as small dimension, light weight, electromagnetic interference resistance, potential for remote sensing and design flexibilities make PCFs stand out for sensing applications [6]. In PCFs, by modifying the structural design parameters such as hole to hole distances, air hole dimensions, air hole shape, and pattern geometry unique

propagation features can be achieved. Previously conducted studies have demonstrated that by choosing the suitable design parameters it is possible to control some key propagation properties such as dispersion, birefringence, nonlinearity, and confinement losses for wide wavelength range [5, 7-8].

Previous studies have shown that liquid and gas sensors are the most desirable sensing samples in industrial and biological practices [4, 6]. In recent years, especially PCF based liquid sensors have attracted great interest especially in bio/chemical solutions. Generally, water, alcohol or benzene types of materials are the major solutions utilized in these processes [9].


Hollow core Photonic Band Gap (PBG) and index guiding are the two well-known guiding mechanisms in PCF structures. It is well known that PBG type PCFs are an ideal candidate for gas sensing applications. Contrariwise, index guided PCF structures are appropriate for liquid sensing solutions. In PCF based sensors evanescent field is the vital component for sensing performance [10-11].

The optical mode of the index guided PCF structures are mainly confined at the centre of the core region. In index guided PCF based sensor designs, biological and chemical liquid samples (analyte) are hosted by air holes around the core region. In this regard, the sensing process is determined by the evanescent interaction between the guided mode and the analyte. Due to direct interaction of PCF core and sensing material (analyte) the evanescent PCF structures are considered as intrinsic sensors [10-11].

Initially, Monro et al. [11] proposed the evanescent-wave PCF sensor formation. Furthermore, Cordeiro et al. [10] has verified that the PCF structure with analyte filled holes are boosting the light mode interaction with the analyte to be detected.

Recent studies have demonstrated that different core and cladding configurations of PCF based sensors can be used for various sensing applications. Naeem et al. [12] have experimentally shown that two core PCF can be employed for strain and temperature sensing. Moreover, temperature and stress sensors with alcohol filled holes are experimentally demonstrated by Shi et al. [13] In addition, PCF structures for water-ethanol [14] and water, ethanol and benzene [15] solutions are proposed theoretically. On the other hand, PCF structures in the terahertz spectrum are proposed theoretically for sensing blood components [16] and analytes such as ethanol and benzene. Birefringence feature is commonly intended for polarization control in fiber based optical sensors and devices [17]. Additionally, highly birefringent fibers are emerging the polarization mode interference [18-19]. Such

HUSEYIN ADEMGIL, is with Department of Computer Engineering European University of Lefke, Lefke, Northern Cyprus, TR-10 Mersin Turkey, (e-mail: hademgil@eul.edu.tr).

 <https://orcid.org/0000-0002-2520-8567>

Manuscript received August 21, 2019; accepted September 25, 2019.
DOI: [10.17694/bajece.608334](https://doi.org/10.17694/bajece.608334)

fiber structures are ideal candidate for hydrostatic pressure sensors [18]. The numerical results in Islam et al. [19] have shown that high birefringent chemical sensor with relatively high sensitivities can be achieved with rectangular structured air holes. However, on the experimental point of view, due to sharp edges of rectangular air holes and complex design parameters this structure can be challenging [19].

The intention of this work is to propose highly birefringent PCF based sensor with dual analyte sensing ability. Unlike previous studies, in order to achieve birefringence, the symmetry of the PCF structure is broken with different analyte arrangements in the core region. Best to my knowledge, dual analyte sensing with such PCF core arrangements have not been studied and compared in the past.

II. DESIGN AND METHODOLOGY

Fig. 1 expresses the geometry of the proposed design of PCF based optical sensors. In this work, numerical simulations are engaged with commercial software Comsol Multiphysics. FV-FEM is engaged to analyze the various propagation characteristics of PCF based optical sensor [7, 11]. Perfectly match layers (PML) are engaged as boundary condition. The composition of the cladding air holes is hexagonal, where silica ($n = 1.45$) is used as background material. The cladding zone is formed by 3 rings of 54 circular air hole, where the hole diameter (d) is fixed to $2\mu\text{m}$, and hole to hole distance (Λ) of cladding air holes is set to $2.4\mu\text{m}$. Proposed PCF structure contains 19 analyte filled holes in the core region where center to center distance of core holes is set to $1.4\mu\text{m}$. As can be seen from the fig. 1, internal core holes are filled with two different analytes. Structures are denoted as *Z-Model*, *X-Model*, and *V-Model* where core hole diameters are fixed to $1.2\mu\text{m}$. The simultaneous detection of two liquid analytes is possible with this core arrangement. Due to the asymmetric refractive index (RI) arrangement in the core region, considered amount of birefringence can also be achieved. The numerical results are calculated for two types of analyte arrangement (*A-Type* and *B-Type*). Analyte fillings are detailed in Table 1, where, Water ($n_a = 1.33$) and Ethanol ($n_a = 1.354$) are considered.

The structural geometry of PCF is very precise as the cross sectional area is divided into sub-domains. Numerical simulations of proposed PCF models are carried out with triangular sub-domains. The FV-FEM procedure is applied with Maxwell's equations. The well-known wave equation is written as [7, 10-11, 20];

$$\nabla \times ([s]^{-1} \nabla \times E - k_0^2 n^2 [s] E) = 0 \quad (1)$$

the electric field vector is denoted by E , where the wave number in free space is $k_0 = 2\pi/\lambda$ and n is the refractive index of the silica background. The anisotropic PML is described with $[s]$ and λ is the operating wavelength.

Birefringence is one of the key parameters that is beneficial and advantageous for various applications of optical fibers. Birefringence is defined as the absolute difference of effective refractive index of x and y polarized fundamental modes [17-19]:

$$B = |n_{eff}^x - n_{eff}^y| \quad (2)$$

n_{eff}^x , and n_{eff}^y are denoting the effective refractive indices of polarization modes.

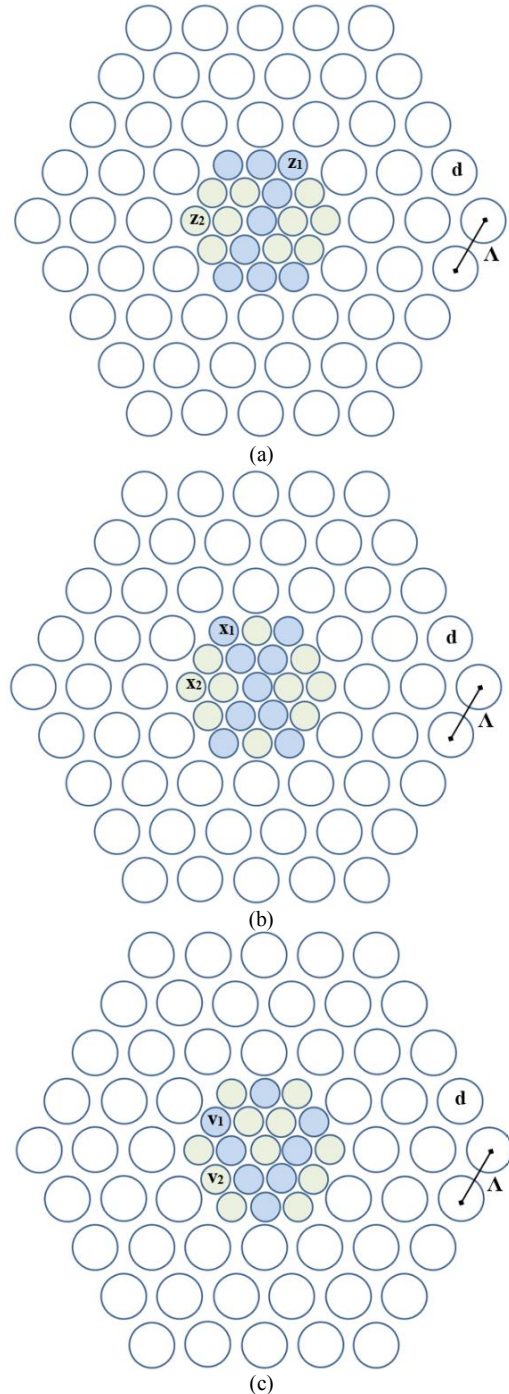


Fig.1. Schematic view of; (a) *Z-Model* PCF sensor, (b) *X-Model* PCF sensor and (c) *V-Model* PCF sensor.

On the other hand, the sensing performance of the liquid sensor can be predictable by calculating the relative sensitivity of the introduced PCF models. Briefly, the interaction volume of the guided mode and liquid analyte is indicating the relative

sensitivity. This key parameter can be mathematically described with the formula below [7, 10,19];

$$r = \frac{n_a}{n_{eff}} \times f \quad (3)$$

The analyte index and the effective refractive index of the guided light mode are denoted with n_a and n_{eff} , respectively. The ratio of the power flow through the specified domain is denoted as f , where it is calculated with the formula given below [7, 10, 19];

$$f = \frac{(sample) \int Re(E_x H_y^* - E_y H_x^*) dx dy}{(total) \int Re(E_x H_y^* - E_y H_x^*) dx dy} \times 100 \quad (4)$$

III. NUMERICAL RESULTS AND DISCUSSION

The dual analyte PCF models with analyte parameters presented in Table I are numerically investigated. It can be seen from the table that, in A-Type sensor models the primary core holes (X_1 , Z_1 , and V_1) and secondary core holes (X_2 , Z_2 , and V_2) are filled with 1.33 and 1.354 RI analyte, respectively. On the other hand, analyte filling of core holes are contrary in B- Type sensor models.

TABLE I
PCF SENSOR ANALYTE ARRANGEMENTS

Sensor Model		A – Type Sensor	B – Type Sensor
Z – Model	Z_1	1.33	1.354
	Z_2	1.354	1.33
X – Model	X_1	1.33	1.354
	X_2	1.354	1.33
V – Model	V_1	1.33	1.354
	V_2	1.354	1.33

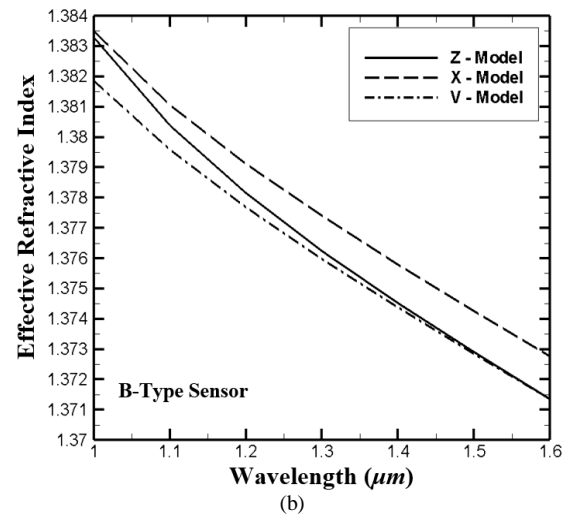
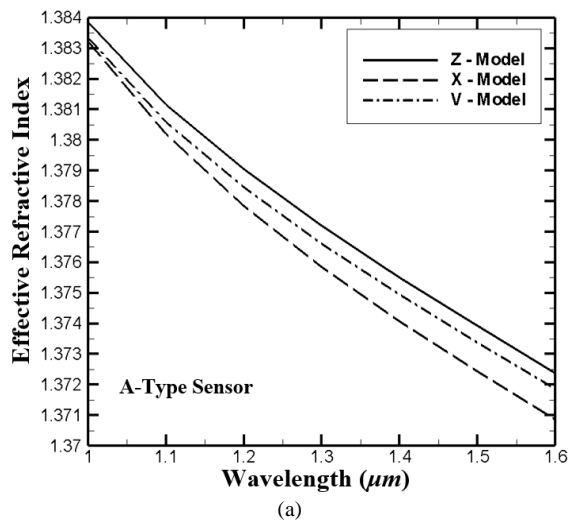


Fig. 2. Variations of effective refractive index for proposed Z, X and V models of PCF sensor.

The variations of effective refractive index of proposed structures relate with operating wavelength is presented in Fig. 2. These figures indicate that the effective refractive index of both types of sensors is decreasing with increasing wavelength. In A- Type sensor, effective index of Z-Model is higher than other models, where X-Model reaches higher levels in B-Type sensor. It is worth noting that the effective refractive index of both sensor types are varying between 1.384- 1.371 for wide wavelength range.

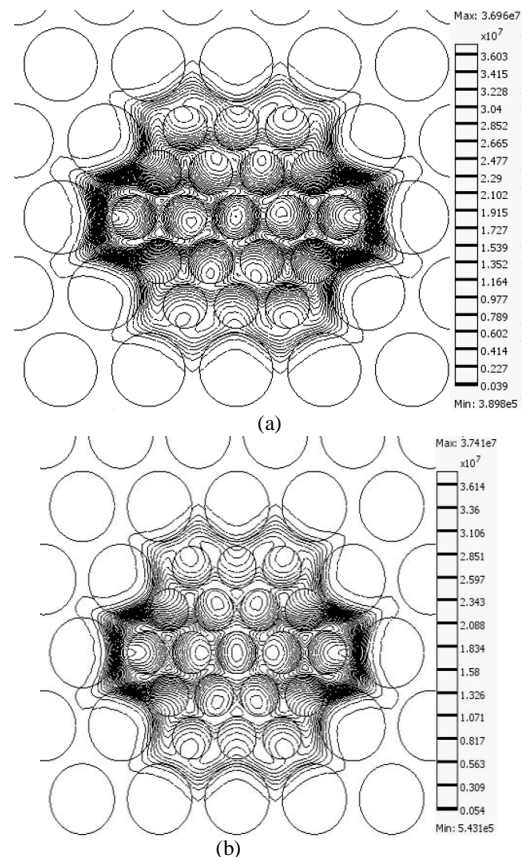


Fig. 3. Electric field distribution at $\lambda = 1.1 \mu m$, where, (a) Z- Model and (b) X- Model of A- Type PCF sensor.

The amount of light and liquid analyte interactions are presented in Fig. 3. Further, substantial interaction between the guided mode and analyte filled holes are illustrated. Figures show the field profile of the x -polarized HE_{11} mode for Z-Model and X-Model of A-Type PCF sensor. The amount of light mode and analyte interaction can be associated with the level of sensitivity coefficient of the PCF sensor models. The birefringence is the momentous propagation feature of the PCF based optical sensors. The asymmetric liquid fillings of core holes are triggering refractive index differences between the x - y polarized modes. As expected, considered amount of birefringence can be realised in Fig. 4 (a) & (b). The V-Model sensor contains the lowest birefringence values for both sensor types. On the other hand, X-Model and Z-Model show similar performances for both types. However, it is worth noting that birefringence level of proposed structures reaches 2 times improved performance with A-Type sensor. Fig.4 (b) indicates that, at shorter wavelengths birefringence of Z-Model is slightly higher than X-Model where this is contrary at longer wavelengths. This variation can be associated to the effective refractive index differences of x - y polarized modes where x - polarized mode of Z-Model is getting closer to y -polarized mode at longer wavelengths.

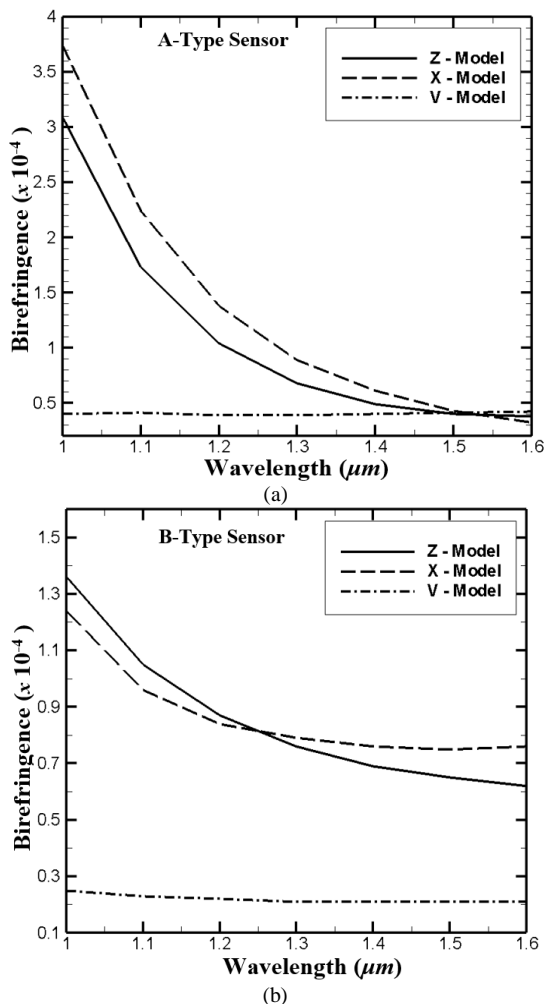


Fig. 4. Variations of birefringence for proposed Z, X and V models of PCF sensor.

Finally, Fig. 5 (a) & (b) shows the variations of sensitivity coefficient with respect to operating wavelength. It can be observed from both figures that the relative sensitivity is increasing at longer wavelengths. This phenomenon can be linked to improved interaction between analyte and guided mode. As expected [7, 10, 19], sensitivity coefficient level of higher index analyte (Ethanol) is moderately higher than lower index analyte (Water). Our proposed structures are almost 2 times more sensitive to Ethanol compared to water at broad wavelength range.

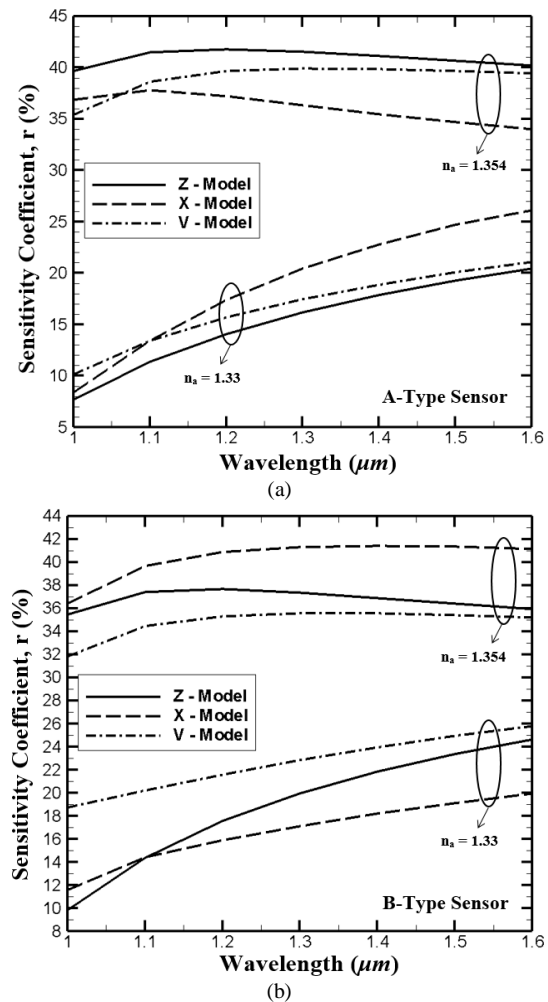


Fig. 5. Variations of sensitivity coefficients for proposed Z, X and V models of PCF sensor.

Table II briefly describes the overall performance of sensor types. It can be seen that X-Model of A-Type sensor shows overall better performance (high birefringence and high sensitivity) for water ($n_a = 1.33$) sensing. On the other hand, Z-Model of A-Type sensor is more desirable for ethanol ($n_a = 1.354$) sensing. More specifically, if we are only considering the sensitivity, Z-Model of A-Type and X-Model of B-Type sensor shows better performance for ethanol ($n_a = 1.354$) sensing. Conversely, X-Model of A-Type and V-Model of B-Type sensor for water ($n_a = 1.33$) sensing.

Previously, various kinds of PCF based structures are designed and analyzed with commercial software packages. Various

numerical methods [7, 10, 19, 21] are employed. These numerical modelling methods are very popular in computational electromagnetism, where they provide accurate and reliable results. However, on experimental point of view development of proposed models still needs to be discussed. In our proposed structures, for high flexibility and low fabrication cost, same diameter air holes with conventional hexagonal arrangement are used in the cladding area. For the same reason, diameter of core holes is kept fixed. Smooth operability of the optical sensor depends on effective analyte (Water & Ethanol) filling. The unique selective filling technique [22-23] permits infiltration of the air holes within PCF structure [24]. Also, previous experimental studies [25-26] have demonstrated that cladding and core holes can be liquid filled. Therefore, fabrication of proposed structures is viable with the current advances in nano-fabrication techniques.

TABLE II
OVERALL PERFORMANCE OF PROPOSED SENSORS

		Biref.	Analyte	Sensing
A-Type Sensor	Z - Model	High	1.33	Low
			1.354	High
	X - Model	High	1.33	High
			1.354	Medium
	V - Model	Low	1.33	Medium
			1.354	High
B-Type Sensor	Z - Model	Medium	1.33	Medium
			1.354	High
	X - Model	Medium	1.33	Low
			1.354	High
	V - Model	Low	1.33	Medium
			1.354	High

IV. CONCLUSION

In this work, numerical analyses of dual analyte PCF based sensor models have been investigated with different core arrangements. This study has been proposed in favour of liquid sensing. Birefringence of proposed sensor types reaches up to 10^{-4} levels, where a sensitivity coefficient of ethanol varies between 32%-42%. On the other hand, a sensitivity coefficient of water varies between 8% - 25%. Numerical results have shown that X-Model and Z-Model of A-Type sensor is an ideal structure for water and ethanol sensing, respectively. Thus, with such steady performance, the proposed dual analyte PCF structure can be employed for liquid sensing and other biological and chemical applications.

REFERENCES

- [1] J. C. Knight, T. A. Birks, P. S. J. Russell, and D. M. Atkin, "All-silica single-mode optical fiber with photonic crystal cladding", *Optics Letters*, vol. 21, no.19,1996, pp. 1547-1549.
- [2] P. St.J. Russell, "Recent Developments in Photonic Crystal Fibres", 2018 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR), pp. Tu2E.1-2.
- [3] X. Qi, S. Chen, Z. Li, T. Liu, Y. Ou, N. Wang, and Jing Hou, "High-power visible-enhanced all-fiber supercontinuum generation in a seven-core photonic crystal fiber pumped at 1016 nm", *Optics Letters*, vol. 43, no. 5, 2018, pp. 1019-1022.
- [4] X. Feng, W. Feng, C. Tao, D. Deng, X. Qin, and R. Chen, "Hydrogen sulfide gas sensor based on graphene-coated tapered photonic crystal fiber interferometer", *Sensors and Actuators B: Chemical*, vol. 247, 2017, pp. 540-545.
- [5] A. Yasli and H. Ademgil, "Geometrical comparison of photonic crystal fiber-based surface plasmon resonance sensors", *Optical Engineering* vol.57, no. 3, 2018, pp. 030801(1-10).
- [6] R. Gao, D. F. Lu, J. Cheng, Y. Jiang, L. Jiang, J. D. Xu, and Z. M. Qi, "Fiber optofluidic biosensor for the label free detection of DNA hybridization and methylation based on an in-line tunable mode coupler", *Biosensors & Bioelectronics*, vol. 86, 2016, pp. 321-329.
- [7] H. Ademgil, "Highly sensitive octagonal photonic crystal fiber based sensor", *Optik - International Journal for Light and Electron Optics*, vol. 125, 2014, pp. 6274-6278.
- [8] V.T. Hoang, B. Siwicki, M. Franczyk, G. Stepniewski, H.L. Van, V.C. Long, M. Klimczak, and R. Buczyński, "Broadband low-dispersion low-nonlinearity photonic crystal fiber dedicated to near-infrared high-power femtosecond pulse delivery", *Optical Fiber Technology*, vol. 42, 2018, pp. 119-125.
- [9] P. B. Kumar, K. Ahmed, D. Vigneswaran, F. Ahmed, S. Roy, D. Abbott, "Quasi photonic crystal fiber based spectroscopic chemical sensor in the terahertz spectrum: design and analysis", *IEEE Sensors Journal*, vol. 18, no. 24, 2018, pp. 9948 - 9954.
- [10] C. M. B. Cordeiro, M. A. R. Franco, G. Chesini, E. C. S. Barretto, R. Lwin, C. H. B. Cruz, and M. C. J. Large, "Microstructured-core optical fibre for evanescent sensing applications," *Optics Express*, vol. 14, no. 26, 2006, pp. 13056-13066.
- [11] T. M. Monro, D. J. Richardson, and P. J. Bennett, "Developing holey fibres for evanescent field devices", *Electronics Letters*, vol. 35, no. 14, 1999, pp. 1188-1189.
- [12] K. Naeem, Y. Chung and B. H. Kim, "Cascade Two- Core PCFs Based In-Line Fiber Interferometer for Simultaneous Measurement of Strain and Temperature", *IEEE Sensors Journal*, vol.19, no.9, 2019, pp.3322-3327.
- [13] F. Q. Shi, Y. Y. Luo, D. R. Chen, J. J. Chen, Z. J. Ren and B. J. Peng, "A Dual- Parameter Sensor Based on the Asymmetry of Alcohol Filling the Photonic Crystal Fiber in Sagnac Loop Temperature", *IEEE Sensors Journal*, vol.18, no.15, 2018, pp.6188-6195.
- [14] C. Zhou, H. K. Zhang, P. Song, J. Wang, C. G. Zhu, P. P. Wang and F. Peng, "Geometrically Structural Parameters Insensitive Fiber Sensor for Detection of Ethanol Concentration", *IEEE Photonics Technology Letters*, vol. 30, no. 23, 2018, pp. 2037-2039.
- [15] M. Suganthya, B. K. Paul, K. Ahmed, Md. I. Islame, Md. A. Jabind, A. N. Bahar, M. S. M. Rajan, "Analysis of optical sensitivity of analytes in aqua solutions", *Optik- International Journal for Light and Electron Optics*, vol. 178, 2019, pp. 970-977.
- [16] K. Ahmed, F. Ahmed, S. Roy, B. K. Paul, Mst. N. Aktar, D. Vigneswaran and Md. S. Islam, "Refractive Index- Based Blood Components Sensing in Terahertz Spectrum", *IEEE Sensors Journal*, vol.19, no.9, 2019, pp.3368-3375.
- [17] Y. Yang, F. Yang, H. Wang, W. Yang and W. Jin, "Temperature-Insensitive Hydrogen Sensor with Polarization-Maintaining Photonic Crystal Fiber-Based Sagnac Interferometer", *IEEE Journal of Lightwave Technology*, vol. 33, no. 12, 2015, pp. 2566-2571.
- [18] P. Hlubina, T. Martynkien, J. Olszewski, P. Mergo, M. Makara, K. Poturaj and W. Urbanczyk, "Spectral-domain measurements of birefringence and sensing characteristics of a side-hole microstructured fiber", *Sensors*, vol.13, no.9, 2013, pp.11424-11438.
- [19] M. S. Islam, J. Sultana, K. Ahmed, M. R. Islam, A. Dinovitser, B. W. Him Ng, and D. Abbott, "A novel approach for spectroscopic chemical identification using photonic crystal fiber in the terahertz regime", *IEEE Sensors Journal*, vol. 18, no. 2, 2018, pp. 575-582.
- [20] B. T. Kuhlmeier, B. J. Eggleton, and D. K. C. Wu, "Fluid-filled solidcore photonic bandgap fibers", *IEEE Journal of Lightwave Technology*, vol. 27, no. 11, 2009, pp. 1617-1630.
- [21] S. Padidar, V. Ahmadi, and M. Ebnali-Heidari, "Design of High Sensitive Pressure and Temperature Sensor Using Photonic Crystal Fiber for Downhole Application", *IEEE Photonics Journal*, vol. 4, no. 5, 2012, pp. 1590-1599.

- [22] F. Wang, W. Yuan, O. Hansen, and O. Bang, "Selective filling of photonic crystal fibers using focused ion beam milled microchannels", *Optics Express*, vol.19, no.18, 2011, pp.17585-17590.
- [23] Y. Huang, Y. Xu, and A. Yariv, "Fabrication of functional microstructured optical fibers through a selective-filling technique", *Applied Physics Letters*, vol. 85, no. 22, 2004, pp. 5182–5184.
- [24] S. Ertman, K. Rutkowska, and T. R. Wolinski, "Recent Progress in Liquid-Crystal Optical Fibers and Their Applications in Photonics", *IEEE Journal of Lightwave Technology*, vol. 37, no.11, 2019, pp. 2516-2526.
- [25] M. Luo, Y. Liu, Z. Wang, T. Han, Z. Wu, J. Guo and W. Huang, "Twin-resonance-coupling and high sensitivity sensing characteristics of a selectively fluid-filled microstructured optical fiber", *Optics Express*, vol. 21, no.25, 2013, pp. 30911-30917.
- [26] R.M. Gerosa, D.H. Spadoti, C.J.S. de Matos, L.S. Menezes and M.A.R. Franco, "Efficient and short-range light coupling to index-matched liquid-filled hole in a solid-core photonic crystal fiber", *Optics Express*, vol.19, no.24, 2011, pp.24687-24698.

BIOGRAPHIES



HUSEYIN ADEMGIL received his B.Sc. Degree in 2005 from Eastern Mediterranean University in Electrical and Electronic Engineering. He received the M.Sc. and PhD degree from the University of Kent, Canterbury, U.K., in 2006 and 2010, respectively. In 2009 July, he joined to the Kosova Telecommunication Company and assigned as a Project Manager. He has contributed in a number of projects such as FFTx and Planning and optimization of GSM Network. In September 2010, he joined European University of Lefke, where he is currently appointed as Head of the Computer Engineering Department. Dr. Ademgil was assigned to the TRNC Information and Communication Technologies Authority (BTHK) as a member of Board of Directors during 2012 May- 2019 May. He has been assigned on July 2019 to the Turkish Cypriot Electricity Authority (KIB-TEK) as a member of Board of Directors. He is a Senior Member of IEEE.

Novel OFDM System Using Orthogonal Pilot Symbols with Subcarrier Index Modulation

Y. ACAR and S. ALDIRMAZ ÇOLAK

Abstract — In this work, two new transmission schemes are proposed to increase the spectral efficiency of orthogonal frequency division multiplexing (OFDM). In practical OFDM systems, channel estimation is usually performed by employing pilot symbols which is based on inserting known symbols in the time-frequency domain. However, pilot symbol designing is one of the bottlenecks of OFDM systems limiting the increase of spectral efficiency. We apply new pilot design structures, which are use orthogonal Walsh-Hadamard codes. To increase the spectral efficiency of OFDM systems, we assign the extra bits to the index of the each orthogonal Walsh-Hadamard codes. Simulation and theoretical results show that proposed methods have better performance than conventional OFDM with higher spectral efficiency. Moreover, no more energy is spent for additional information carried in the indices compared to classical OFDM systems. As a result, proposed methods provide both spectral efficiency and energy efficiency.

Index Terms— Orthogonal frequency division multiplexing (OFDM), subcarrier index modulation (SIM), channel estimation, Walsh-Hadamard codes.

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM) has attracted a lot of attention for broadband wireless communication systems due to robustness against frequency-selective fading channels and simple implementation [1]. OFDM has been adopted by several popular standards such as long term evolution-advanced (LTE-A), worldwide interoperability for microwave access (WiMAX), local area network (LAN), and digital video broadcasting (DVB) [2-3]. The channel estimation is a crucial and challenging procedure for OFDM systems in practical scenarios especially for mentioned standards above. In [4], OFDM performance is presented for pedestrian and vehicular channel models, i.e., practical channels.

Channel estimation (CE) techniques can be roughly divided up into two groups, i.e., using blind-way or pilot symbol aided (PSA) [5, 6]. While blind channel estimation methods are based on second or higher order statistics; pilot-assisted

YUSUF ACAR, is with STM Savunma Teknolojileri Mühendislik ve Ticaret A.Ş., Ankara, Turkey, (e-mail: yusuf.acar@stm.com.tr).

 <https://orcid.org/0000-0002-3956-1097>

SULTAN ALDIRMAZ ÇOLAK, is with Kocaeli University, Kocaeli Turkey, (e-mail: sultan.aldirmaz@kocaeli.edu.tr).

 <https://orcid.org/0000-0001-7154-0723>

Manuscript received June 08, 2019; accepted September 16, 2019.

DOI: [10.17694/bajece.588919](https://doi.org/10.17694/bajece.588919)

channel estimation methods transmit a known signal, pilots, in both transmitter and receiver sides. Blind-way techniques have higher spectral efficiency compared to that of PSA; however their performance are poor due to the data interference. Moreover, their computational complexity is higher than that of PSA techniques. Several PSA-CE schemes for OFDM applications have been investigated [7, 8, 9, 10]. Since pilot symbols are transmitted instead of data symbols, spectral efficiency of PSA-CE methods is lower than blind-way CE techniques. Interested readers are referred to [11] and the references therein for a review of OFDM channel estimation techniques.

Spatial modulation (SM), which has been recently proposed for multi-input multi-output (MIMO) systems, is a very distinct and useful modulation scheme [12]. Using the idea in SM, it was demonstrated by sub-carrier index modulation (SIM) (i.e., OFDM-IM) that spectral efficiency could be increased for OFDM systems [13, 14, 15]. SIM concepts have attracted increasing attention as a promising technique for the next generation wireless communication systems in the last decade. SIM method allows to transmit additional bits to the conventional modulation schemes by mapping the data bits to the indices of the different medias. For more details on this important field, the interested researchers are referred to survey on SIM schemes in [16].

Recently, two blind CE techniques for OFDM systems have been proposed by inspired the SIM technique [17]. The positions of the pilot signals in the frequency domain of the OFDM signal (before IFFT operation) transmit additional data bits in addition to the modulated symbols. Thus, the spectral efficiency of the system is increased.

The Walsh-Hadamard matrix [18] or pseudo-codes are key components of modern information technology such as code division multiple access (CDMA) communication systems. Walsh-Hadamard matrices are square matrices whose entries are -1 or $+1$, and columns or rows are orthogonal to each other. Assume that we have $N_p \times N_p$ Walsh-Hadamard matrix as follows:

$$\mathbf{C} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,N_p} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,N_p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N_p,1} & c_{N_p,2} & \cdots & c_{N_p,N_p} \end{bmatrix}$$

where N_p denotes number of pilot symbols.

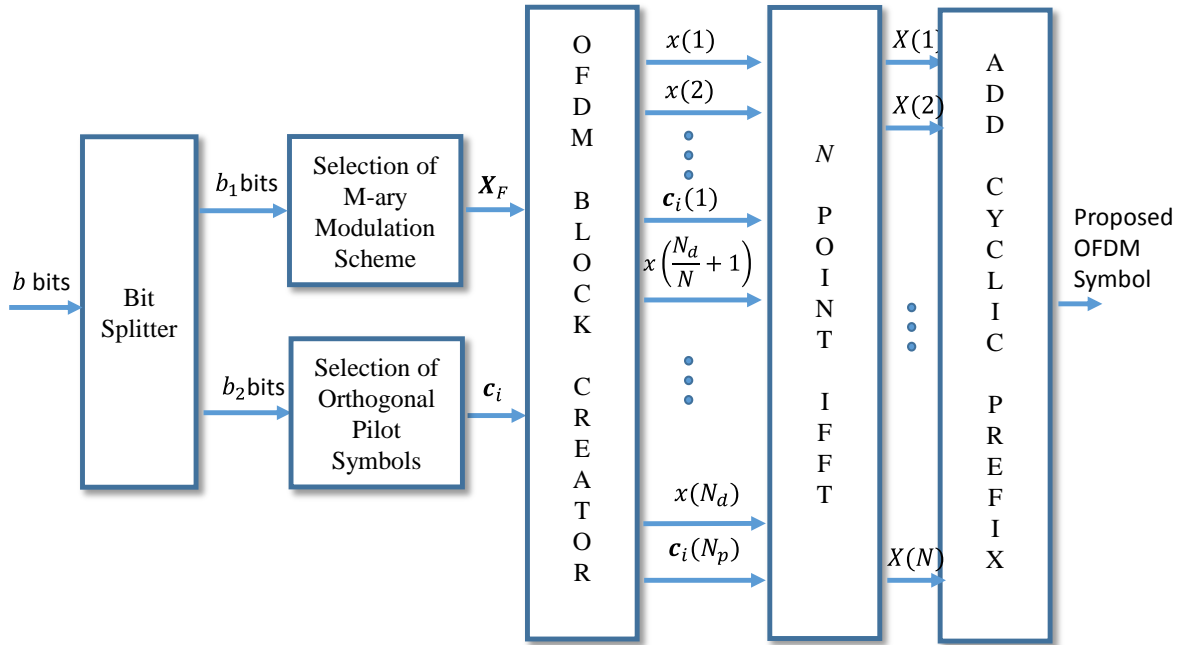


Fig.1. Block Diagram of the Proposed Methods

We can re-written Walsh-Hadamard matrix as $C = [c_1 \ c_2 \ \dots \ c_{N_p}]^T$ where $c_1 \ c_2 \ \dots \ c_{N_p}$ are $1 \times N_p$ orthogonal vector. Inspired by SIM concept and orthogonality of Walsh-Hadamard codes, we proposed two novel channel estimation techniques to increase the spectral efficiency of OFDM systems. The basic idea of the proposed techniques is to select pilot symbol vectors based on the incoming bits from possible pilot vector candidates.

In the proposed methods, to take advantage of orthogonality of Walsh-Hadamard codes, we assign additional bits to the each orthogonal code. For example, for Walsh-Hadamard codes matrix C with length $N_p = 4$, we simply assign the information bits in each row vector as given in Table 1. For instance, if incoming bits for orthogonal Walsh-Hadamard code word is $[0 \ 1]$, we select $c_3 = [+1 \ +1 \ -1 \ -1]$ as a pilot symbols in the related OFDM symbol. These designs also can be extending for higher N_p values. Proposed method is capable of enhancing the spectral efficiency without extra energy consumption compared to conventional OFDM by assigning extra bits to the index of orthogonal pilot symbols. As a result, the index of the Walsh-Hadamard codes convey an additional information and complete (realize) channel estimation. Moreover, proposed techniques do not need any prior knowledge at the receiver and extra energy consumption.

This paper is organized as follows. Section II demonstrates the principle of proposed schemes. Section III introduces the spectral efficiency of the proposed schemes. The receiver design and simulation results are given in Section IV and

Section V, respectively. Finally, conclusion is presented in Section VI.

II. PROPOSED METHODS

The general system block diagram is given in Fig.1. First, incoming b bits are split into b_1 and b_2 bits in each OFDM symbol. Then, b_1 bits are mapped onto the M -ary signal constellation and the remaining b_2 bits are used to select a pilot signal from a predefined set of orthogonal pilot vectors. We proposed two methods which are inserted the selected pilot tones considering the whole OFDM and subgroups. In the first system, only one pilot signal is selected according to the index bits (b_2), while in the second system, OFDM is divided into subgroups, and the pilot signal is selected according to the index bit for each subgroup. This allows more bits to be moved in an OFDM symbol. The proposed schemes are explained as follows:

A) **Orthogonal Pilot Code Index Modulation (OPC-IM):**

In the proposed Orthogonal Pilot Code Index Modulation (OPC-IM) scheme, $b_2 = \lfloor \log_2(N_{p_i}) \rfloor$ bits are assigned to the index of orthogonal pilot tone for each OFDM symbol where N_{p_i} is number of the pilot tone. Fig. 2 shows an illustrative frame structure of the proposed OPC-IM scheme where total number of subcarrier $N = 32$, total data symbols $N_d = 24$ and number of pilot tones $N_{p_i} = 8$. For this parameter set, additionally $b_2 = \lfloor \log_2(8) \rfloor = 3$ bits are carried. For instance, an incoming bits are $[0 \ 0 \ 0 \ 1 \ 0 \ \dots \ 1 \ 1 \ 0]$. First three bits $[0 \ 0 \ 0$

0] are mapped to orthogonal pilot code index, (*i.e.*, the code \mathbf{c}_1 will be used for pilot symbols according to the Table II) then the remained bits [1 0 \cdots 1 1 0] are used for mapping

TABLE I
A LOOK-UP TABLE FOR $N_p = 4$

Bits	Pilot Tone Codes
[0, 0]	$\mathbf{c}_1 = [+1, +1, +1, +1]$
[0, 1]	$\mathbf{c}_2 = [+1, -1, +1, -1]$
[1, 0]	$\mathbf{c}_3 = [+1, +1, -1, -1]$
[1, 1]	$\mathbf{c}_4 = [+1, -1, -1, +1]$

TABLE II
A LOOK-UP TABLE FOR $N_p = 8$

Bits	Pilot Tone Codes
[0, 0, 0]	$\mathbf{c}_1 = [+1, +1, +1, +1, +1, +1, +1, +1]$
[0, 0, 1]	$\mathbf{c}_2 = [+1, -1, +1, -1, +1, -1, +1, -1]$
[0, 1, 0]	$\mathbf{c}_3 = [+1, +1, -1, -1, +1, +1, -1, -1]$
[0, 1, 1]	$\mathbf{c}_4 = [+1, -1, -1, +1, +1, -1, -1, +1]$
[1, 0, 0]	$\mathbf{c}_5 = [+1, +1, +1, +1, -1, -1, -1, -1]$
[1, 0, 1]	$\mathbf{c}_6 = [+1, -1, +1, -1, -1, +1, -1, +1]$
[1, 1, 0]	$\mathbf{c}_7 = [+1, +1, -1, -1, -1, -1, +1, +1]$
[1, 1, 1]	$\mathbf{c}_8 = [+1, -1, -1, +1, -1, +1, +1, -1]$

according to M -ary modulation scheme. Therefore, in the proposed OPC-IM, in contrast to classical OFDM, pilot symbols are used to not only estimate the channel response, but also convey additional data bits through the index domain while maintaining same power consumption (with OFDM system with PSA-CE).

B) Subblock Orthogonal Pilot Code Index Modulation (SOPC-IM):

To increase the spectral efficiency of proposed OPC-IM method, an OFDM symbol is divided into subblocks with length $k = N/s$, where s represents number of subblock and N represents the length of the inverse fast Fourier transform (IFFT). Then, we perform OPC-IM method for each subblock. The process in all subblocks are independent and same of each other. An illustrative frame structure of the proposed Subblock Orthogonal Pilot Code Index Modulation (SOPC-IM) scheme where total number of subcarriers $N = 32$, total data symbols $N_d = 24$, number of pilot tones for each subblock $N_{p_2} = 4$ and number of subblock $s = 2$ is given in Fig. 3. As seen from this figure, there are two subblocks and each of them use four orthogonal pilot tones for CE. For example, assume that an incoming bit sequence of is [1 0 0 1 0 1 1 \cdots 1 1 0]. For the first subblock, the first two bits [1 0] are mapped to orthogonal pilot code index 3 (*i.e.*, the code \mathbf{c}_3 will be used for pilot symbols in first subblock) and then incoming

bits [0 1] are used for subblock 2 (*i.e.*, the code \mathbf{c}_2 will be used for pilot symbols in second subblock) according to Table 1. Therefore, the proposed SOPC-IM method has more spectral efficiency than OPC-IM thanks to independent pilot tone selection for each subblock. Finally, remaining bits [0 1 1 \cdots 1 1 0] are used for M -ary modulation scheme as usual.

III. SPECTRAL EFFICIENCY

The spectral efficiency of classical OFDM, without taking into account the cyclic prefix (CP) overhead, is given as follows:

$$\eta_{OFDM} = \frac{b_1}{N} = \frac{(N - N_{p_1}) \log_2 M}{N} \quad (1)$$

It is clear that, pilot tones cause the loss of the spectral efficiency. The spectral efficiency of the proposed OPC-IM and SOPC-IM are given in (2) and (3), respectively.

$$\eta_{OPC-IM} = \frac{b_1 + b_2}{N} = \frac{(N - N_{p_1}) \log_2 M + \log_2(N_{p_1})}{N} \quad (2)$$

$$\eta_{SOPC-IM} = \frac{b_1 + b_2}{N} = \frac{(N - kN_{p_2}) \log_2 M + k \log_2(N_{p_2})}{N} \quad (3)$$

Table III shows the spectral efficiency of the proposed methods for different modulation size. It is clear that the proposed SOPC-IM method has %16.66 more spectral efficiency than classical OFDM system with BPSK modulation scheme.

TABLE III
SPECTRAL EFFICIENCIES FOR $N=256$, $N_{p_1} = 64$, $N_{p_2} = 4$ AND $k=16$.

	OFDM	OPC-IM	SOPC-IM
BPSK	0.7500	0.7734	0.8750
QPSK	1.500	1.5234	1.6250
16-QAM	3.000	3.0234	3.1250

IV. RECEIVER DESIGN

Let \mathbf{x}_F be the transmitted OFDM symbol as follows:

$$\mathbf{x}_F = [\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{c}_i(1), \dots, \mathbf{x}(d), \mathbf{x}(d+1), \dots, \mathbf{c}_i(p), \dots, \mathbf{x}(N_d), \mathbf{c}_i(N_p)] \quad (4)$$

where $d = 1, 2, 3, \dots, N_d$, $p = 1, 2, 3, \dots, N_p$, \mathbf{x} is data symbol vector and \mathbf{c}_i is selected Walsh-Hadamard code. After applying IFFT, the OFDM symbol given as:

$$\mathbf{x}_F = \mathbf{W}_N^H \mathbf{x}_T \quad (5)$$

where \mathbf{W}_N is the FFT matrix with $\mathbf{W}_N^H \mathbf{W}_N = \mathbf{N} \mathbf{I}_N$. Then, guard interval (or CP) is added to \mathbf{x}_F . The signal is converted from

parallel to serial then the resulting OFDM symbol is transmitted through the channel. The vector of channel impulse response (CIR) coefficients are given as $\mathbf{h} = [h_1, h_2, \dots, h_{N_L}]^T$, whose elements have complex Gaussian distribution as $CN(0, 1/N_L)$ where N_L is the channel length. Hence, the received signal is given as

$$y_F(f) = h_F(f) + w_F(f) \tag{6}$$

where $w_F(f)$ and $h_F(f)$ are the noise samples and the channel fading coefficients in the frequency domain, respectively.

codes as follows: The proposed methods are first used to detect the index of Walsh-Hadamard codes as follows:

$$\frac{1}{N_p} \sum_{n=1}^{N_p} \mathbf{c}_i(n) \mathbf{c}_j(n) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{8}$$

After determining the orthogonal pilot code \mathbf{c}_j , the frequency response of channel at the pilot positions can be estimated by using least square (LS) method as $\hat{h}(n_c) = y_F(n_c) / \mathbf{c}_j(n_p)$. Finally, cubic SPLINE interpolation technique is used in the process of constructing the whole channel response [6]. Then the first b_2 bits are demodulated according to the detection

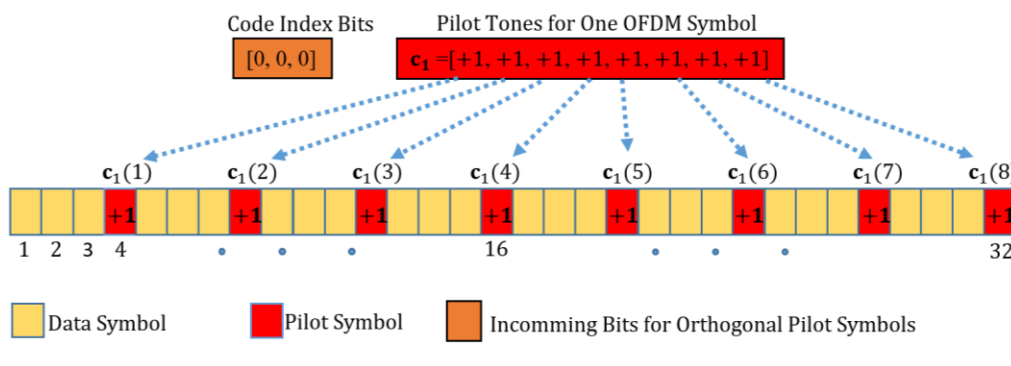


Fig.2. An illustrative frame structure of the proposed OPC-IM scheme with $N = 32$, $N_d = 24$ and $N_{p1} = 8$.

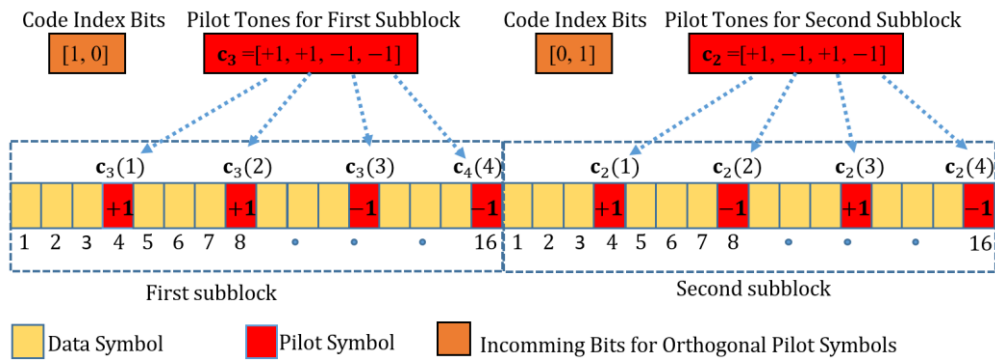


Fig.3. An illustrative frame structure of the proposed SOPC-IM scheme with $N = 32$, $N_d = 24$ and $N_{p2} = 4$ for each subblock.

Then, the received signals at pilot subcarriers can be expressed as follows:

$$y_F(n_c) = \mathbf{c}_i(n_p) h_F(n_c) + w_F(n_c) \tag{7}$$

where $n_p = 1, 2, \dots, N_p$, $n_c = N/N_p, 2N/N_p, \dots, N$, and \mathbf{c}_i is the selected orthogonal pilot symbol according to index bits b_2 . Then by correlating received signal (7) with Walsh-Hadamard matrix \mathbf{C} the orthogonal index of the orthogonal pilot code can be determined easily thanks to orthogonality of

result in (7) using the look-up table for index of codes. To obtain the transmitted information b_1 bits on each OFDM symbols, the receiver uses hard decision decoding for M -ary symbols. As a result, proposed methods not only complete channel estimation without a loss of power, but also transmit additional information bits without extra bandwidth.

V. SIMULATION RESULTS

Bit error rate (BER) performance of the proposed techniques are evaluated for different N_d and N_p values over

frequency selective Rayleigh fading channels with length $N_L = 10$. Signal to noise ratio (SNR) is described as E_b/N_0 where N_0 is the noise variance and E_b is energy per bit. We assume that the receiver perfectly determines the index bits of Walsh-Hadamard codes. The BERs of conventional OFDM and OPC-IM are given in Fig. 4 as functions of the SNR for PSA channel estimation. As shown in from Fig. 4, the proposed OPC-IM method achieves almost same performance as the classical OFDM with different modulation scheme such as BPSK, QPSK, 8-QAM, 16-QAM and 64-QAM. Fig. 5 illustrates the BER of SOPC-IM and conventional OFDM as functions of the SNR. As shown that proposed SOPC-IM method achieves better performance than the conventional OFDM for lower modulation size. Because, at lower modulation size, the percentage of the bits which are carried by the index of orthogonal pilot tones are higher than that of the lower modulation size.

Table IV shows the BER gain of proposed systems compared to the classical pilot-aided OFDM system at a specific E_b/N_0 value. The BER gain (ζ_ϕ) at $E_b/N_0 = 30$ dB is

$$\zeta_\phi = \frac{BER_{\phi, 30dB} - BER_{\phi, 30dB}}{BER_{\phi, 30dB}}$$

where ϕ denote classical pilot-aided OFDM system and the proposed system ($\phi \in \{OPC-IM, SOPC-IM\}$), respectively. As can be seen from Table IV, BER performance of the proposed systems slightly outperforms the classical pilot-aided OFDM system. The performance gap between SOPC-IM system and the classical OFDM system is also seen from Fig. 5.

Fig. 6 depicts the effect of increasing the total number of the subcarriers N on the spectral efficiency of the proposed techniques and conventional OFDM scheme. It is demonstrated that spectral efficiency for proposed methods are higher than conventional OFDM. For example, spectral efficiency for $N = 1024$ and 8-QAM is 2.25 bpcu for conventional OFDM while the proposed methods have $\eta_{SOPC} = 2.29$ bpcu and $\eta_{OPC} = 2.37$ bpcu for OPC-IM and SOPC-IM respectively. Hence, proposed methods have higher spectral efficiency than conventional OFDM.

TABLE IV
BER GAIN OF THE PROPOSED SYSTEMS COMPARED TO THE CLASSICAL PILOT-AIDED OFDM SYSTEM AT $E_b/N_0=30$ dB.

	BER Gain of the OPC-IM system at $E_b/N_0=30$ dB (1×10^{-2})	BER Gain of the SOPC-IM system at $E_b/N_0=30$ dB (1×10^{-2})
BPSK	1.82	16.67
QPSK	0.91	8.33
8-QAM	0.61	5.56
16-QAM	0.46	4.17
64-QAM	0.30	2.78

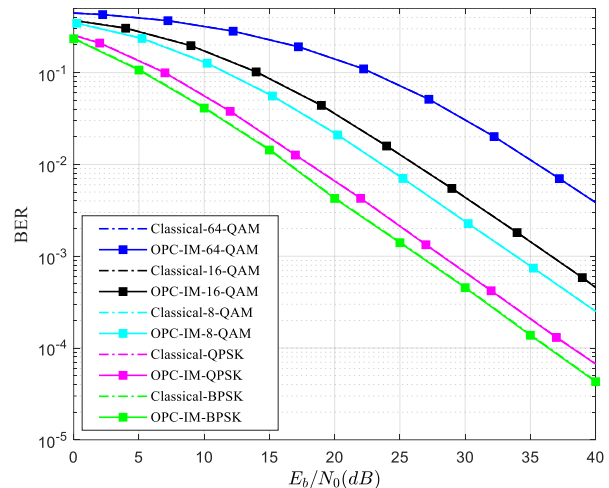


Fig.4. BER performances of proposed OPC-IM scheme and conventional OFDM with $N_d = 512$, and $N_{p1} = 128$.

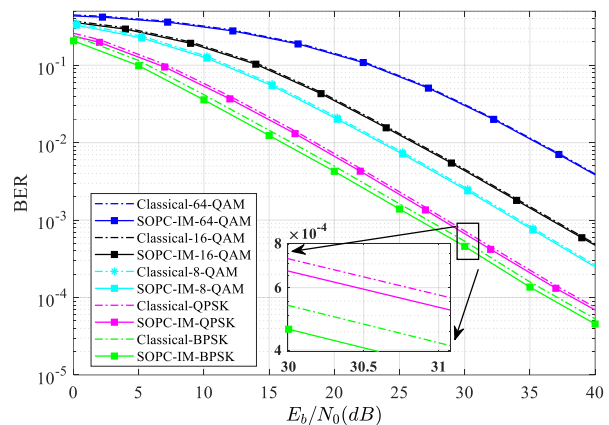


Fig.5. BER performances of proposed SOPC-IM scheme and conventional OFDM with $N_d = 512$, $k=32$ and $N_{p2} = 4$ for each subblock.

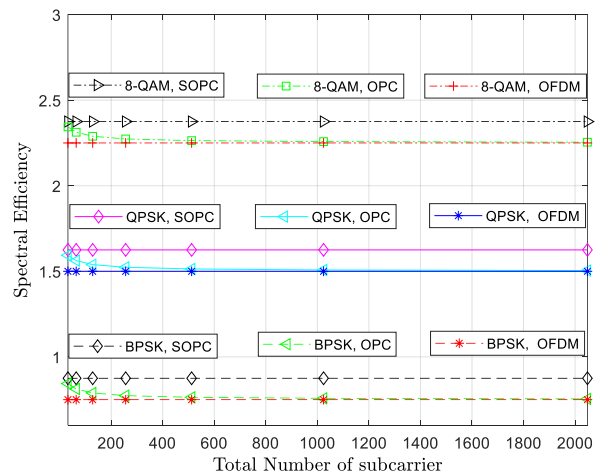


Fig.6. Spectral efficiency with varying N values and modulation schemes.

VI. CONCLUSIONS

In this work, new pilot design approaches are proposed to enhance the spectral efficiency of conventional OFDM system. To obtain the high spectral efficiency, our designs use Walsh-Hadamard codes. Using these codes in each OFDM symbols shows that the proposed methods have higher spectral efficiency, better BER performance and no extra energy consumption with tolerable complexity compared to currently best known OFDM scheme.

REFERENCES

- [1] R. Nee, P. Ramjee, "OFDM for wireless multimedia communications," Artech House, Inc., 2000.
- [2] K. Fazel, S. Kaiser. Multi-carrier and spread spectrum systems: from OFDM and MC-CDMA to LTE and WiMAX. John Wiley & Sons, 2008.
- [3] M. Ergen, Mobile broadband: including WiMAX and LTE. Springer Science & Business Media, 2009.
- [4] B. Kamislioglu and A.Akbal, "Effect of doubly selective channels about bit error probability in OFDM and FBMC," European Journal of Technique, Vol. 7 No.2, 2017, pp. 96-101.
- [5] M. C. Necker and G. L. Stuber, "Totally blind channel estimation for ofdm on fast varying mobile radio channels," IEEE Trans. Wireless Comm., Vol.3, No.5, 2004, pp. 1514-1525.
- [6] Y. Acar, H. Dogan, E. Basar, and E. Panayirci, "Interpolation based pilot aided channel estimation for stbc spatial modulation and performance analysis under imperfect CSI," IET Comm., Vol.10, No.14, 2016, pp. 1820-1828.
- [7] Y. Acar, H. Dogan, and E. Panayirci, "Pilot symbol aided channel estimation for spatial modulation-OFDM systems and its performance analysis with different types of interpolations", Wireless Personal Communications, Vol.94, No.3, 2017, pp. 1387-1404.
- [8] Y. Acar, S. Aldırmaz and E. Bařar, "Channel estimation for OFDM-IM systems," Turkish Journal of Electrical Engineering & Computer Sciences, 2019, 27.3: 1908-1921.
- [9] B. Özbek, R. Yılmaz . "The Adaptive Channel Estimation For STBC-OFDM Systems". IU-Journal of Electrical & Electronics Engineering 5 (2011): 1333-1340.
- [10] B. Kamışlıođlu, A. AKBAL, "LSE Channel Estimation and Performance Analysis of OFDM Systems". Fırat University Turkish Journal of Science and Technology 12 (2017) : 53-57
- [11] Y. Liu, et al. "Channel estimation for OFDM. IEEE Communications Surveys & Tutorials, 2014, 16.4: 1891-1908.
- [12] R. Mesleh, H. Haas, C.W. Ahn., and S. Yun, "Spatial modulation-a new low complexity spectral efficiency enhancing technique", 2006 First International Conference on Communications and Networking, China (pp. 1-5). IEEE., October 2006.
- [13] Y. Acar, "Subblock Aided OFDM with Index Modulation," Balkan Journal of Electrical and Computer Engineering, 7 (2) , 156-161, (2019). DOI: 10.17694/bajece.514324
- [14] S. Aldırmaz, Y. Acar, and E. Basar. "Adaptive dual-mode OFDM with index modulation." Physical Communication 30 (2018): 15-25.
- [15] Y. Acar, and T. Cooklev. "High performance OFDM with index modulation." Physical Communication 32 (2019): 192-199.

[16] E. Basar, M. Wen, R. Mesleh, M. Di Renzo, Y. Xiao, and H. Haas, "Index modulation techniques for next-generation wireless networks," IEEE Access, Vol.5, 2017, pp. 16693-16746.

[17] Q. Wang, G. Dou, X. He, R. Deng, J. Gao, "Novel OFDM System Using Data-Nulling Superimposed Pilots With Subcarrier Index Modulation", IEEE Communications Letters, Vol. 22, No.10, 2018, pp. 2164-2167.

[18] D. A. Bell, "Walsh functions and Hadamard matrixes," IET Electronics Letter, Vol.2, No.9, 1966, pp. 340-341.

BIOGRAPHY

YUSUF ACAR received the B.S.E. degree (with honors), M.S.E. degree and Ph.D. degrees in Electrical and Electronics Engineering from Istanbul University, Istanbul, Turkey, in 2008, 2011 and 2015, respectively. In 2015, he joined the Department of Electrical and Electronics Engineering, Istanbul Kùltür University as an Assistant Professor. He was a researcher at the Purdue University Fort Wayne (PFW), USA, between the September 2017 and September 2018. He is currently a senior system engineering at STM Savunma Teknolojileri Mühendislik ve Ticaret A.Ş., Ankara, Turkey. His general research interests cover communication theory, estimation theory, statistical signal processing, information theory and software-defined platforms. His current research activities are focused on radar signal processing.



SULTAN ALDIRMAZ ÇOLAK received the B.S degree in Electronics and Communications Engineering from Kocaeli University, Kocaeli 2004 and the M.S. and PhD degrees in Yıldız Technical University (YTU), Istanbul 2006 and 2012, respectively. She was a visiting research scholar in the Department of Electrical and Computer Engineering of University of South Florida for the spring and summer of 2009. She is currently an Assistant Professor in the Electronics and Communications Engineering Department of Kocaeli University, Kocaeli, Turkey. Her research interests are in time-frequency signal processing, and communications theory.

Comparisons on Intrusion Detection and Prevention Systems in Distributed Databases

M.GUCLU, C.BAKIR, V.HAKKOYMAZ and B.DIRI


Abstract—With the use of distributed systems, different users can instantly access data from different locations and perform some operations on the data. However, the unauthorized access of multiple users to the system from different points at the same time can lead to dangerous results in terms of data security and confidentiality of the data. This study is based on intrusion detection and prevention systems built on distributed databases and classifies the methods used to analyze and evaluate successes comparatively. It is observed that the artificial immunity algorithm we have described in artificial intelligence techniques, which is one of the methods classified as three categories, gives more successful results compared to the other techniques mentioned in the data mining and statistical methods.

Index Terms—Artificial intelligence method, Data mining, Distributed database, Intrusion detection system, Intrusion prevention system, Statistical method.


I. INTRODUCTION

DATA ACCESS and data communication have become very easy thanks to the developing technology. Today, as a result of increasing of the internet usage rate, expanding of the area used and increasing of the diversity of the work done, the issue of security has become a serious research topic. With the use of distributed systems, different users can gain instant access to data from different locations and perform a number


MEHMET GUCLU, is with Department of Computer Engineering University of Yildiz Technical University, Istanbul, Turkey, (e-mail: mehmetguclu007@gmail.com).

 <https://orcid.org/0000-0001-2306-6008>


CIGDEM BAKIR, is with Department of Computer Engineering University of Yildiz Technical University, Istanbul, Turkey, (e-mail: cigdem.bakir@igdir.edu.tr).

 <https://orcid.org/0000-0001-8482-2412>

VELI HAKKOYMAZ, is with Department of Computer Engineering University of Yildiz Technical University, Istanbul, Turkey, (e-mail: yhkoymaz@yildiz.edu.tr).

 <https://orcid.org/0000-0002-3245-4440>

BANU DIRI, is with Department of Computer Engineering University of Yildiz Technical University, Istanbul, Turkey, (e-mail: diri@yildiz.edu.tr).

 <https://orcid.org/0000-0002-4052-0049>

Manuscript received August 12, 2019; accepted October 16, 2019.
DOI: [10.17694/bajece.605134](https://doi.org/10.17694/bajece.605134)

of operations on the data. Especially in the area of security, a new threat is encountered every day, and accordingly, a rapid development in security measures is taking place. In this sense, many new applications have been developed for the purpose of ensuring the security of the computer systems, preventing unauthorized access, and developing mechanisms for authentication and access controls within the scope of the information security. The acceleration of development processes in internet and communication areas has also led to more systems that malicious attackers can harm, and accordingly, it also revealed the possibility that these attackers may have obtained much more information. For this reason, significant increases in the number of intrusions/attacks and in the use of intrusion detection and prevention methods have been observed.

Often the vast majority of intrusions are carried out by exploiting the vulnerabilities of the used systems. In order to prevent such intrusions, it is necessary to create a safe environment and ensure that intrusions are detected and prevented in a timely manner. No matter how secure a system is, what is important here is to detect attacks early and prevent possible infiltrations. In this context, there are various intrusion detection and prevention systems that have been developed for each possible type of attack from different sources and have been the subject of research and development studies from past to present. In this study, the point that the intrusion detection and prevention systems reached during their development process was mentioned, and by classifying the systems in question, within-class and between-classes performance evaluations were carried out. It should be also noted that in this study, the words “attack” and “intrusion” were used interchangeably.

The rest of the article is organized as follows: related studies in the 2nd section, data set introduction in the 3rd section, material and method in the 4th section, findings in the 5th section, evaluation in the 6th section, and the discussion and conclusion are given in the final section.

II. RELATED WORKS

Quickprop neural networks, a multi-agent statistical prediction system, have been developed to predict database attacks beforehand and to detect vulnerabilities before an attack occurs [1]. In this study, the Pearson correlation coefficient was used to calculate hidden layers and it was tried to identify users who did not have authority over a bank data. Abnormal and incorrect user behavior occurred in a short period was able to be detected.

Another study carried out to detect incorrect user behaviors is the use of genetic algorithms [2]. Based on neural networks, genetic algorithm makes classification by the rules obtained by creating a variety of rules from network properties. The results of this study were presented as comparative to other studies. However, as in the previous one, this study also offered a solution for attacks that could occur in the short term.

The study of predicting and preventing intrusions by using the Hidden Markov Model is another study [3]. The Hidden Markov model is a classification algorithm performed to find hidden States based on given states. Distributed data communicate with each other on very large networks and therefore they are open to serious attacks. In this study, risk analysis was carried out using fuzzy technique and packet rate sent as dangerous was determined. In addition, attacks that would pose risks for distributed environments were attempted to be identified.

Deng and colleagues developed a Support Vector Machine (SVM)-based system for [4] wireless and ad hoc networks intrusion detection system. Security issues for wireless and ad hoc networks were attempted to be solved by the SVM method. In that study, hierarchical and complementary distributed systems were developed for two common Denial of Service (DoS) attacks and performance measurements were made. The effect of DoS attack on speed, distance changes between communication distances, and system performance of location information was observed.

For intrusion prediction and detection systems, Jemili et al. [5] introduced a new approach based on hybrid propagation that combines probabilities with the Bayesian network. The hybrid propagation approach is used to notice abnormal connections occurring both normally and in the network. The purpose of that study was to reveal possible attack plans, scenarios and the relationships between them. In addition, this study combined the host-based intrusion detection systems and the network-based intrusion detection systems to ensure data consistency.

Hu et al. [6] developed a method combining Particle Swarm Optimization (PSO) and SVM techniques for network attacks that often changed and were unnoticed due to the variances occurring in the network. They also attempted to ensure finding of the attacks for each node of the distributed database by using the Gaussian Mixture Model which was an Adaboost-based intrusion detection algorithm. However, this approach was not able to fully identify all types of attacks, particularly new types of attacks.

Abraham et al. [7] aimed the use of genetic algorithms in the intrusion detection systems. In the study, an automatic intrusion detection program was developed by using the given features. The output program was small and simple. Whereas all the features of many machine learning methods are used for intrusion detection, few features of genetic programming was used in the proposed study. A study developing an intrusion detection system for wired networks by using genetic programming in the future was proposed.

In their study, using the artificial neural network, Sağıroğlu et al. [8] aimed to determine what attack method the packets flowing on a network applied. In order to find “Neptune” and

“the ping of death” from these attacks, an artificial neural network model with Multi-Layer Sensor was used. By taking the DARPA Data Sets as an example, they trained for the networks. As a result of that study, detection of DoS attacks that could come from the internet has been successfully achieved.

There are many methods developed and used to date for intrusion detection systems, such as data mining methods, artificial neural networks, and statistical methods.

III. DATASET DEFINITION

One of the most widely used datasets in applications related to intrusion detection systems is the “KDD Cup’99” (Knowledge Discovery and Data Mining Cup’99) dataset. On this dataset, the applicability of multilayer artificial neural networks has been tested and performance analyses have been conducted with parallel programming [9]. The KDD-Cup 1999 data is a data set containing different types of attacks, such as DoS, R2L, U2R, and Probe. This data set contains a total of 972780 samples. In DoS attack-type, the attacker keeps the server engaged by constantly sending fake requests to the server. In this case, the server becomes unable to respond to the requests of the users requesting by formal means. In the R2L (Remote to Local Attack) attack type, the attacker, who does not have access permission, sends packets over the network. Thus, he/she uses the necessary data without permission by providing access to the system. In U2R (User to Root Attack) attack type, the attacker seizes the user's password and accesses the system as if he/she is an administrator. In Probing Attacks, on the other hand, the attacker disrupts the security controls in the system and receives the data he wants over the network.

The KDD-Cup 1999 dataset is a dataset derived from the DARPA BSM (Defense Advanced Research Projects Agency Basic Security Module) data. The DARPA BSM data was generated as a result of simulation of a network traffic belonging to the American Air Force and has been widely used in Intrusion Detection. It includes a total of 38 different types of attacks, including 7 weeks of training and 2 weeks of test data. The contents of the data set consist of tcpdump files taken from night and day network listeners of the attacked machines, log records taken from the attacked machines, and files taken from the security module [10].

IV. MATERIAL AND METHODS

The methods used in the Intrusion Detection and prevention systems in distributed databases are gathered into three main groups as shown in Table 1: Data mining methods, Statistical methods, and Artificial Intelligence methods. The techniques in each method are included in the following subsections. In addition, the performance values of the techniques were evaluated as within-group and between-groups.

TABLE I
INTRUSION DETECTION AND PREVENTION SYSTEMS

Methods Used	
Data Mining	K-Mean Clustering k-Nearest Neighbor (k-NN) Decision Tree Support Vector Machine (SVM) Association Rules
Statistical Methods	Learning Vector Quantization (LVQ) Hidden Markov Models (HMM) Naive Bayes Fuzzy Logic
Artificial Intelligence Methods	Genetic Algorithm Artificial Neural Networks Artificial Immune Techniques

A. Data Mining Methods

Under this group, K-Mean Clustering, k-Nearest Neighbor (k-NN), Decision Tree, Support Vector Machine (SVM), Rules of Association are widely used.

1) K-Mean Clustering

The K-mean clustering technique is one of the non-educational learning methods that group objects according to their similarities. K-mean clustering, which is a division-based method, calculates N objects based on k distance of cluster center and incorporates the object into the cluster where it is near the center of the cluster. The cluster center is initially determined by averaging one or more random instances, and is recalculated in each iteration. Each time, similarities of all data is found according to the new cluster centers. In this way, similar objects are taken into the same cluster as other objects are taken into different clusters. These steps iteratively repeat. The steps end when the clustering error rate (objective function) is minimal. In attack detection and prevention, k-Mean divides N attack types into k amount of clusters. While the within-cluster similarities of the clusters obtained as a result of the division process are maximum, their between-cluster similarities should be minimum. The success and performance of this method vary according to the number of k clusters randomly selected at the beginning, cluster centers, and similarity criteria used. For this reason, it has not produced very successful results in finding the false alarm rate [11]. The false alarm rate is the proportion between classifying a normal data (un-attacked) falsely as attacked data and classifying an attacked data as normal data.

The K-mean method has been used to reduce computational complexity and increase the classification success in the intrusion detection and prevention systems [12]. In the studies conducted on the KDD-Cup 1999 data, the K-Mean Clustering method achieved approximately 96% success [11]. On the other hand, using the KDD-Cup 1999 data, Nadiammai et al. [13] achieved 92.05% success with the K-Mean Clustering method.

2) k-Nearest Neighbor

The k-Nearest Neighbor (kNN) method is the oldest and simplest educational classification method. It calculates the distance between the given input vectors and selects the class of k nearest neighbors. Different classes may exist for different k values. For this reason, the k parameter is very important for classification time and classification accuracy [10]. In this method, the distances of the data, whose class is intended to be found, to all known data are calculated. Euclidean distance is generally used as the distance criterion. Randomly, k neighbor number is determined. The class in which the data is included is determined by looking at the nearest k neighbors. Unlike the K-Mean Clustering method, the classes of the data set used are known in this method, and the newly arrived data is found by looking at these classes. In intrusion detection, the kNN algorithm has been used for the classification of data samples belonging to normal and aggressive species [10]. The distance of the data, whose class to be found, to all data is calculated. By looking at the mean of k data, the attack class of the data at hand is determined. The number of neighbors (k) and whether the number of classes of the data is balanced affect the success and performance of this method.

DARPA attempted to detect samples belonging to the attacker class by looking at the frequency of system calls on the BSM data and provided a low false positive rate. The false positive rate is the rate of finding an attack-free data as an attack-type, that is, it is the rate of false classification of a data.

kNN classifier has been used to reduce the false alarm rate that will identify false and non-normal data in intrusion detection and prevention [14]. In this study, 5 close neighborhoods of false alarm models of normal and aggressive data were taken and a high degree of success was achieved. In the studies conducted on the KDD-Cup 1999 data, kNN (the nearest neighbor method) achieved approximately 97% success [15]. In addition, in the study conducted by Aburomman et al. [16] on KDD-Cup 1999 data, 96% success was also achieved. On the other hand, Chen and his colleagues [17] achieved 91.96% success in their study carried out on the KDD-Cup 1999 data.

3) Decision Tree

The decision tree technique is one of the first classification algorithms used in data mining. Classification results are obtained more easily and more quickly. In this technique, columns in each data set show features, while values for those features are defined in each row. In addition, classes are defined for each record by depending on the values of the features. The first approach in the decision tree is to select features. Then the classes are divided according to this selected feature and this process is iteratively repeated. Each node shows a feature and these nodes have child nodes [18]. In short, the decision tree technique is performed in two stages. In the first stage, a tree consisting of root and child nodes is created. In the second stage, various classification rules are issued according to the structure of this tree. These classification rules show the

nodes remaining between the root node of the tree and its leaves. Figure 1 shows how to create a sample decision tree.

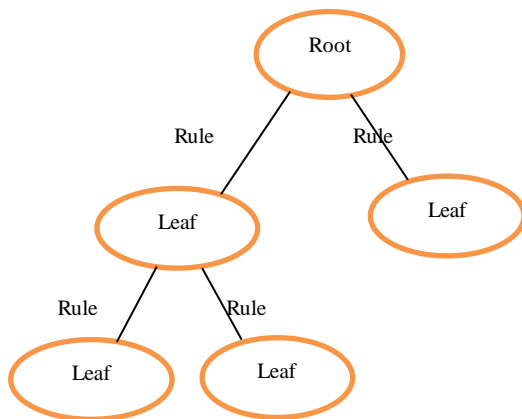


Fig.1. The creation of the decision tree

In network-based intrusion detection and prevention systems, each node displays the types of attacks or normal events in the user or data. The decision tree solves the classification problem by modeling the data set. It predicts future types of attacks based on the model created in false intrusion detection. Provides high performance in real-time intrusion detection. It uses a variety of rule-based models. In the detection of new types of attacks, it gives quite successful results. In the studies conducted on the KDD-Cup 1999 data, the decision tree method achieved approximately 96% success [18]. Rachburee et al. [19] carried out an intrusion detection system on the KDD-Cup 1999 data. In that study, feature selection was made in this data set by the Chi-Square method, and attack types were classified by Decision Trees and Artificial Neural Networks.

4) Support Vector Machines

In analyzing data, the support vector machines (SVM) is a technique based on the supervised learning used in classification. Although SVM was initially used to solve the two-class data problems, it was later expanded and used in the multi-class data problems. In a two-class data problem, a model is created to separate the two classes from each other. This model is obtained by creating a function. Which class the new incoming data belongs to is determined according to this function. The goal in the SVM technique is to find the function that will obtain the most appropriate hyperplane that will separate the two classes from each other. In addition, the support vectors belonging to the two classes must be as maximum as possible to separate the classes from each other. SVM is a classification algorithm that determines the class of each training vector in high-dimensional space. SVM determines the output of the system as well as classes and hyperplane that will determine the support vectors of the data. At the time of training, it determines support vectors by linear, polynomial, or sesamoid functions. The fact that

SVM separates the classes varies depending on various parameters [20].

SVM was used to classify types of attacks on DARPA BSM data in attack detection and Prevention [21]. SVM has two important roles in attack detection. The first and most important is to train the system by providing real-time performance for attack detection, and to calculate the success of the system. The second role is to overcome the scaling problems that may occur in the system. Furthermore, SVM provides very successful results in high-dimensional space and complex classification problems [21]. In the studies conducted on the KDD-Cup 1999 data, the SVM method achieved approximately 99% success [15]. Chen et al. [17] achieved 92.46% success in their study conducted on the KDD-Cup 1999 data. In addition, 93.9% success was also achieved in the study carried out on the KDD-Cup 1999 data by Aburomman et al. [16].

5) Association Rules

The association rules algorithm is a data mining technique that examines association behaviors between data by looking at the historical data. Assuming that T is transaction, let us express all transactions in the database with $\{ T_1, T_2, \dots, T_n \}$ and transaction objects with $\{ i_1, i_2, \dots, i_m \}$. The rule between the data is $X \rightarrow Y (c, s)$. Here, s refers to the support between the data and c refers to the confidence interval. Support indicates the frequency of using an object for all objects. Confidence indicates the frequency of using of an object with the other object. s specifies the percentage of X and Y transactions performed together; c specifies the ratio [22]. Association rules are most commonly used in product sales. For example, more sales can be made to the customer by finding the products sold together in the markets. The first step in this algorithm is to find frequently used data objects for each transaction and set a threshold value for the s confidence interval. The second step is to create appropriate rules for the data set. The main problem in the apriori algorithm is to construct a large number of rules. However, rules created on these data objects can be restricted by selecting frequently used data objects. The algorithm steps are as follows:

- Minimum support and confidence interval are determined
- Support value of each object is found.
- The smallest support value is compared to the support value of each object, and those smaller than the smallest value are discarded from the set of objects.
- Binary association rules are created and the same operations are repeated.

A new approach was carried out on the KDD-99 data set using association rules for intrusion detection and prevention [22]. With the set of fuzzy relational rule, a new classification approach was created that determines different classes. They also aimed to perform an effective algorithm that could make measurements on new data objects.

In the study conducted by Tajbakhsh et al. [22], 91% success was achieved. In the studies carried out on the KDD-Cup 1999 data, the association rule method achieved approximately 96% success [7].

Security analyses were conducted on event records by establishing relational rules between read and write transactions [23]. In this study, the performance of the system was measured by trying to determine reliable read and write transactions on a very large database.

B. Statistical Methods

Under this group, Learning Vector Quantization (LVQ), Hidden Markov Models (HMM), Naive Bayes, Fuzzy Logic techniques are used widely.

1) Learning Vector Quantization

Learning Vector Quantization (LVQ) is a supervised classification method. This technique is accomplished according to the Kohonen learning rule. Unlike other classification algorithms, the LVQ technique finds that the n-dimensional input vector can be represented by vectors of which process element. The finding of this vector varies depending on the learning rate and the maximum number of training. The process element closest to the input vector is rewarded, and the class of the input vector is approximated to that process element. Thus, this processing element is rewarded. If not, it is punished by being removed. This way feature vectors are updated.

In LVQ intrusion detection and Prevention Systems, a layer is used to classify the attack types given as input to the system. This layer is independent of the input vectors and collects the input vectors, which are similar to each other, in the same class [24]. LVQ networks have two layers, the first and the next layers. The first layer trains the attack types given as input vectors for classification. The second layer is converted to a target layer specified by the user. The classes that are trained show the subclasses of the system and the classes that must occur at the target. LVQ has been successfully used in many applications from telecommunications to robotics [25]. Also, this classification algorithm is intuitive and has several forms such as LVQ2, LVQ3. The cost function is determined by calculating the distance between classes [25]. The learning rate and number of iterations determine performance and it is selected iteratively.

LVQ was used to classify attack types [24]. They used two layers to classify 5 types of attacks: Normal, DoS, U2R, R2L, and probe. These layers determine the subclass and the master class. The LVQ method has achieved about 81% success in the studies conducted on the KDD-Cup 1999 data [26]. Degang et al [27] achieved 76.3% success by classifying attack types through using LVQ after normalizing data on the same data set and making feature selection.

2) Hidden Markov Model

In the Hidden Markov Model (HMM), it is attempted to predict future states that may occur when present states are given as input to the system. HMM is a stochastic process

since it produces a different output each time it is run. Also, in Markov models, the system can move from its own state to another state, depending on the probability distribution, or remain in the same state. The probabilities that occur in the state are called transition probabilities. In HMM, unlike the normal Markov model, states are not seen by the observer. However, transitions that depend on the state can be seen.

Intrusion detection systems are defense mechanisms that detect bad packets in network traffic in the distributed systems. The intrusion prevention systems are divided into two, network-based and host-based intrusion prevention systems. The main purpose of the intrusion prevention systems is to observe suspicious flow and suspicious packets in normal network traffic. Also, by rearranging the path of suspicious network traffic, it ensures the preventing of attacks such as DoS attacks. It also observes all network performance and high packet processing rates and attempts to reduce false positive rates at DoS stage [3].

HMM is widely used in many areas such as bioinformatics, handwriting recognition, image processing, and audio processing. HMM was used for attack prediction prevention in distributed database and risk assessment was done [3]. There are two stochastic processes in HMM. One is the state of the system ($x_t; t=1,2,\dots$) and the other is the observable processes ($x_t; t=1,2,\dots$). \mathbf{T} refers to observations that are consecutive among intrusion detection representatives.

HMM units are as follows [3]:

- $S = \{ s_1, s_2, \dots, s_N \}$ identifies possible states in the system. There are 4 states in this study: Normal (N), Intrusion Attempt (IA), suspicious activities occurring in the network-Intrusion in Progress (IP), and occurring one or more attacks in the system-Successful Attack (SA).
- The observations that occur in the system are expressed by $V = \{ v_1, v_2, \dots, v_M \}$. There are three observations in this study: observation of no suspicious activity (N), observation of a suspicious activity in the network (P), and observation of the successful realization of a suspicious activity (SA).
- The first distributed vector is defined in the system as $\pi = \{ \pi_i \}$ ve $\pi_i = P(x_i=i)$. This system is assumed to be N-state.
- The transition probability matrix is $P = \{ p_{ij} \}$ ve $p_{ij} = P(x_t=j | x_{t-1}=i)$. This shows the interaction between the system and the users entering the system without permission.
- The observable probability matrix defines security or quality for representatives of each intrusion detection system.

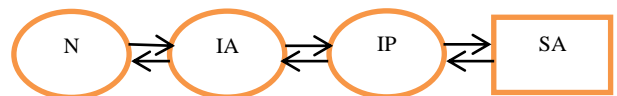


Fig.2. Security states used in HMM

The security modeling of the Hidden Markov model in networks is as in Figure 2 [3]. States indicate security states in circles. If there is a damage to the system, it is shown with SA. Observations are independent of current states. From each iteration result, the calculated probability distributions are updated [28].

3) Naive Bayes Classifiers

The Naive Bayes technique is a technique based on Bayesian probability. The Naive Bayes classifier handles events independent of each other. The most important reason for this is that the probability of a feature is not affected by the probability of the other features. It is widely used in classifying text documents and in classifying whether incoming mails to an e-mail are spam or not. It calculates the probabilities of all data to determine which class the data belongs to. The probability of one data depends on the probability of the occurrence of the realizing state of the other data and the probability of all data being. Whichever class has a greater probability value, the data is included in that class. The results of the Naive Bayes classifier technique are generally correct. However, the fact that the data is noisy can produce incorrect results due to some reasons, such as variance. By reducing input features, Naive Bayes allows important features to be found. Thus, effective and active intrusion detection system can be created [31].

With the Naive Bayes classification method, it was attempted to detect attacks occurring on the network [32]. They achieved very successful results in the study that they conducted in order to detect new types of attacks occurring in the network on the KDD'cup'99 data set. Deshmukh et al. [29] achieved 88.02% success in their study carried out on the KDD-Cup 1999 data [29]. The Naive Bayes method have achieved about 85% success in the studies performed on the KDD-Cup 1999 data [33].

4) Fuzzy Logic

Fuzzy Logic techniques have been used in the field of computer security since the 90s. There is no concept of certainty in this technique. A data is included in more than one class according to the degree of membership. Degree of membership is between 0 and 1. For example, the fact that air or water is warm, cold or mild varies according to people. For this reason, certain value ranges cover these three classes. The steps of Fuzzy Logic technique are given below:

- All data is defined.
- Membership functions are determined.
- The rules to be applied are defined.
- The rules are evaluated and combined.
- The data is classified according to the membership values shown.

In mixed systems, it continuously analyzes incoming data to ensure computer security. Additionally, this technique attempts to detect user signatures or attacks with classical pattern recognition. It detects events that are faulty or abnormal. It consists of two stages. The first stage is rule generation, while the second stage is detection and prevention of attacks. Because Fuzzy Logic addresses low,

medium, and high attack types in a way to cover each other, it overcomes many sharp boundary problems and reduces errors that occur as false positives. It ensures system optimization in real-time systems [34].

Using the Fuzzy Logic technique, Tian et al. [35] divided the large data sets into sub-datasets and performed a performance analysis of the data set by following the TCP data for different data sets. In the studies conducted on the KDD-Cup 1999 data, the Fuzzy Logic method has achieved approximately 94% success [30]. Nadiammai et al. [13], on the other hand, achieved 81.54% success using the K-average clustering method on the KDD-Cup 1999 data.

C. Artificial Intelligence Methods

Under this group, Genetic Algorithm, Artificial Neural Networks, and Artificial Immune Techniques are widely used methods.

1) Genetic Algorithm

Genetic algorithms are methods based on biological processes and used in solving search and optimization and also used to be able to make modeling.

Genetic algorithms are used to solve problems that are difficult or impossible to solve by conventional methods. Since genetic algorithms are a random search method, they work over multiple sets of solutions instead of searching a single solution for the optimal solution of the problem. Therefore, results in problem solving are not always best. The reason why genetic algorithm is preferred is that genetic algorithm does not need any information about the nature of the problem. The basic steps of the genetic algorithm are as follows [36]:

- A random population is created.
- By applying genetic processes (selection, crossover), new individuals are created from this population. Selection and crossover processes create new individuals by exchanging genes from individuals in the population.
- Among these individuals, the most appropriate individuals who can solve the problem are selected.
- The population length is the same for all iterations, and for future iteration the highest probability of the function is chosen. Iteration stops when it comes to the optimal threshold.

In distributed systems, the network structure is denoted by $G=(N,E)$ weighted graph. E indicates communication links between nodes, while N indicates nodes. In multiple network structures, the variable T represents cost, while the network structure is denoted by $T=(N_T,E_T)$. $P_T(s,u)$ denotes the path from s source to u destination node. The cost of T is calculated as in Equation 1 [36].

$$C(T_s) = \sum_{e \in E_T} C(e), e \in E_T \quad (1)$$

On the other hand, the minimum bandwidth from s source node to u target node is calculated as in Equation 2.

$$B_T = \min(B(\epsilon), \epsilon \in E_T) \tag{2}$$

In studies conducted on the KDD-Cup 1999 data, the genetic algorithm has achieved approximately 85% success [37].

2) *Artificial Neural Networks*

Artificial Neural Networks (ANN) are information systems that model the human brain and classify data through learning. It was developed based on the working principle of the human brain. In other words, ANN is an information processing structure developed with a logic similar to biological neural networks and linked to each other by weights.

An ANN consists of input, output, and hidden layers. While the data is imported to the artificial neural network by the input layer, it was exported by the output layer. The layers between the input and the output layer constitutes the hidden layers.

Neurons in feedforward artificial neural networks are connected only forward. The neuron refers to all data connected together. Each layer of the neuron network contains the connection of the next layer, and these connections are not backward. That is, there is a hierarchical structure between neurons, and neurons in one layer only transmit data to the next layer. The forward transition consists of activation flow reaching the output layer and input samples. Activation functions such as Sigmoid and Gaussian function can be used. In the back-transition phase, the actual output on the network is compared to the target output and the error in the output units is calculated [38]. For intrusion detection and prevention systems, this structure is calculated as in Equations 3, 4, and 5 [39].

$$x(t) = f(W^A x_c(t) + W^B u(t - 1)) \tag{3}$$

$$x_c(t) = x(t - 1) \tag{4}$$

$$y(t) = g(W^C X(t)) \tag{5}$$

where $x(t)$ is the hidden layer output, $y(t)$ is the output of the output layer, $u(t-1)$ is the input of the network, WA is the weight of the connection between units and the hidden layer, WB is the weight of the connection between the input and output layer, WC is the weight of the connection between the hidden and output layer, $f(\cdot)$ and $g(\cdot)$ are the activation code between the hidden layer and the output layer [39].

The backpropagation network shows how the neuron is trained. ANN is a type of the supervised learning. When the supervised method is used, the network is provided with both sample inputs and expected outputs. The expected outputs for the given networks are compared to the actual outputs. If the expected outputs are used, the error is calculated and the weights of the various layers are

adjusted backward from the output layer to the input layer. The network updates its coefficients to obtain the expected output. In ANN, the error in the output layer is calculated as a result of each iteration and this error is transmitted from the output layer to the input layer toward all neurons and the weights are rearranged according to the margin of error. This margin of error is distributed to the neurons preceding the neuron itself in proportion to their weight.

ANN is a method through which security vulnerabilities can be solved in distributed systems. In distributed computations, ANN consists of nodes, connections between nodes, and processing units. The connection between the two units consists of one unit's weights that affect the other unit. These units move from the input nodes to the output nodes by gathering and passing through the threshold value. ANN is implemented in different network securities, medicine, marketing, banking and finance, telecommunication, operations management and other industries [38].

Tong and colleagues used the ANN model in the intrusion detection system. With Elman neural networks, they have attempted to detect both faulty and anomaly attacks. These neural networks have the content of the nodes; the content of each node receives its input from the single hidden layer, and the output for each node is connected to the hidden layer [39]. Mahit et al. [40] achieved about 94% success in their study conducted with ANN on the KDD-Cup 1999 data. In addition, 98.5% success was achieved in the study carried out on the KDD-Cup 1999 data by Aburomman et al. [16].

3) *Artificial immune system*

We can identify the immune system as a protective mechanism that brings the body in defense by protecting it against diseases. The artificial immune system (AIM), which has emerged by inspiring from the biological definition of the immune system and its working logic, has been observed to be frequently involved in research studies, especially in recent years as an artificial intelligence-based method. One of these areas of research is also the use of AIM in intrusion detection. It can be said that the first source of inspiration for this system, which we call artificial immunity, is computer viruses.

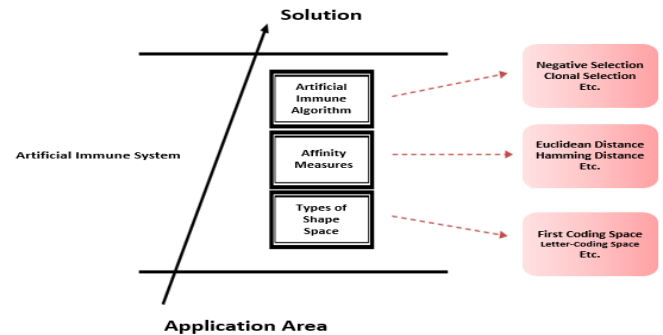


Fig.3. The Layered Structure of AIM [41]

The artificial immune system (AIM) has a layered structure and it is shown in Figure 3 [41]. The similarity between antigen and antibodies used as the basic part of the system is ensured by affinity measurement, such as recognition and representation in the shape space. Euclidean, Hamming, or Manhattan distances are used for the affinity measurement. By mutating the individuals with a low-affinity measurement more, the best antibody that will recognize an antigen is found, and the antibody is replicated by mutation at certain rates (clonal selection algorithm). The reproduced cells are added to the antibody cells. The antigen is presented to the antibody cells. Antibodies that best identify the antigen and are above a certain threshold value are taken and raced among them (similarity rates are calculated with Affinity measurement). As a result of the race, a certain number of individuals with low affinity are taken and added to the memory cell. In this way, the memory cell that will represent the best antibody cells against an antigen is produced. This process is also applied separately for other antigens to be presented to the system [42]. The artificial immune method achieved approximately 99% success in the study of Shem et al. [42] conducted on the KDD-Cup 1999 data. Chen et al. [17] also achieved 92.41% success in the study they carried out on the KDD-Cup 1999 data.

V. EXPERIMENTAL STUDY

In this study, distributed database intrusion detection and prevention methods were compared according to some metrics such as accuracy, speed, performance, and size of the data set used. Accordingly, the advantages and disadvantages of the used methods are shown in Table 2.

TABLE II
ADVANTAGES AND DISADVANTAGES OF THE METHODS USED IN DISTRIBUTED SYSTEMS

Data Mining Methods	<p>Advantage: it is very useful for small data and its success is high. It is safe, and its results are close to precise.</p> <p>Disadvantage: Its application to the system brings with it a number of computational and time complexities.</p>
Statistical Methods	<p>Advantage: It generally does not need prior information to observe Normal movements. It is therefore active in finding new types of attacks.</p> <p>Disadvantage: Determining the parameters and metrics required to calculate the success of the system are quite difficult and vary for each set of data.</p>
Artificial Intelligence Methods	<p>Advantage: It is flexible and can be applied to almost all data sets.</p> <p>Disadvantage: It consumes a high amount of resources for the intrusion detection and prevention systems.</p>

Artificial intelligence techniques and data mining methods provide more successful results than statistical methods. In addition, artificial intelligence methods are more easily

implemented for big data. By combining these techniques, it may be possible to achieve safer and faster data transmission.

VI. RESULTS AND EVALUATION

With the development of technology, data access and data communication has become quite easy. Thanks to the distributed systems, quick and easy access to data from anywhere can be ensured. However, the fact that more than one user wants to access the system from different locations at the same time presents a number of problems such as data security, data confidentiality, service continuity, authorization, and secure storage of the data. Problems of the users' access to data, listening to the network, blocking the service, insecure networks, obtaining and storing of the important information by unauthorized people constitute a few of these problems. In order to overcome these problems, various systems developed based on the intrusion detection and prevention techniques applied in the field of data mining, statistical and artificial intelligence are used.

The performance results of the techniques, which are used for intrusion detection and prevention systems in distributed databases, on the KDD-Cup-1999 data are given in Table 3 by compiling from various studies. In the studies conducted, it was observed that the use of noise removal, feature extraction, and selection methods increased classification success, and with these methods, the data at hand is made more useful. Therefore, the results of the studies vary depending on the techniques of normalization, noise removal, feature extraction, and feature selection.

TABLE III
SUCCESFULL OF THE METHODS USED IN DISTRIBUTED SYSTEMS

Methods	Techniques	Success Rate (%)
Data Mining Methods	Support Vector Machines	99.18
	k-Nearest Neighborhood	97.04
	Decision Tree	96.83
	Association Rules	96.9
	K-Mean Clustering	96.9
Statistical Methods	Fuzzy Logic	94.92
	Hidden Markov Models	93.4
	Naive Bayes	88.02
	Learning Vector Quantization	81
Artificial Intelligence Methods	Artificial Immune Techniques	99.74
	Artificial Neural Networks	94
	Genetic Algorithm	85.7

VII. DISCUSSION AND CONCLUSION

Today, computers are widely used in many parts of education, economy, military and business life. A lot of valuable information that is confidential, private and needs to be protected is shared among users over the network. The rapid development of the technology and the sharing of information leads to occurring of attacks against the network and computer security. Emerging of many security problems such as encryption, authorization, damaging the system by unauthorized persons, crashing the system also necessitates

taking the required security measures. In this case, the intrusion detection and prevention systems have emerged to detect users who are abusing the system and to prevent abnormal behaviors.

As a result, artificial intelligence techniques, particularly the artificial immune technique, have given highly successful results in the intrusion detection and prevention systems for distributed databases. The effective use of artificial intelligence techniques in the intrusion detection and prevention is of great importance for future studies. It should not be ruled out that more successful results can be achieved by using hybrid methods, including especially improved artificial intelligence techniques.

REFERENCES

- [1] P. Ramasubramanian, A. Kannan, "Multi-Agent based Quickprop Neural Network Short-term Forecasting Framework for Database Intrusion Prediction System", CiteSeerX, 2014.
- [2] P. Romasubramanian, A. Kannan, A. "A genetic-algorithm based neural network short-term forecasting framework for database intrusion prediction system", *Soft Computing*, Vol., 8, pp. 699-714, 2006.
- [3] K. Haslum, A. Abraham, "Disp: A framework for distributed intrusion prediction and prevention using hidden markov models and online fuzzy risk assessment", 3rd International Symposium on Information Assurance and Security, pp.183-190, 2007.
- [4] H. Deng, Q. Zeng, Q. "SVM-based detection system for wireless ad hoc networks", *Vehicular Technology Conference*, Vol.3, pp. 2147-2151, 2003.
- [5] F. Jemili, M. Zaghdoud, "Hybrid Intrusion Detection and Prediction multiAgent System, HIDPAS", (IJCSIS) *International Journal of Computer Science and Information Security*, Vol.5, 1, pp. 62-71, 2009.
- [6] W. Hu, G. Jun, "Online Adaboost-Based Parameterized Methods for Dynamic Distributed network Intrusion Detection", *IEEE Transactions on CyberNetics*, Vol.44, 3, pp. 66-82, 2014.
- [7] A. Abraham, C. Grosan, C. Martiv-Vide, "Evolutionary design of intrusion detection programs", *Int. Journal of Network Security*, Vol. 4, pp. 328-339, 2007.
- [8] Ş. Sağıroğlu, E.N. Yolaçan, U. Yavanoğlu, "Zeki Saldırı Tespit Sistemi Tasarımı ve Gerçekleştirilmesi", Ankara, 2011.
- [9] M.Z. Yıldırım, A. Çavuşoğlu, B. Şen, İ. Budak, İ. "Yapay Sinir Ağları ile Ağ Üzerinde Saldırı Tespiti ve Paralel Optimizasyonu", XVI. Akademik Bilişim, Mersin, 2011.
- [10] Y. Liao, V.R. Vemuri., "Use of K-Nearest Neighbor classifier for intrusion detection", *Elsevier Computers&Security*, Vol. 21, 5, pp.439-448, 2002.
- [11] M. Jianliang, "The Application on Intrusion Detection based on K-Means Cluster Algorithm", *International Forum on Information Technology and Applications*, pp. 150-152, 2009.
- [12] K.M. Faraoun, A. Boukelif, "Neural Networks learning improvement using the K-Means clustering algorithm to detect network intrusions", *International Journal of Computer and Information Engineering*, Vol. 1, 10, pp. 3138-3145, 2007.
- [13] G.U. Nadiammai, M. Hemalathen, "An evaluation of clustering technique over intrusion detection system", *ICACCI '12 Proceedings of the International Conference on Advances in Computing, Communications and Informatics*, pp. 1054-1060, 2002.
- [14] K. Law, F. Kwok, "IDS False Alarm Filtering using KNN Classifier, Springer Information Security Applications Lecture Notes in Computer Science", pp.114-121, 2004.
- [15] A. Adetunmbi, "Network Intrusion Based on Rough set and k-Nearest Neighbour", *International Journal of Computing ICT Research*, Vol. 2, 1, 2008.
- [16] A. Aburonman, M. Reaz, "A novel SVM-kNN-PSO ensemble method for intrusion detection system", *Elsevier Applied Soft Computing*, Vol.38, pp. 360-372, 2016.
- [17] M. Chen, P. Chang, J. Wu, "A population-based incremental learning approach with artificial immune system for network intrusion detection", *Elsevier Engineering Applications of Artificial Intelligence*, 51, pp. 171-181, 2016.
- [18] A. Peddabachigari, A. Abraham, "Intrusion detection systems using decision trees and support vector machines", *International Journal of Applied Science and Computations*, pp.1-16, 2004.
- [19] N. Rachburee, N. Punlumjeak, "Big Data Analytics: Feature Selection and Machine Learning for Intrusion Detection on Microsoft Azure Platform", *Journal of Telecommunication Electronic and Computer Engineering*, Vol. 9, 1-4, pp. 1-5, 2017.
- [20] A. Sung, S. Mukkamala, "Identifying import features for Intrusion Detection using Support Vector Machines and Neural Networks", *Proceedings of the 2003 Symposium Applications and the Internet (Saint'03)*, 2003.
- [21] S. Mukkamala, G. Janoski, "Intrusion Detection using Neural Networks and Support Vector Machines", *IJCNN'02 Proceedings of the 2002 International Joint Conference on*, Vol. 2, pp. 1702-1707, 2002.
- [22] A. Tajbakhsh, M. Rahmati, "Intrusion detection using fuzzy association rules", *Elsevier Applied Soft Computing*, Vol. 9, pp. 462-469, 2009.
- [23] Y. Hu, B. Panda, "A data mining approach for Database Intrusion Detection", *ACM Symposium on Applied Computing*, pp. 711-716, 2004.
- [24] R. Noum, Z. Al-Sultani, "Learning Vector Quantization (LVQ) and k-Nearest Neighbor for Intrusion Classification", *World of Computer Science and Information Technology Journal (WCSIT)*, Vol. 2, 3, pp. 105-109, 2012.
- [25] B. Hamman, D. Hoffman, "Learning vector Quantization for (dis-)similarities", *Elsevier Neurocomputing*, Vol. 131, pp. 43-51, 2014.
- [26] E. Soleiman, A. Fatarat, "Using Learning Vector Quantization (LVQ) in Intrusion Detection Systems", *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, Vol. 1, 10, 2014.
- [27] Y. Degang, C. Guo, C. "Learning Vector Quantization Neural Network Method for Network Intrusion Detection", *Wuhan University Journal of Natural Sciences*, Vol. 12, 1, pp. 147-150, 2007.
- [28] L.R. Rabier, "A tutorial on Hidden Markov Models and Selected applications speech recognition", *Ready in Speech Recognition*, pp. 267-296, 1990.
- [29] D. Deshmukh, T. Ghorpade, P. Padiya, "Improving Classification Using Preprocessing and Machine Learning Algorithms on NSL-KDD Dataset", *2015 International Conference on Communication, Information & Computing Technology (ICCICT)*, 2015.
- [30] R. Shanmugavadivu, N. Nagarajan, "Network Intrusion Detection System using Fuzzy Logic", *Indian Journal of Computer Science and Engineering (IJCSCE)*, Vol. 2, 1, pp. 101-111, 2014.
- [31] D.S. Mukherjee, N. Sharma, "Intrusion Detection using Naive Bayes Classifier with Feature Reduction", *Elsevier Procedia Technology*, Vol. 4, pp. 119-128, 2012.
- [32] S. Sharma, "An Improved Network Intrusion Detection Technique based on k-means clustering via Naive Bayes Classification", *IEEE-International Conference on Advances In Engineering, Science and Management (ICAESM-2012)*, pp. 417-422, 2012.
- [33] M. Panda, M. Patra, "Network Intrusion Detection using Naive Bayes", *IJCSNS International Journal of Computer Science and Network Security*, Vol. 7, 12, pp. 258-263, 2007.
- [34] A. El-Semany, "A Framework for Hybrid Fuzzy Logic Intrusion Detection Systems", *IEEE International Conference on Fuzzy Systems*, pp. 325-330, 2005.
- [35] J. Tian, "Intrusion detection combining Multiple Decision Trees by Fuzzy Logic", *Proceedings of the sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'05)*, 2005.
- [36] S. Janakiraman, V. Vasudevan, "An Intelligent Distributed Intrusion Detection System using Genetic Algorithm", *JCIT Journal of Convergence Information Technology*, Vol. 4, 1, 2009.
- [37] M. Hassan, "Network Intrusion Detection System Using Genetic Algorithm and Fuzzy Logic", *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 1, 7, pp. 435-444, 2013.
- [38] W. Chen, S.H. Hsu, "Application of SVM and ANN for intrusion detection", *Elsevier Computers&Operations Research*, 32, 2617-2634, 2005.
- [39] X. Tong, Z. Wang, "A research using hybrid RBF/Elman neural networks for intrusion detection system secure model", *Elsevier Computer Physics Communications*, Vol. 180, pp.1795-1801, 2009.
- [40] D. Mahit, "Using Artificial Neural Network Classification and Inversion of Intrusion in Classification and Intrusion Detection System,

International Journal of Innovative in Computer and Communication Engineering, Vol. 3, 2, pp. 1102-1108, 2015.

- [41] L.Castro, J. Timmis, "Artificial immune systems as a novel soft computing paradigm", *Soft computing*, Springer, Vol. 7, 8, pp. 526–544, 2003.
- [42] J.Shen, J. Wang, "Network Intrusion Detection by Artificial Immune System", *IEEE Power and Energy General Meeting*, pp.1-8, 2011.
- [43] C. Bakir, V. Hakkoymaz, "Veritabanı Güvenliğinde Saldırı Tahmini ve Tespiti için Kullanıcıların Sınıflandırılması", *ISCTurkey2015 8.Uluslararası Bilgi Güvenliği ve Kriptoloji Konferansı (VIII. Int'l Conference on Information Security and Cryptology)*, pp. 28-33, 2015.



BANU DIRI received the B.S. degrees in computer engineering from the University of Yildiz Technical University, in 1987 and the M.S. degree in computer engineering from Yildiz Technical University, in 1990, Ph.D.degree in Yildiz Technical University in 1999. She works as a

professor at the Yildiz Technical University. Her research interests include data mining, natural language processing, machine learning and artificial intelligence.

BIOGRAPHIES



MEHMET GUCLU received the B.S. degrees in computer engineering from the University of Yildiz Technical University, in 2009 and the M.S. degree in computer engineering from Yildiz Technical University, İstanbul, in 2013. He started his PhD in 2013 at Yildiz Technical University and still

continues. His research interests include information security, data mining and machine learning.



CIGDEM BAKIR received the B.S. degrees in computer engineering from the University of Sakarya, in 2010 and the M.S. degree in computer engineering from Yildiz Technical University, İstanbul, in 2014. She started his PhD in 2014 at Yildiz Technical University and still continues. Since 2012, she was a

Research Assistant with the Yildiz Technical University. She works a Research Assistant at Iğdir University. Her research interests include recommendation systems, information security, machine learning and data mining.



VELI HAKKOYMAZ received B.S. degrees in computer engineering from Hacettepe University, in 1987 and M.S.degree in Computer Science from University of Pittsburgh (PA), In 1992, Ph.D.degree in CWRU (OH) in 1997. In 2011, he received the title of Associate Professor. His research

interests include database management systems, computer architecture, operating systems and distributed systems.

Determination of the Optimum Hybrid Renewable Power System: A case study of Istanbul Gedik University Gedik Vocational School

O. AKAR, U. K. TERZI, T. SONMEZOCAK and B. K. TUNCALP


Abstract—In this study, Hybrid Renewable Power System (HRPS) has been designed to meet the energy requirement of Istanbul Gedik University Vocational School (IGUVS) in an optimum way. The energy requirement will be achieved mainly with wind and solar power generation system. The hybrid regenerative power system will be requested from the network when it can't meet its energy needs. The optimum configuration of grid-connected solar/wind hybrid power generation system will be determined by considering the wind speed data and solar radiation data of the location of IGUVS. Moreover, sensitivity analysis will be carried out taking into consideration the wind speed and solar radiation values. If the optimum hybrid power generation system specified is used, the carbon emission values obtained will be examined.

Index Terms—Building Based Renewable Energy Sources; Hybrid Power Systems; Distributed Generation.


I. INTRODUCTION

TODAY, the production and consumption of energy have become a global problem worldwide, which is an independent problem that every country is experiencing.


Onur AKAR, is with Department of Electronics and Automation, Gedik Vocational School, Istanbul Gedik University, Istanbul, Turkey, (e-mail: onur.akar@gedik.edu.tr).

 <https://orcid.org/0000-0001-9695-886X>


Umit K. TERZI, is with Department of Electrical-Electronics Engineering Technology Faculty, Marmara University Istanbul, Turkey, (e-mail: terzi@marmara.edu.tr).

 <https://orcid.org/0000-0001-6739-7717>

B. Koray TUNCALP, is with Department of Electrical Electronics Engineering University of Halic, Istanbul, Turkey, (e-mail: koraytuncalp@halic.edu.tr).

 <https://orcid.org/0000-0003-1632-8900>

Temel SONMEZOCAK, is with Department of Electricity and Energy, Gedik Vocational School, Istanbul Gedik University, Istanbul, Turkey, (e-mail: temel.sonmezocak@gedik.edu.tr).

 <https://orcid.org/0000-0003-4831-9005>

Manuscript received August 12, 2019; accepted October 16, 2019.
DOI: [10.17694/bajece.623632](https://doi.org/10.17694/bajece.623632)

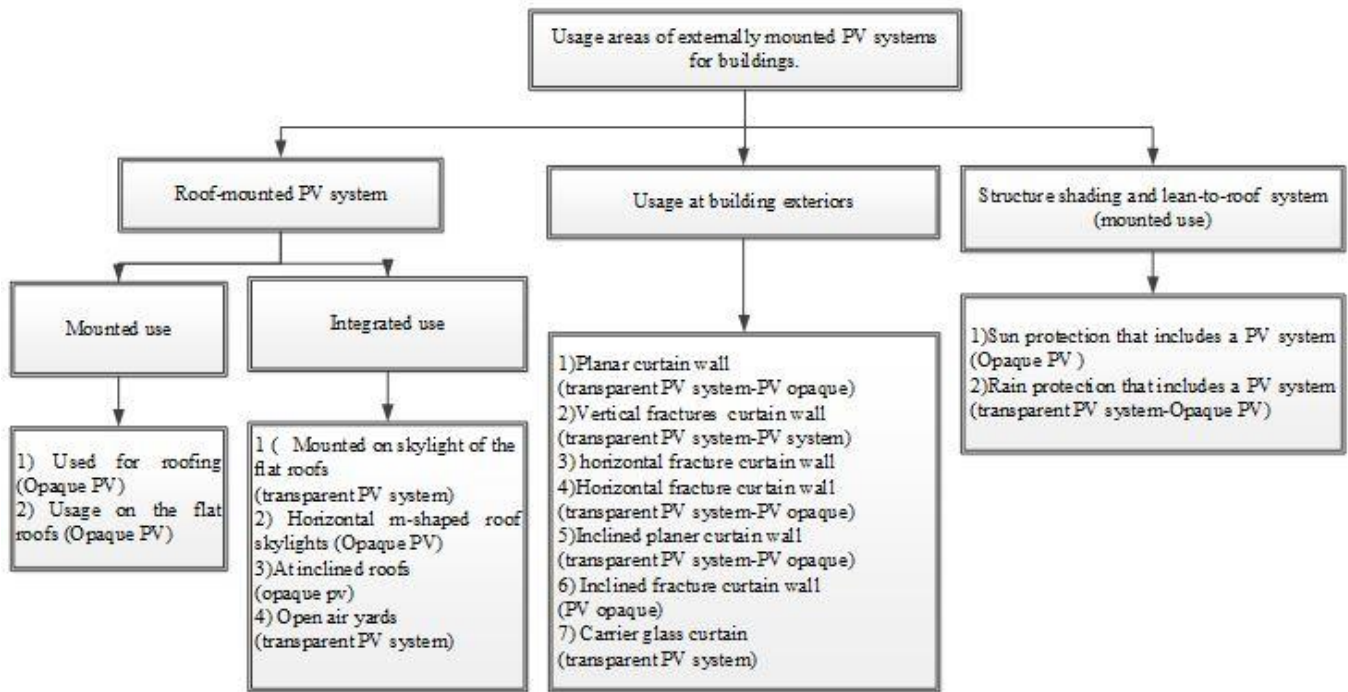
75.5% of the world's energy sources are fossil-based conventional energy sources. Due to factors such as the depletion of conventional energy sources in recent years, the generation of carbon emissions from these sources and the emission of greenhouse gases, countries have begun searching for alternatives to meet their energy needs. When countries plan their energy policies, they are mainly concerned with two critical parameters [1-3]. These are ensuring the use of renewable energy sources and taking measures for efficient use of energy. To benefit from renewable energy sources, the potential of renewable energy sources (wind, solar biomass, hydraulics, etc.) must be sufficient. The feasibility study should be carried out for whatever renewable energy source is used before making an investment decision [4]. Another parameter is to determine the policies that will enable us to avoid unnecessary energy consumption through the efficient use of energy. The share of energy consumed in buildings in the world can reach 45% -50% of the total energy consumption of the countries. It is critical to construct buildings that consume less energy nowadays due to factors such as the energy crisis since the 1970s and the effect of the increase in the unit cost of energy resulting from the constantly increasing energy production costs to the end users [5, 6]. It is projected by the international energy authorities that the use of renewable energy will steadily increase [7, 8]. Countries, especially developed countries, are creating energy maps to identify their potential for renewable energy sources. Renewable energy sources-based energy generation systems can generate power both as part of or independent of national networks. Many buildings use hybrid power generating systems, some independent of national network but more commonly supported by the national network. Therefore, each building must be transformed into an active structure that can produce its own energy, instead of being a passive energy consumer [9]. As it is well known, the biggest handicap of renewable energy sources is the lack of continuity. This discontinuity problem stems from the fact that the energy produced by the renewable energy sources depends on the natural conditions [10]. It is not always possible to achieve continuity of energy produced from one source. It is thought that this problem will be minimized by the hybrid energy system by using more than one renewable energy source together [11, 12].

If the choice of renewable energy sources to be used on the buildings is made by taking into account the geographical location and climate conditions of the building with technical and economic analysis, the buildings may produce part of or all of their energy needs and may even be able to produce more than their own need and reduce the demand on the national network. Moreover, the transfer of excess of the generated electricity to the interconnection network with auxiliary technical equipment provides economical benefits to both the producers and the countries.

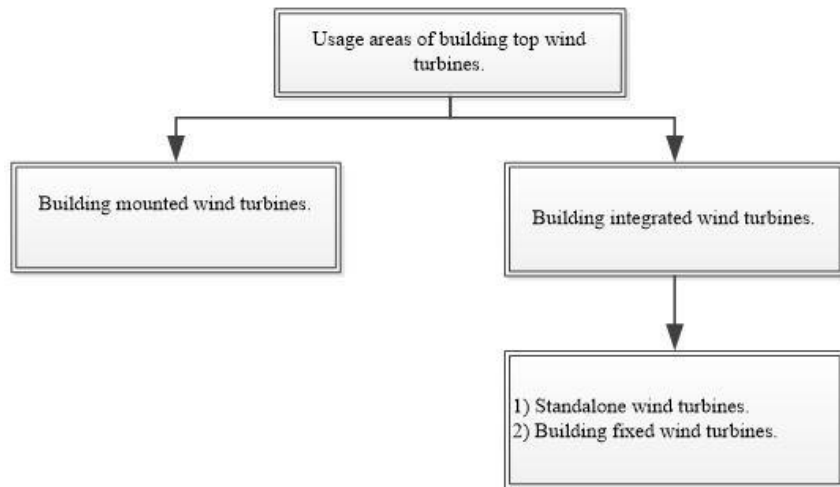
II. USE OF RENEWABLE ENERGY SOURCES IN BUILDINGS

A. Electricity Generation in Building Shell with PV Panels

PV panels were first applied to the buildings as an additional system in 1981, and later PV panels were produced which can be used directly as the roof covering. Since 1992, R&D studies have shown that PV panels can be used effectively in building vertical shells, and pilot applications are increasingly used in the form of building facades [13]. Nowadays PV panels can be produced in two different kinds, semi-permeable or opaque, and various colors. In practice, PV panels are applied directly on the roof, vertical walls and auxiliary areas such as rain covers or sun shades [9]. In Fig. 1 (a), the usage area of PV materials on the building is shown hierarchically.



(a)



(b)

Fig.1. (a)Usage areas of externally mounted PV systems for buildings (b) Usage areas of building top wind turbines.

B. Electricity Generation in Building Shell with Wind Generator

Medium and small-scale horizontal and vertical axis wind turbines are used in the buildings. Wind turbines can be located at a suitable point in the surroundings of the building, or they can be placed in the horizontal shell (roof) of the buildings. There are many applications for the use of integrated wind turbines in high buildings such as multi-story skyscrapers or residences [10]. Small-scale wind turbines have nominal generation values ranging from 50 W to 20 kW. At low wind speeds, these turbines continue to produce electricity at lower values [14]. As shown in Fig. 1 (b), wind turbines are placed on buildings either as an integral part of the building when it is constructed or mounted later on as an addition [9, 15].

Wind turbines, according to their construction, are divided into two categories; horizontal wind turbines and vertical wind turbines (straight or angled). The rotor of the horizontal axis wind turbines is parallel to the ground. They are economical if they are installed in areas with high wind speeds [16]. Horizontal axis turbines have drawbacks due to the flexing system and generally cannot follow rapid and comprehensive changes. The wings of the vertical axis wind turbine are fixed perpendicular to the central axis. They are usually used in areas with low wind speeds. In such wind turbines, the direction of the wind is not important. Thanks to the wing structures, they can easily use the winds coming from every direction. They are mounted directly on the ground (soil level, building base, etc.). The construction is simple and light, easy to service with low post cost [2]. In parallel with the development of wind turbines, the presence of small vertical wind turbines, which can adapt to the direction of the wind on the constructions and reduce losses, have made it possible to prefer these types of turbines for use on the buildings. Principal vertical axis wind turbines are Darrieus, Giromill, and Savonius [16]. It is understood that the most appropriate wind turbine that can be used in the light of the above information may be a small vertical wind turbine.

C. Electricity Generation by HRPSs

Because renewable energy sources and energy production depend on natural conditions, continuity in the produced energy is not always ensured. For this reason, the creation of a hybrid power system using multiple sources of energy is considered. Hybrid renewable power systems (solar, wind, biomass, hydraulics, etc.) are designed as complementary units to meet the energy demand continuously [17, 18]. To design a hybrid renewable power system, the energy generation system needs to know the potential of the existing renewable energy sources to be utilized. In addition, we need to know the energy demand at the facility. Therefore, energy planners need to investigate the sun, wind and other potential sources of energy in the region concerned, in addition to the planned use of energy. This approach can be used to design a hybrid power system capable of optimally meeting the energy demands of the respective facility or settlement [19].

III. HRPS COMPONENTS FOR IGUVS

A. Energy Claim

IGUVS is a 4-story building located in Istanbul Pendik. The latitude and longitude values at the time of construction are 40° 54' 3.6" and 29° 13' 8.4" respectively. Fig. 2 shows the direction of the building and the front view of the structure. The front of the building is facing south-southwest. The elevation is 14 m up to the starting point of the building, and the corner heights are 16 m. The width of the front is 52 m. The total active area to be used for the PV panels is calculated as about 400 m² when the windows and other gaps are deducted from the front edge of the structure. Satellite and front view of the IGUVS building can be seen in Fig. 2.



Fig.2. Satellite and front view of the IGUVS building.

The electricity consumption of the building was evaluated and the electricity consumption value was entered to the HOMER software as data on an hourly basis per year. The daily and hourly load values of IGUVS by using these values and HOMER software are shown in Fig. 3. HOMER software is a software that performs operations such as modeling, analysis and simulation of power systems, and sensitivity analysis, and was developed within the American National Renewable Energy Laboratory (NREL). The software allows the design of power systems as well as the comparison of power systems with each other. The hourly load data belonging to IGUVS was obtained by considering the monthly bills of the school. The daily rate is calculated as 376 kWh/day. The daily peak load is also calculated as 24 kW. While the average hourly minimum load demand of IGUVS is 9.40 kWh between 19:00 - 06:00 hours, maximum hourly load demand is 17 kWh between 16: 00-17: 00 hours.

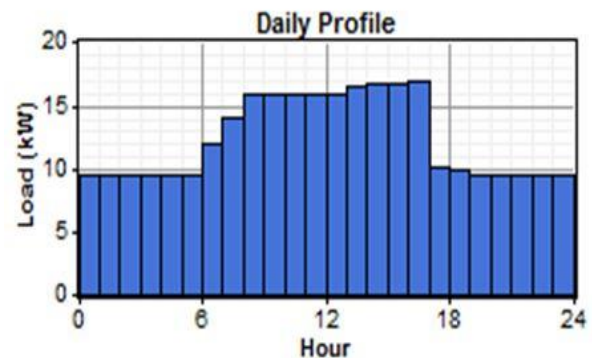


Fig.3. Hourly energy demand of the building.

B. Potential of Solar Energy

Monthly average solar radiation (kWh/m^2) values were taken from the nearest meteorological station to the province of Istanbul for the last 20 years [20]. Using these values, the solar radiation values and the clearness index of IGVUS using HOMER software are shown in Fig. 4. The annual average solar radiation value is calculated as $4.028 \text{ kWh}/\text{m}^2/\text{d}$. At HOMER, solar radiation data is synthesized using the Graham algorithm. The Graham algorithm is used by the HOMER program to obtain the solar radiation data entered into the program on a monthly average using latitude and longitude coordinates. The annual average clearness index is set at approximately 0.502.

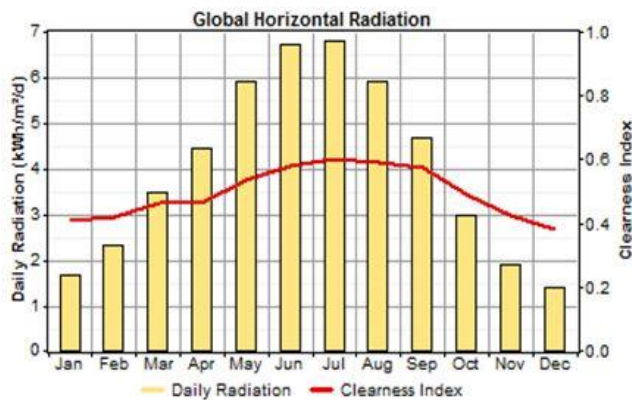


Fig.4. Monthly average solar radiation values.

C. Wind Power Potential

The monthly average wind speed (m/s) values for the last 20 years of Istanbul province are taken from the nearest meteorological station to the region [20]. The monthly average wind speed values of IGVUS using these values and HOMER software are shown in Fig. 5.

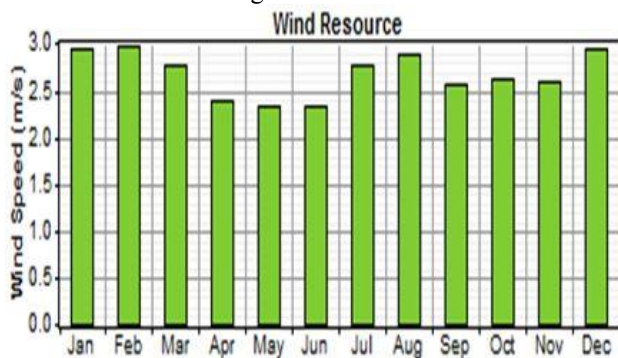


Fig.5. Average monthly wind speed.

D. Hybrid Power System Components

The components to be used in electricity production and their characteristics will be given in detail. The hybrid power system shown in Fig. 6 constitutes the energy production part, the wind turbine, and the solar panel unit. The system is powered by a grid to meet energy demands when wind and/or sun are inadequate [9]. In this hybrid power system, energy production is provided by wind turbines and solar panels. The

energy will be supplied from the hybrid power system when wind and/or the sun is built on the level that can produce electricity. When there is not enough wind or sun, the building buys the energy demand from the network. Therefore, a clean electricity generation system with continuity has been established. Energy produced more than the demand will be sold to the national network. Two-way counters are used in these systems for this purpose. The aim here is the measurement of electricity consumption or the sale of excess energy sold to the network. Turkish law numbered 6094 and titled "Law on the Amendment of the Law on the Use of Renewable Energy Sources for the Purpose of Electricity Generation Production" allows the sale of the excess electricity to the network eliminating the need for a battery bank for energy storage in this system, which leads to a decrease in system installation cost [21, 22].

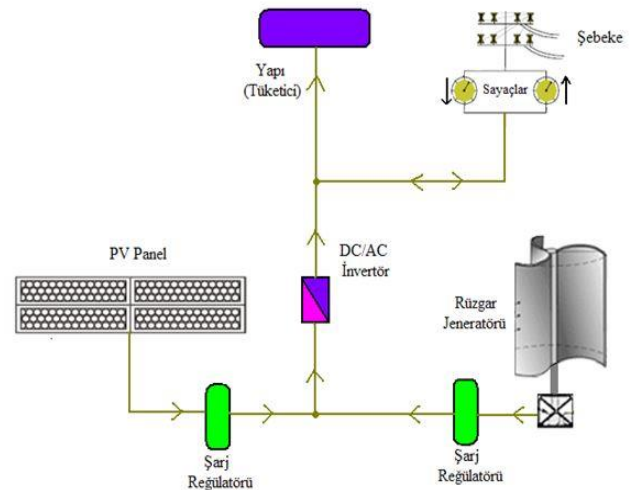


Fig.6. Single line schematic of a wind-solar hybrid power system.

1) PV Panel

The PV panel used in the hybrid power system is Sharp ND-Q235FD4-235 W. PV panel output power is 235 W. The PV panel is one of the main energy providers in the hybrid system. The PV panel only generates energy during the time period from 06:00 to 18:00 and does not generate energy at other time intervals. The cost of the PV panel is about \$ 500, and the replacement cost is equal to the initial cost. The cost of operation and maintenance is negligible [23]. The number of PV panels varies from 0 to 10.

2) Wind Turbine

The wind turbine used in the hybrid power system is one of the main energy providers of the system as is the PV panel. The wind turbine suggested here is a model that is not in the library of the HOMER software. The Savonius wind turbine with the vertical axis is used. The Savonius wind turbine consists of the data obtained as a result of tests performed in the laboratory environment with power parameters. It was entered as an external addition to HOMER. The Helix S594 model was used as the wind turbine. The initial cost of the wind turbine is about \$ 17,500. The output power of the wind

turbine is 4.5 kW. Wind turbine replacement costs were estimated at approximately \$ 15,000 and operation and maintenance costs at \$ 500 per year [24].

3) Converter

The power of the converter was chosen at the peak power of 24 kW. It is used to transfer the power obtained from PV panels. The conversion efficiency is 90% and the initial cost is 655 \$/kW. The replacement cost is equal to the initial cost. Operation and maintenance cost is negligible. Meanwell DC/AC Solar Inverter is preferred as the converter [25].

4) Network

If the renewable energy sources in the hybrid power generation system cannot meet the load demand, the grid will meet the demand. Thus, continuity of energy is ensured. There are three different tariffs, which are called "triple tariff" [26].

E. Operating Characteristics of the Hybrid Power System

Operating characteristics of the hybrid power system are as follows:

Wind turbine and PV panel are the main energy providers. While the wind turbine is directly feeding the load, power produced by PV panels reaches the load by means of the power converter.

If the wind turbine and the PV panel cannot meet the demand, the grid will meet the demand.

Project life is set at 25 years.

The renewable energy ratio in the hybrid power generation system is expressed as a minimum of 40%.

According to The Central Bank of the Republic of Turkey, the annual interest rate is 12% [27].

IV. DETERMINATION OF OPTIMUM HRPS FOR IGUVS

A. Sensitivity Analysis of HRPS

It is the evaluation of which hybrid system is more optimal from alternative hybrid power generation systems when the wind speed values change between 2.67 m/s and 5.5 m/s and the solar radiation values change between 3 kWh/m²/d and 5.5 kWh/m²/d. The wind speed values and solar radiation values used in the sensitivity analysis are given in detail in Table 1.

TABLE I
EVALUATION OF WHICH HYBRID SYSTEM IS MORE OPTIMAL IN WHICH CASE

Solar radiation values (kWh/m ² /d)	Wind speed values (m/s)
4.028	2.670
3.000	3.500
5.000	4.500
6.000	5.500

Besides, the optimum hybrid power generation system and optimal system configurations obtained in correspondence with the values given in Table 1 are detailed in Fig. 7. The optimum wind speed and solar radiation data for IGUVS are

4.028 kWh/m²/d and 2.670 m/s, respectively. Considering these values, it is seen that the optimum hybrid renewable power generation system is a grid-connected wind/solar power system. The optimum configuration of this system was obtained as a 30 kW PV panel, 22.5 kW wind turbine, 50 kW Grid, and a 24 kW converter.

Solar (kWh/m ² /d)	Wind (m/s)	PV (kW)	Wind (kW)	Conv. (kW)	Grid (kW)	Initial Capital	Operating Cost (\$/yr)	Total NPC	COE (\$/kWh)	Ren. Frac.	Capacity Shortage
4.028	2.670	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.56	0.00
4.028	3.500	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.67	0.00
4.028	4.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.63	0.00
4.028	5.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.72	0.00
4.028	6.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.76	0.00
3.000	2.670	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.51	0.00
3.000	3.500	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.64	0.00
3.000	4.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.63	0.00
3.000	5.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.72	0.00
3.000	6.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.76	0.00
5.000	2.670	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.59	0.00
5.000	3.500	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.70	0.00
5.000	4.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.63	0.00
5.000	5.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.72	0.00
5.000	6.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.76	0.00
6.000	2.670	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.60	0.00
6.000	3.500	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.71	0.00
6.000	4.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.63	0.00
6.000	5.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.72	0.00
6.000	6.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.76	0.00
7.000	2.670	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.61	0.00
7.000	3.500	30	5	24	50	\$ 235,720	3,953	\$ 286,247	0.164	0.71	0.00
7.000	4.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.63	0.00
7.000	5.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.72	0.00
7.000	6.500	30	5	24	50	\$ 40,720	1,573	\$ 60,824	0.035	0.76	0.00

Fig.7. Optimum hybrid power system options for each wind speed and solar radiation data.

Fig. 8 shows that the optimum system is the Grid/Wind hybrid power generation system when the wind speed is 4.5 m/s or higher with all values of the solar radiation. With lower wind speed, the optimum hybrid system will be the Grid/Wind/PV system.

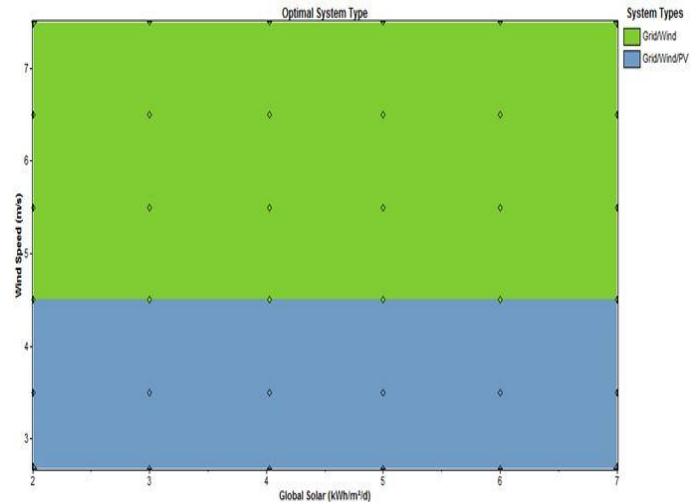


Fig.8 Sensitivity analysis.

In these cases, the total installation cost and the unit cost of energy are the lowest. It is the optimum value with a unit energy cost of 0.164 \$/kWh.

B. Economic Analysis of the Hybrid Renewable Power System

As a result of optimization performed considering the optimum values of wind speed and solar radiation data for Gedik Vocational School, unit energy cost is calculated as

0.164 \$/kWh. The initial capital was \$ 235,720 and the operating cost was \$ 4000. The total net present cost was \$ 286,242. The distribution of this cost among the components of the hybrid renewable power generation system is shown in Fig. 9 in detail.

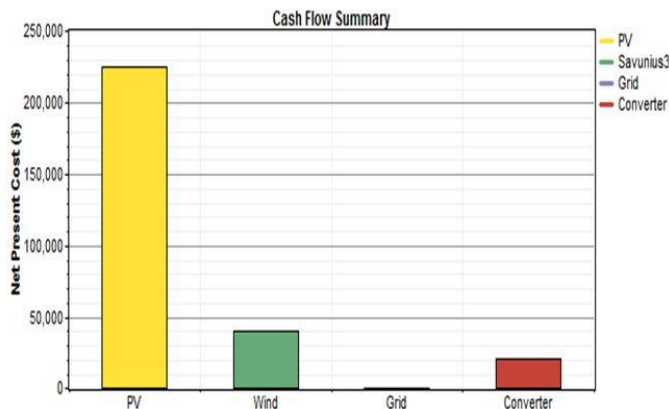


Fig.9 Cost summary.

C. Analysis of Electricity Generation and Consumption

As shown in Fig. 10, annual electricity generation from PV panels is 35,575 kWh/year while yearly electricity generation from wind turbines is 49,294 kWh/year.

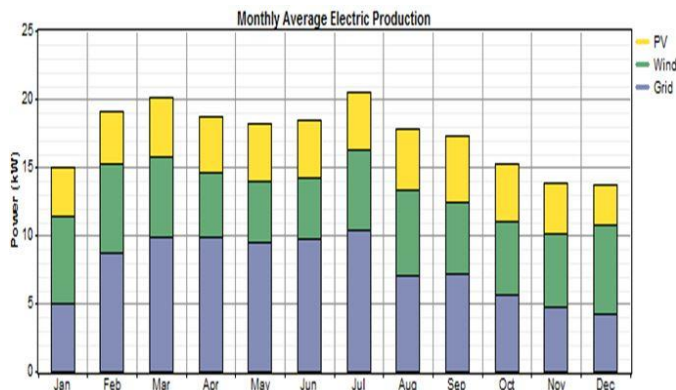


Fig.10 Monthly average electricity production by source types

Finally, the annual amount of energy supplied from the network stands at 66,834 kWh/year. The total amount of electricity needed for the building is 151,703 kWh/year. In this case, the PV panels provide 23% of the demand while the wind turbine provides 32% and the network 44%. In other words, 44% of the electricity demand is generated by the grid and 53.4% by the renewable energy sources. The annual electricity consumption of the building is 136,510 kWh/year, and the sales-ready excess electricity is 5,367 kWh/year. The residual energy is 2,405 kWh/year, 1.59%, which is in the acceptable range and corresponds to the other losses of the system.

D. Emission Status of Hybrid Power System

The construction requires 136,510 kWh/year of electrical energy per year. Emission values can easily be calculated in

HOMER software. Before simulating the power system, HOMER determines the emissions factor (kg of pollutant emitted per unit of fuel consumed) for each pollutant. After the simulation, it calculates the annual emissions of that pollutant by multiplying the emissions factor by the total annual fuel consumption [28]. When this energy is supplied from the network, atmospheric emissions of about 236 kg of greenhouse gas emission per day are achieved, with CO₂ emissions of 86,274 kg/year for SO₂ and 183 kg/year for NO₂. However, with the grid-connected hybrid renewable solar/wind power system, these values are 38047 kg/year CO₂ emissions of 168 kg/year SO₂ and 82.4 kg/year NO, 106 kg/day of carbon dioxide emissions per day. When the proposed grid-connected renewable solar/wind power system is preferred, there is a 45% reduction in greenhouse gas emission. It is also a fact that the system is a renewable, clean and environmentally friendly hybrid power system.

V. CONCLUSION

The findings obtained as a result of this study are given below:

It is planned to design vertical wind turbine on top of the Gedik Vocational School building and PV panels to be vertically mounted, unlike conventional horizontal installation. In this context, system modeling has been carried out.

Optimum wind speed and solar radiation data for Gedik Vocational School are 4.028 kWh/m²/d and 2.670 m/s respectively. Considering these values, it seems that the optimum hybrid renewable power generation system is a grid-connected wind / solar power system.

Optimum system configuration is a 30 kW PV panel, 22.5 kW wind turbine, 50 kW Grid, and 24 kW converter.

The unit energy cost of the optimal hybrid power generation system is calculated as 0.164 \$/kWh. Besides, the initial cost was \$ 235,720, and the operating cost was \$ 4,000. The total net present cost was \$ 286,242.

A reduction in greenhouse gas emission of 45% is obtained when a grid-connected hybrid renewable solar/wind power system is preferred. It is also a fact that the system is a renewable, clean and environmentally friendly hybrid power system.

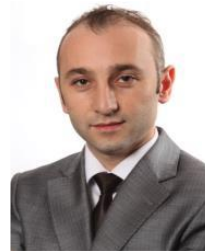
With the optimum hybrid power generation system, the renewable hybrid energy system meets the yearly electricity demand required by the building, with 23% solar panel, 32% wind turbine, and the renewable hybrid energy system meets 55% of the annual electricity demand of the designed hybrid system.

As a result of the sensitivity analysis, it is seen that the optimum system is the Grid/Wind hybrid power generation system when the wind speed is 4.5 m/s and above with all values of the solar radiation, and the optimum hybrid system in all other cases is the Grid/Wind/PV system.

REFERENCES

- [1] J.S. Ojo, P.A. Owolawi, A.M. Atoye, "Designing a Green Power Delivery System for Base Transceiver Stations in Southwestern Nigeria", SAIEE Africa Research Journal, Volume 110, March 2019, pp.19-25.
- [2] S. Sayın, "The Significance of the use of renewable energy in our country's building sector and the opportunities of utilizing of solar energy in buildings", Selcuk Univ. M.Sc. Thesis; 2006.
- [3] T. Adefarati, R.C. Bansal, "Application of renewable energy resources in a microgrid power system", The Journal of Engineering, 2019, pp.5308-5313.
- [4] I. Das, C.A. Cañizares, "Renewable Energy Integration in Diesel-Based Microgrids at the Canadian Arctic", Proceedings of the IEEE, Volume 107, Sept. 2019, pp. 1838-1856.
- [5] M. Altın, "Research on the Architectural Use of Photovoltaic (PV) Components in Turkey from the Viewpoint of Building Shape", Dokuz Eylul Univ. Ph.D. Thesis; 2005.
- [6] H.U.R. Habib, S. Wang, M.R. Elkadeem, M.F. Elmorshedy, "Design Optimization and Model Predictive Control of a Standalone Hybrid Renewable Energy System: A Case Study on a Small Residential Load in Pakistan", IEEE Access, Volume 7, August 2019, pp. 117369-117390.
- [7] IEA Renewable energy: 2018. <https://www.iea.org/topics/renewables/Accessed08.07.2018>
- [8] R. Singh, R.C. Bansal, "Optimization of an Autonomous Hybrid Renewable Energy System Using Reformed Electric System Cascade Analysis", IEEE Transactions on Industrial Informatics, Volume 15, Jan. 2019, pp.399-409.
- [9] O. Akar, "Technical and Economical Analysis of Renewable Energy Resources Based Electricity Generation And Usage In Smart Buildings", Marmara Univ. M.Sc. Thesis; 2011.
- [10] I. Yuksek, T. Esin, "Possibilities of Using Renewable Energy Resources in Buildings", 5th International Advanced Technologies Symposium (IATS' 09), Karabuk, 2009.
- [11] A. Parida, S. Choudhury, D. Chatterjee, "Microgrid Based Hybrid Energy Co-Operative for Grid-Isolated Remote Rural Village Power Supply for East Coast Zone of India", IEEE Transactions on Sustainable Energy, Volume 9, no 3, July 2018, pp.1375-1383.
- [12] A. Kumar, A.R. Singh, Y. Deng, X. He, P. Kumar, R.C. Bansal, "Multiyear Load Growth Based Techno-Financial Evaluation of a Microgrid for an Academic Institution", IEEE Access, Volume 6, June, 2018, pp.37533-37555.
- [13] G. Celebi, "Principles of Use of Photovoltaic Panels in Building Vertical Shells", Gazi University Journal of Faculty of Engineering and Architecture, 2002, pp. 17-33.
- [14] G. Kocaslán, "Energy resources in Turkey and the research of wind energy as an alternative technology", Istanbul Univ. M.Sc. Thesis, 2002.
- [15] M.H. Gunel, H.E. Ilgin, "Use of Wind Energy as an Architectural Design Criteria", Journal of Ege Architecture. Izmir, 6-11; 2008.
- [16] G.V.J. Bussel, S.M. Mertens, "Small Wind Turbines for The Built Environment. The Fourth European & African Conference on Wind Engineering", ITAMASCR, 11-15 July, 1-9; 2005.
- [17] B. Dursun, "Determination of the optimum hybrid renewable power generating systems for Kavakli Campus of Kırklareli University", Turkey. Renewable and Sustainable Energy Reviews. 2012; 16, pp.6183-6190.
- [18] S.B. Efe, B. Kocaman, "Physical Realization and Analysis of Renewable Energy-Based Hybrid System", 5th International Symposium on Innovative Technologies in Engineering and Science 29-30 September 2017 (ISITES2017) Baku, Azerbaijan, 2017.
- [19] B. Dursun, C. Gokcol, I. Umut, E. Ucar, S. Kocabey, "Techno-Economic Evaluation of a Hybrid PV-Wind Power Generation System", Intern. Journ. of Green Energy, 2013; pp. 117-136.
- [20] Istanbul Meteorology 1st Division Directorate, Monthly average wind speed and global radiation values of the last 20 years, Turkish State Meteorological Service Report, Istanbul, 2011.
- [21] N. Koca, "Using hybrid energy at home", Sakarya Univ. M.Sc. Thesis, 2006.
- [22] O. Akar, U.K. Terzi, "Building Based Micro Distributed Power Generation Systems", 3e Electrotech Journal. 223, pp. 210-220, 2013.22
- [23] Sharp Electronics Corporation, 2015. http://files.sharpusa.com/Downloads/Solar/Products/sol_dow_NDQ235F4.pdf. Accessed 08.07.2018
- [24] Swepic, 2018. <http://www.swepinc.com/pdf/Wind%20Turbines/Vertical%20Axis%20Wind%20Turbines/Helixwind%20S594.pdf>. Accessed 08.07.2018
- [25] Mean Well, 2018. <http://www.meanwell.com/webapp/product/search.aspx?prod=tn-3000>, Taiwan, Accessed 08.07.2018
- [26] P. Yilmaz, M.H. Hocaoglu, A.E.S. Konukman "A Pre-feasibility Case Study on Integrated Resource Planning Including Renewables", Energy. Policy, 2008.
- [27] The Central Bank of the Republic of Turkey Exchange Rates, 2018. <http://www.tcmb.gov.tr/wps/wcm/connect/en/tcmb+en>. Accessed 08.07.2018
- [28] How HOMER Calculates Emissions, 2019. https://www.homerenergy.com/products/pro/docs/latest/how_homer_calculates_emissions.html. Accessed 28.10.2019

BIOGRAPHIES



Onur AKAR was born in Giresun in 1981. He received his B.S. degree from Electrical Education Department of Technical Education Faculty of University of Marmara, Istanbul, in 2005 and his M.Sc. degree from Electrical Education Department of Institute of Pure and Applied Sciences, University of Marmara, Istanbul, in 2011. He also

received B.S degree from Electrical and Electronics Engineering Department of Engineering Faculty of Karadeniz Technical University, Trabzon, in 2017. He has been currently attending doctorate programme in Electrical and Electronics Engineering department of Institute of Pure and Applied Sciences of Marmara University, Istanbul.

He has been working as a Lecturer at Mechatronics Program of Vocational High School of Istanbul Gedik University since 2010. He served as the Head of Electrical Department of Vocational High School of Gedik University in Istanbul between 2012 and 2015. He is the author of 2 article and one book chapter. His research interests include Control Systems, Renewable Energy Systems and Power Systems.



Umit K. TERZI was born in Zonguldak in 1968. He received his B.S. degree from Electrical Education Department of Technical Education Faculty of University of Marmara, Istanbul, in 1989 and his M.Sc. and Ph.D. degree from Electrical Education Department of Institute of Pure and Applied Sciences, University of Marmara, Istanbul, in 1994 and 2000 respectively.

From 1989 to 1996, he worked as a research assistant, from 1996 to 2000 as a lecturer, from 2000 to 2013 as an Assistant Prof. Dr. for university of Marmara. Since 2013 he has been working as an Associate Prof. Dr. for Electrical and Electronics department of Technology Faculty of Marmara University and Electrical Education Department of Technical Education Faculty where he is head of department.

His research interests include Electrical Machinery, Power Systems, Energy Transmission and Distribution, Renewable Energy Systems.



B. Koray TUNCALP was born in Istanbul in 1962. He received his B.S. degree from Electrical Education Department of Technical Education Faculty of University of Marmara, Istanbul, in 1984 and his M.Sc. and Ph.D. degree from Electrical Education Department of Institute of Pure and Applied Sciences, University of Marmara, Istanbul, in 1988 and 1999 respectively.

From 1984 to 1993, he worked as a research assistant, from 1993 to 1999 as a lecturer, from 1999 to 2001 as an Assistant Prof. Dr., from 2001 to 2003 as an Associate Prof. Dr., from 2003 to 2014 as an Prof. Dr. in two departments for University of Marmara. Since 2015 he has been working as an Prof. Dr. for Electrical and Electronics Engineering Department of School of Engineering, also is Director of Vocational School of Halic University.

His research interests include Electrical Energy Measurement, Mechatronics, Instrumentation, E-learning, Vocational and Technical Education.



Temel SONMEZOCAK was born in Edirne in 1978. He graduated his B.S. degree from Electrical Education Department of Technical Education Faculty of University of Marmara, Istanbul, in 2001. In 2015, he completed his M.Sc. degree at Istanbul Aydin University, Department of Electrical and Electronics Engineering. In 2016, he started his Ph.D. program in

Electronics Engineering at Yildiz Technical University and still continues. Since 2018, he has been working as a lecturer in Electrical department of Istanbul Gedik University.

His research interests include lighting systems, power electronics and renewable sources.

Publication Ethics

The journal publishes original papers in the extensive field of Electrical-electronics and Computer engineering. To that end, it is essential that all who participate in producing the journal conduct themselves as authors, reviewers, editors, and publishers in accord with the highest level of professional ethics and standards. Plagiarism or self-plagiarism constitutes unethical scientific behavior and is never acceptable.

By submitting a manuscript to this journal, each author explicitly confirms that the manuscript meets the highest ethical standards for authors and coauthors

The undersigned hereby assign(s) to *Balkan Journal of Electrical & Computer Engineering* (BAJECE) copyright ownership in the above Paper, effective if and when the Paper is accepted for publication by BAJECE and to the extent transferable under applicable national law. This assignment gives BAJECE the right to register copyright to the Paper in its name as claimant and to publish the Paper in any print or electronic medium.

Authors, or their employers in the case of works made for hire, retain the following rights:

1. All proprietary rights other than copyright, including patent rights.
2. The right to make and distribute copies of the Paper for internal purposes.
3. The right to use the material for lecture or classroom purposes.
4. The right to prepare derivative publications based on the Paper, including books or book chapters, journal papers, and magazine articles, provided that publication of a derivative work occurs subsequent to the official date of publication by BAJECE.
5. The right to post an author-prepared version or an official version (preferred version) of the published paper on an internal or external server controlled exclusively by the author/employer, provided that (a) such posting is noncommercial in nature and the paper is made available to users without charge; (b) a copyright notice and full citation appear with the paper, and (c) a link to BAJECE's official online version of the abstract is provided using the DOI (Document Object Identifier) link.



ISSN: 2147- 284X
Year: October 2019
Volume: 7
Issue: 4

CONTENTS

D. İzci; Constructing an Electronic Circuitry for Label-free Hall Biosensors	366-372
Ş. Eraslan, S. Karabulut, M. C. Atalay, Y. Yeşilada; Evaluation of Visualisation of Scanpath Trend Analysis (ViSTA) Tool.....	373-383
H. Gençoğlu, T. Yerlikaya; Three Part Hybrid Encryption Schema.....	384-390
D. Köseoğlu, O.Chouseinoglu; Opinion Analysis of Software Developers Working Onsite on Public Sector Software Projects: An Exploratory Study.....	391-398
M. F. Aslan, K. Sabancı, A. Durdu; Comparison of Contourlet and Time-Invariant Contourlet Transform Performance for Different Types of Noises.....	399-404
B. Dogan; An alignment-free method for bulk comparison of protein sequences from different species.....	405-416
M. Akgün; An Active Genomic Data Recovery Attack.....	417-423
H. F. Carlak, E. Kayar; Volt / VAR Regulation in Energy Transmission Systems Using SVC and STATCOM Devices.....	424-433
H. Ademgil; Simultaneous Sensing of Dual Analyte Photonic Crystal Fiber Based Liquid Sensor.....	434-439
Y. Acar, S. Aldırmaz Çolak; Novel OFDM System Using Orthogonal Pilot Symbols with Subcarrier Index Modulation.....	440-445
C. Bakır, V. Hakkoymaz, B. Diri, M. Güçlü; Comparisons on Intrusion Detection and Prevention Systems in Distributed Databases.....	446-455
O. Akar, Ü. K. Terzi, B. K. Tunçalp, T. Sönmezocak; Determination of the Optimum Hybrid Renewable Power System: A case study of Istanbul Gedik University Gedik Vocational School.....	456-463

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING

(An International Peer Reviewed, Indexed and Open Access Journal)

Contact

Batman University
Department of Electrical-Electronics Engineering
Bati Raman Campus Batman-Turkey

Web: <http://dergipark.gov.tr/bajece>
<http://www.bajece.com>
e-mail: bajece@hotmail.com

