ISCTurkey 2015

EUROPEAN CYBER SECURITY MONTH

VIII. ULUSLARARASI BİLGİ GÜVENLİĞİ VE KRİPTOLOJİ KONFERANSI

VIIIth INTERNATIONAL CONFERENCE ON INFORMATION SECURITY AND CRYPTOLOGY

Cyber Security and Critical Infrastructure Siber Güvenlik ve Kritik Altyapılar

30-31 October 2015,
METU Cultural and Convention Center, Ankara, Turkey

30-31 Ekim 2015,
ODTÜ Kongre ve Kültür Merkezi, Ankara, Türkiye

Destekleyen Kuruluş

enisa
European Network and Information Security Agency

Organizers Düzenleyen Kuruluşlar

BİLGİ GÜVENLİĞİ DERNEĞİ

BTK
BİLGİ TEKNOLOJİLERİ VE İLETİŞİM KURUMU

ODTÜ

www.iscturkey.org

BİLDİRİLER KİTABI

BİLGİ GÜVENLİĞİ DERNEĞİ

8. ULUSLARARASI BİLGİ GÜVENLİĞİ VE KRİPTOLOJİ KONFERANSI

30-31 EKİM 2015

BİLDİRİLER KİTABI

ISBN: 978-605-86904-3-1

İçindekiler

ANALYSIS OF INFORMATION LEAKAGES ON LASER PRINTERS IN THE MEDIA OF ELECTROMAGNETIC RADIATION AND LINE CONDUCTIONS

Cihan Ulaş, Ulaş Aşık, and Cantürk Karadeniz

Abstract — In this paper, the emissions of a laser printer, which may process classified information, are investigated in the media of electromagnetic radiation (ER), Power Line Conductors (PLC), and Signal Line Conductors (SLC). First, the candidate frequency points of CE are examined in the frequency domain. Second, the emitted signal is AM-demodulated with the proper bandwidth, and then sampled by a high storage oscilloscope in these frequency points. Third, the collected data is converted to 2D image by applying signal and image processing techniques. In addition, this study introduces some practical measurement methods to reveal the possible CEs of laser printers. Finally, the procedure of the image reconstruction of CEs of the laser printer data is explained in detail.

Index Terms — compromising emanations, information leakages, printers, TEMPEST, electromagnetic radiation, power and signal line conductors.

I. INTRODUCTION

ELECTRONIC equipment naturally emits electromagnetic (EM) waves during its regular operation. Unintentional intelligence bearing signals may disclose the processed information which might be transmitted, received, or processed by any information processing equipment. If the information is classified as confidential, a serious information security weakness is occurred. There have been many studies on the subject of Compromising Emanations (CE) and information leakages caused by the information technology equipment such as computers displays, keyboards, and printers.

Harold Joseph Highland mentioned about the computer security risk of electromagnetic radiation in 1967 [1]; however, the first detailed open publication about compromising emanation risks was released by a Swedish government committee in 1984 [2]. Wim van Eck reconstructed Cathode Ray Tubes (CRT) screen information and displayed on a television monitor by using commercial equipment in 1985 [3]. Moreover, information leakages of other computer units and peripherals, such as keyboard and printers have been studied in the literature.

Keyboards are mostly used as input devices for confidential data such as passwords and text documents entry. Measurements and analyses on CE of keyboards were started with Han in early 1990s [4]. Vuagnoux et al. used an effective method to deal with the keyboards emissions and recovered keyboard entry from a distance around 20 meters with 95% success [5]. Zhang studied on compromising mechanism of the keyboards and compared the emanations among various keyboard types [6]. Kuhn studied on many researches in the field of CE of CRT displays, laptop displays, and flat panel displays. Kuhn reconstructed a CRT display image from three meters away [7]. Then laptop displays and flat panel displays are studied, and target display images are reconstructed successfully [8, 9].

Printers are also used as output devices for computer systems. Acoustic emanations of printers were studied in 1991, and the letters “W” and “J” were distinguished successfully [10]. An attack method has been presented which is based on the recording of the sound of a dot matrix printer processing English text [11]. Up to %72 of the printed words were recovered and the attack achieved recognition rate up to 95 % with the assumption of the knowledge about the text. Tosaka et al. studied the CE of laser printers, and they measured the magnetic field of a laser printer in the near field and achieved to reconstruct the printed image [12]. Przesmycki used some special Test Patterns (TP) to improve to the measurements of CE of monochromatic laser printers [13]. He presented the oscillograms of three lines that placed on different places of white sheet.

In this paper, CE of a laser printer is investigated in the media of power line conductors, signal line conductors, and electric radiation (ER). While the most of the studies focuses on the measurements of CE in the media of ER, in this study, it is also shown that the risk of information leakages in the media of power and signal lines (like USB) cannot be ruled out. In addition, we introduce a practical approach of searching CE using a conventional spectrum analyzer. In this approach, it is shown that the configuration of resolution bandwidth, frequency span and the sweep time has to be applied properly. Moreover, to be able to analyze and detect the CE frequency points more conveniently, new image patterns are proposed in addition to ones introduced by Przesmycki [13].

In the next section, the emission measurement setups in ER, PLC, and SLC are given. In Section 3, the approach for the searching CE and the evaluation of the TP are discussed. The image reconstruction method from the CE of the printer emissions is explained section 4. Finally, the paper is concluded in Section 5.

II. EMISSION MEASUREMENT SETUP

Visual assessment of emissions is really difficult in the case of a laser printer. The video signal is sent at a specific time in the printing process and activation time is limited. The same is true for the laser exposure system. Another difficulty is the noise produced by the sub-system used at the printing process, which could mask the CE.

The method used in this study is based on the study descripted in [13], which consist of using a spectrum analyzer and an oscilloscope. If the settings of wideband receiver system are not appropriate, detection of CE is almost impossible. Signal detection process consists in scanning the whole frequency range of the measurement setup and searching for any variation in the emanation that is related to the TP. When relation is determined, emanation is demodulated according to its amplitude modulation (AM)

as descripted in [13] and [5]. The spectrum analyzer is used to identify the data-related emanations and a spectrum analyzer and an oscilloscope are used together to decide whether the relation is a compromising emanation or not. In this way, the number of frequency points to be investigated is reduced significantly.

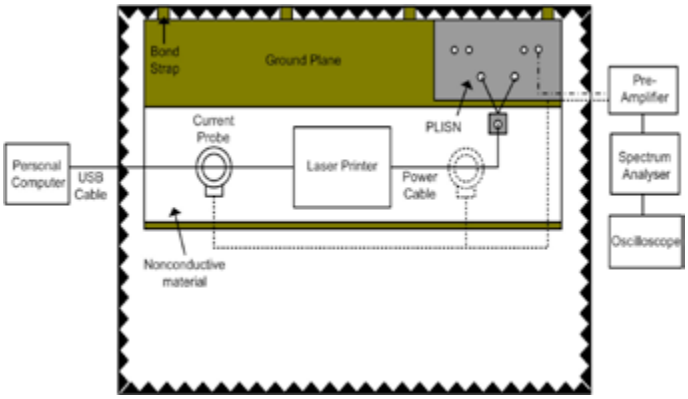


Fig. 1. Measurement setup of the laser printer's power leads and USB cable in the FAR.

Tests are conducted in a Fully Anechoic Room (FAR) as shown in Fig. 1 and Fig. 2. The method is also compared with the method proposed by Przesmycki [13] in the media of Electric Radiation (ER). A personal computer used to send TP and the measurement system is located in the control room. The laser printer is placed on a table in the FAR as shown in Fig. 1 and Fig. 2. FSET 22 is used as spectrum analyzer and the video output of the receiver is connected to the high storage oscilloscope. Here, the video output provides the AM-demodulated data, which is then sampled by the oscilloscope.

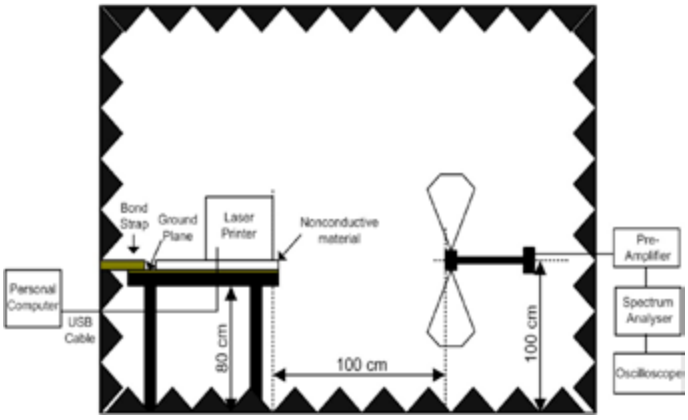


Fig. 2. Measurement setup for ER tests and antenna position in the FAR.

Searches are also performed for CE conducted on power line and signal line over 100 kHz to 1 GHz. General line-conduction measurement setup is shown in Fig. 1. While SLC measurements are performed using only current probe, PLC measurements are performed using either a current probe or a Power Line Impedance Stabilization Network (PLISN). In PLC measurements, both PLISN and current probe are used to differentiate emanations from the power cable and power leads.

ER measurements are performed over the frequency range 10 kHz to 2 GHz. In ER measurements, three types of antenna is utilized, which are rod antenna in the frequency range of 10 kHz - 30 MHz, biconical antenna in the frequency range of 30 MHz - 300 MHz, and log periodic antenna in the frequency range of 300 MHz - 2 GHz. In all measurements with biconical and log periodic antennas, they are polarized vertically and horizontally and positioned 1 meter away from the front edge of the setup boundary and 1 meter above the floor as shown in Fig. 2.

In our study, we printed the same pattern multiple times to ensure the emanations from the laser printer to be an unchanging input signal. Analysis with the swept spectrum analyzer needs time ranging 5 milliseconds to several seconds to sweep across the frequency span. This approach is based on the assumption that the input signal is not changed significantly in the time it takes to complete a sweep of the analyzer. We determined that 10 pages per pattern are adequate to assure the emanations as a static input signal.

Another way to improve the input signal quality is to use appropriate TP whose characteristics could be easily detected from the emanations by visual assessment. TP shall be simple but at the same time shall be complex enough to be noticeable in the noisy spectrum. The TP (IV, V, and VI) used in this study are shown in Fig. 3. The emissions from printing Pattern I are used as reference to distinguish data-related emissions from the other patterns in every measurement as in [13].

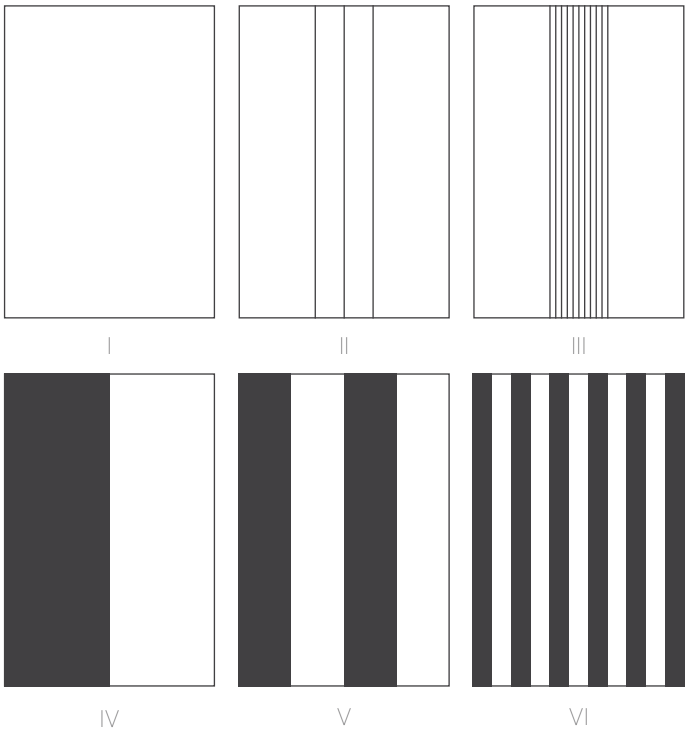


Fig. 3. Test patterns used in this study.

III. THE APPROACH FOR SEARCHING CE OF LASER PRINTERS AND THE EVALUATION OF TEST PATTERNS

In this section, we describe our approach to capture the data-related emanations and to evaluate whether the data-related emanation is a compromising emanation or not. The method is used in ER, PLC, and SLC media, and the obtained results are given. The analyses carried out in frequency domain are explained in this section.

The print speed of current laser printers is nearly 30 pages per minute. It means that it is necessary to sweep frequency span multiple times in 2 seconds to catch the CE. Sweep time (ST) is the most critical parameter in the all setting of the analyzer because of an absent of a mechanism to trigger the analyzer as in the case of laser printer tests. The ST is dependent on the resolution bandwidth (RBW), frequency span, and the design of the spectrum analyzer. For a near-Gaussian-shaped analog RBW filters, the relation between ST, span, and the RBW can be obtained through Equations (1-3).

$$t_p = \frac{RBW}{Span} ST \quad (1)$$

where t_p is the time in pass band. This time can be approximated to rise time tr of the filter which is inversely proportional to the bandwidth of the filter as

$$t_r = k \frac{1}{RBW} \quad (2)$$

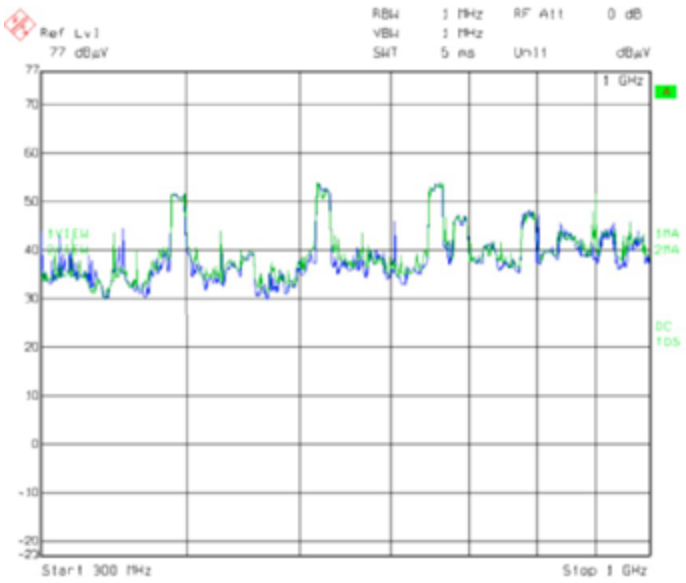
If the terms $t_r = t_p$ are equalized and solved for ST, the following relation is obtained.

$$ST = k \frac{Span}{RBW^2} \quad (3)$$

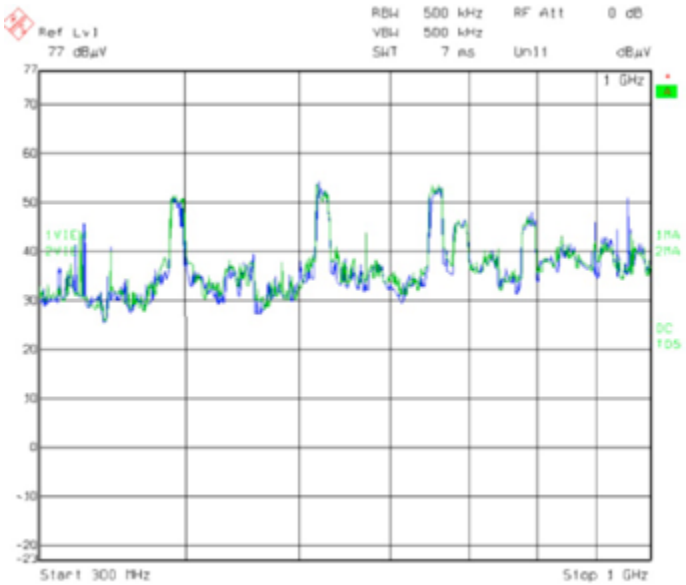
where k is the constant of proportionality.

There is a tradeoff among frequency selectivity, which can be improved by reducing RBW, signal-to-noise ratio (SNR), frequency span and measurement speed as seen in Equation 3. As RBW is narrowed at a fixed frequency span, the displayed average noise level of the spectrum analyzer is lowered. SNR and the selectivity are improved but the sweep time and trace update rate are degraded. For modulated signals, it is important to set the RBW wide enough to include the sidebands of the signal to make the measurement accurate. The optimum choice of the spectrum analyzer setting depends heavily on the characteristics of the signals of interest. In this study, we try to determine the most suitable spectrum analyzer setting to capture the CE from the laser printer.

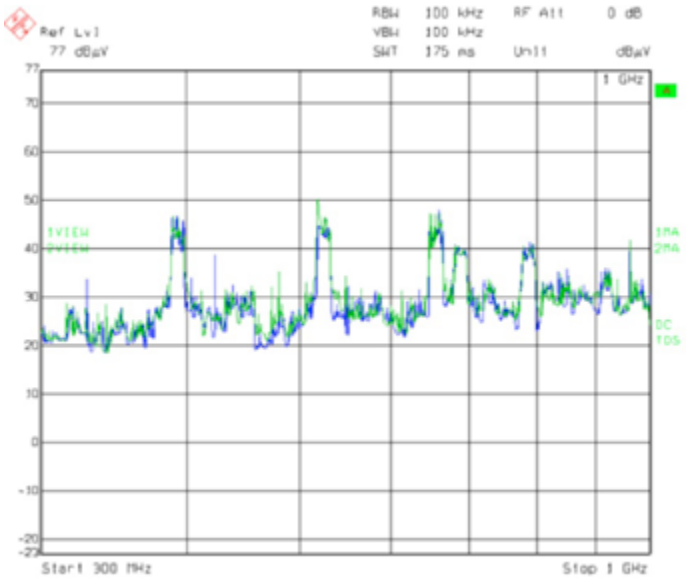
In the first study, we set RBW to 500 kHz as in [13] and sweep the 300 MHz – 1 GHz frequency band. Pattern I and Pattern II are used for the test as [13]. Fig. 4(b) gives the measurement results. As seen from the figure that there is no apparent frequency points or bands to be evaluated as the data-related emanation. We change RBW value and repeat the measurements to find a narrower frequency range to start with. Measurement results are presented in Fig. 4. At first step we chose RBW as 1 MHz. Emissions related to Pattern I and Pattern II are become more similar than the emissions measured when we set RBW to 500 kHz as seen in Fig. 4(a). Assigning the data-related emanation is become more difficult as we increase RBW value. We decide to use RBW values lower than 500 kHz.



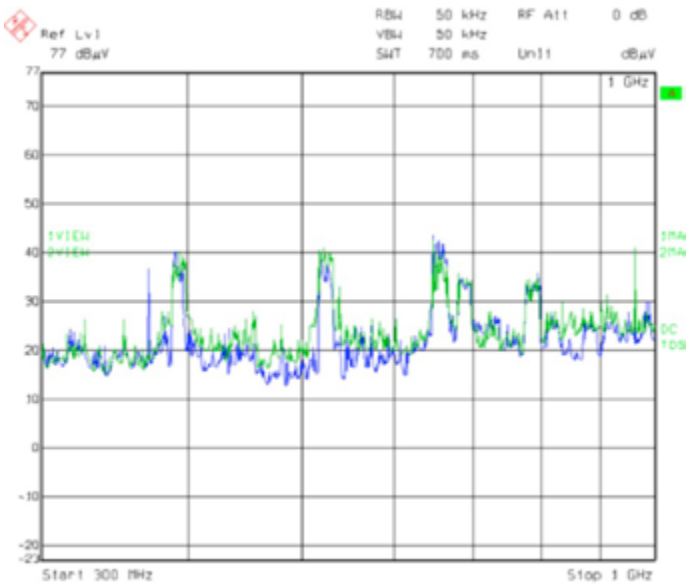
a



b



c



d

Fig. 4. ER test results while printing patterns, Blue: Pattern I, Green: Pattern II; (a) RBW = 1 MHz. (b) RBW = 500 kHz. (c) RBW = 100 kHz. (d) RBW = 50 kHz.

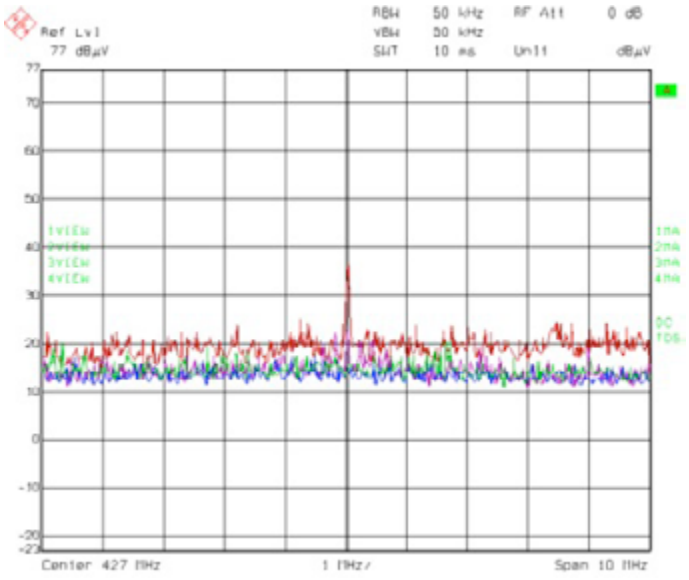
The measurement results as RBW is set to 100 kHz and 50 kHz are presented respectively in Fig. 4(c) and Fig. 4(d). There is no improvement on determining the date-related emanation between the measurement results as RBW is set to 100 kHz and 500 kHz as seen in Fig. 4(c) and Fig. 4(b). Emissions related to Pattern I and Pattern II could only be distinguishable when RBW is set to 50 kHz as seen in Fig. 4(d).

The proposed minimum value for RBW is 50 kHz because lower values make the sweep time of 300 MHz – 1GHz frequency span longer than print time of one page. As we examine Fig. 4(d), pattern emissions are explicitly differentiated from each other in following sub-bands: 400 MHz – 435 MHz, 460 MHz – 620 MHz and 820 MHz – 880 MHz.

In the second study, we make searches in the sub-bands listed above to capture the emanations related to the video signal. We choose the span value between multiples of 5 MHz to cover the sub-bands properly. We start with 10 MHz span and try to find the optimum span value in the following studies. The results are given in Fig. 5. Emanations related to empty page and Pattern II are nearly the same in all sub-bands. The difference between empty page and Pattern III is negligible in all span steps. Possibly, the number of lines and the thickness of each line at Pattern II and Pattern III would not be enough to radiate the data-related emanations that would pass the noise level at 50 kHz RBW.

TP IV, V, VI are compared with TP I, II, III. Although the difference among empty page and TP I, II, III are not distinguishable, for the proposed patterns the emission difference is around 15 dB in all the data-related frequency points as shown in Fig. 5. Pattern VI is the most suitable candidate for searching the data-related frequency points in oscilloscope to decide whether it is a compromising emanation or not since it can be identified under the high level noise due its high frequency content. Therefore, in the rest of the study, Pattern VI is used. For this particular case, while the emissions in 400 MHz – 435 MHz and 460 MHz – 620 MHz subbands contains the data-related emanations, in 820 MHz – 880 MHz sub-band, any data-related emanation frequency point isn't found. This difference might be result from the laser assembly system that are not used while printing empty pages.

In the second step, we repeat the tests with empty page and Pattern III by setting RBW to 5 kHz. Maximum frequency span is set to 5 MHz due to the limit on the print time.



a

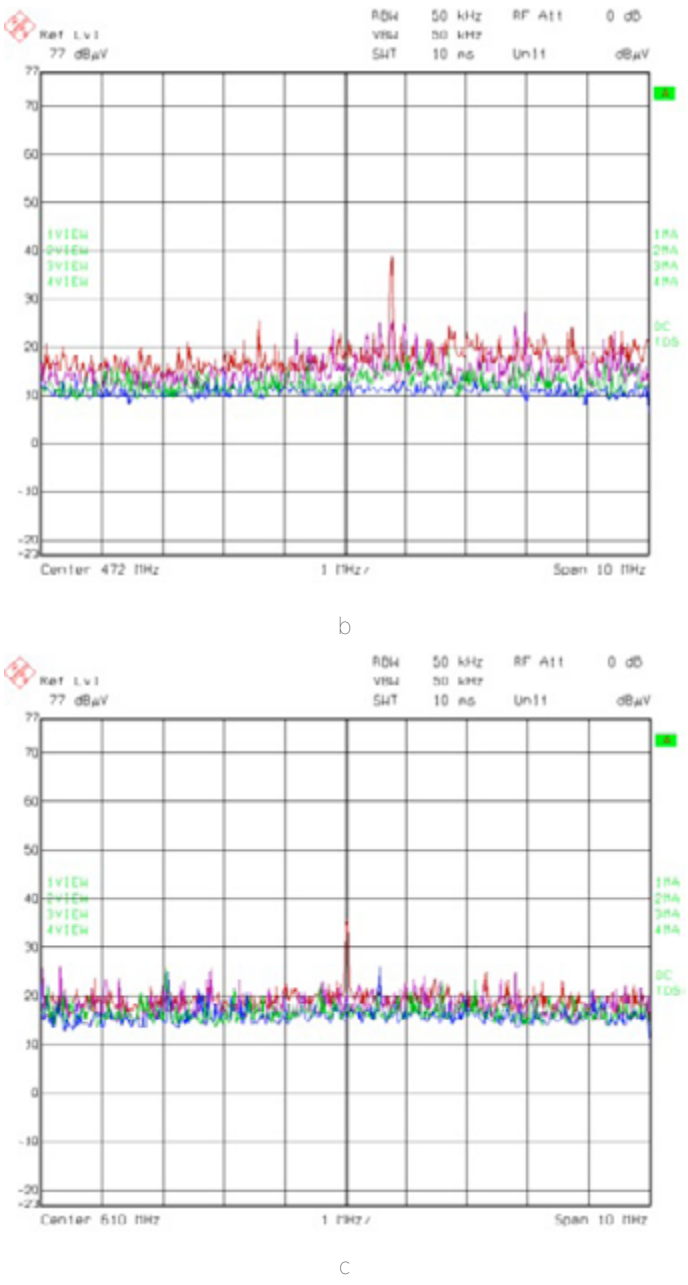


Fig. 5. ER correlated to TP. Blue: Pattern I, Green: Pattern II, Magenta: Patern III, Red: Same for Pattern IV, V, and VI.

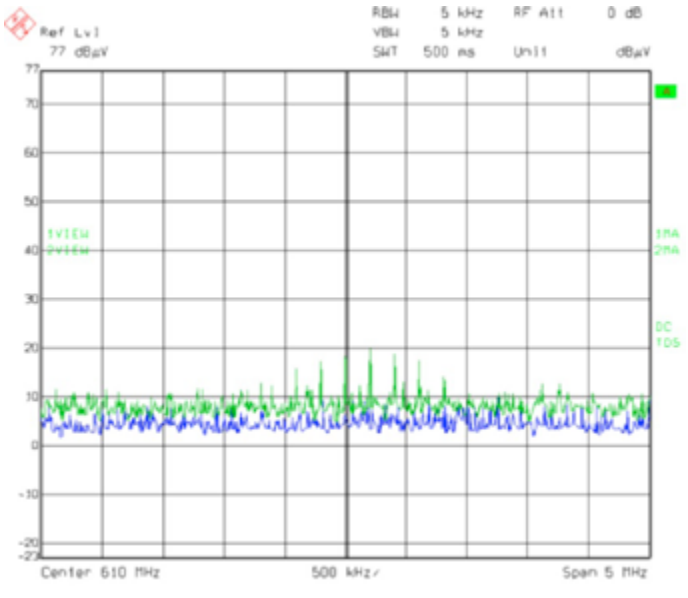


Fig. 6. ER correlated to test patterns; Blue: Pattern I, Green: Pattern III.

The results are presented in Fig. 6. SNR is improved as 10 dB and peaks related to Test Pattern (TP) become visible. In some sub-bands improving SNR as 10 dB is not enough to make peaks visible. We narrow the span and decrease the RBW. This approach increases the search time of the whole frequency band. On the other hand, Pattern VI allows us to increase the RBW and frequency span as shown in Fig. 7. Although the radiations related to Pattern III aren't visible as given in Fig. 6, the radiation related to Pattern VI improves the SNR without lowering the RBW. High RBW allows us to increase the frequency span; thus, the search time of the whole frequency band is decreased significantly. In the experimental tests, it is observed that the span width bigger than 50 MHz makes it difficult to decide whether an emanation is data-related or not because the number of peaks increases. As the number of peaks increases, the number of tests increases to determine whether an emanation at the peak point is data-related or not. We set RBW to 20 kHz to improve SNR. RBW values lower than 20 kHz with 50 MHz span makes sweep time longer than print-time of one page.

In the third study, we aim to determine whether sweeping the whole band as in the first study or sweeping the whole band by dividing into sub-bands. We divide 300 MHz – 1 GHz band into 50 MHz sub-bands and set RBW to 20 kHz. We find the data-related emanation points in the 700 MHz – 750 MHz sub-band which is not explicit in Fig. 4(d). Therefore, it is considered that the best way to search of the data-related emanations of laser printers is to divide the test frequency range into sub-bands and to use RBW as minimum as possible.

IV. RECONSTRUCTION FROM THE EMISSION OF PRINTER DATA

The method proposed in the preceding section provides us the CE in the analyzed frequency points. However, to be able to obtain and certify a human readable document, the AM-demodulated data in these frequencies are sampled with a high storage oscilloscope, and the document printed is reconstructed by applying signal processing techniques.

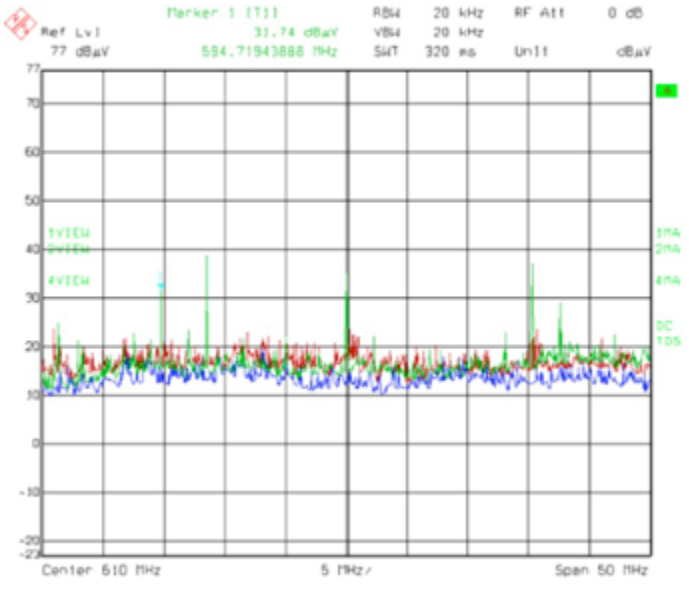


Fig. 7. ER correlated to test patterns; Blue: Pattern I, Green: Pattern VI, Red: Pattern III.

First, in order to understand the structure of printer data, which is directly obtained from the video signal sent to the laser scanner system with a probe, its representation in one dimensional (1D) space is investigated. Second, the row frequency used to convert 1D data to 2D image is calculated. The similar procedure in the second step is carried out for the AM-demodulated data, and the reconstructed image is shown on a computer screen as an image.

A. Representation of the 1D Printer Data

The time domain representation of a 1D printer data obtained from the video signal of the formatter output is given in Fig. 8(a). To be able to capture one image page, one has to collect about 1 second data. For this reason, the data signal is sampled with 10 MHz by an oscilloscope, which provides 10 Million Sample (MS) points. The data printed is a text with the font of Times New Roman and 72 pt. The corresponding frequency domain representation of the printed data is shown in Fig. 8 (b).

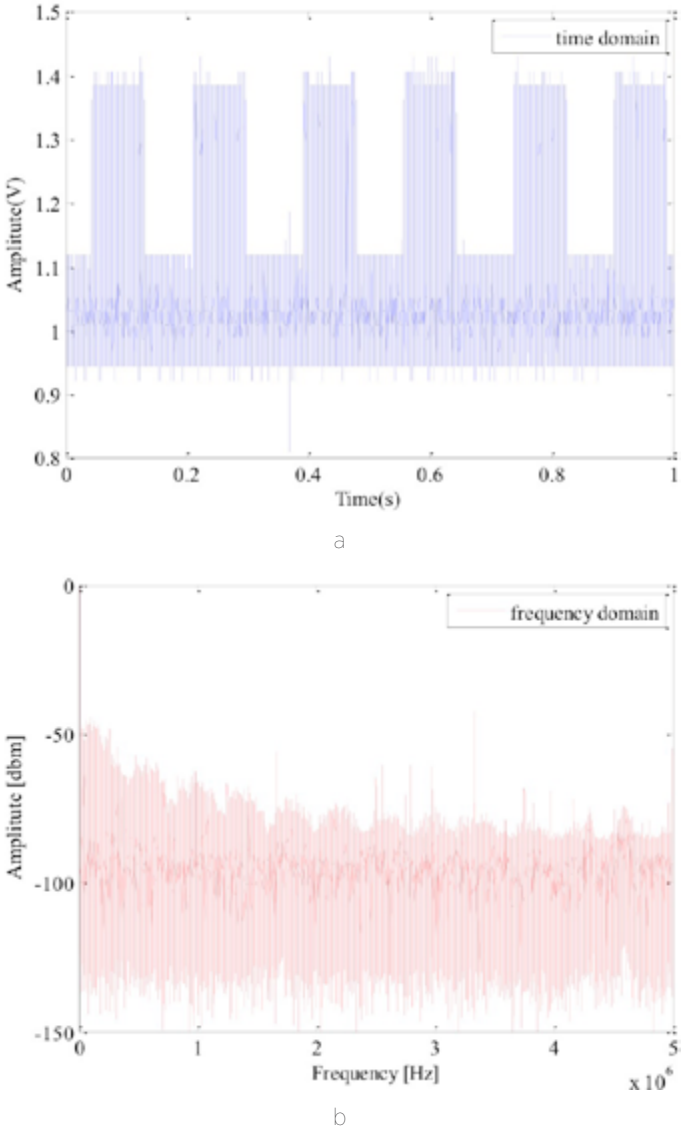


Fig. 8. Printer data singal representation. a) time domain b) frequency domain representations.

As seen from the frequency content, the most of the power accumulated in the first 1 MHz band. The reason is that the 72 pt. text fills almost only the 1 MHz spectrum. Thus, it can also be said that for this particular case, it is sufficient to use 2 MHz sampling frequency to reduce the sample points to 2MS. This 1D data has to be converted to 2D data in order to obtain the printed document image. Therefore, one needs a row frequency for the conversion, and this information is not known by an attacker in advance. However, a simple but effective computation method to find row frequency is introduced in the next section.

B. Extraction of the Row Frequency

To be able to reconstruct the printed data and raster it in 2D, one has to know the row frequency. The terminology of the row frequency is borrowed from video raster concept of the Video Display Units (VDU). The row frequency of a VDU, also known as horizontal frequency, can be obtained by the VESA standards if one knows or guests the resolution of VDU. Unfortunately, apart from the producer, one cannot know the row frequency of the printer since it is not declared in anywhere. However, this row frequency actually is hidden in the 1D data and can be obtained from the frequency spectrum by looking at the low frequencies. Mostly, the first powerful spike after the DC component provides the row frequency. In some cases, the first component can be weaker but by looking at other spike frequencies, it is seen that the other consecutive spike frequencies are the multiple of the main component. For the example given in Fig. 8, we obtain the row frequency as 2407 Hz for the given data as shown in Fig. 9. The other spike frequencies are 4814 Hz, 7221 Hz, and 9628 Hz, respectively.

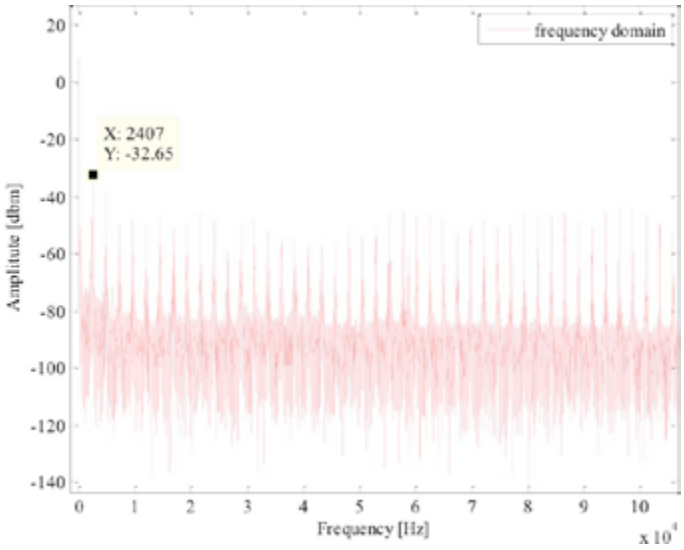


Fig. 9. Extracting the row frequency from the frequency spectrum.

C. Data Reconstruction from the Emissions

The similar procedure of finding the row frequency for the ideal data is applied also for the emissions obtained by the receiver systems. The AM-demodulated data is sampled by an oscilloscope and stored. Then this data is analyzed in the frequency domain to find the row frequency. Similarly, we focus on the first part of the spectrum and measure the row frequency. As shown in Fig. 10, the reconstructed image, the printed document, can be easily read. The row frequency is actually a real number, and the reason of obtaining a rotated image is that the row frequency is rounded to an integer value to be able convert 1D array to 2D image matrix.

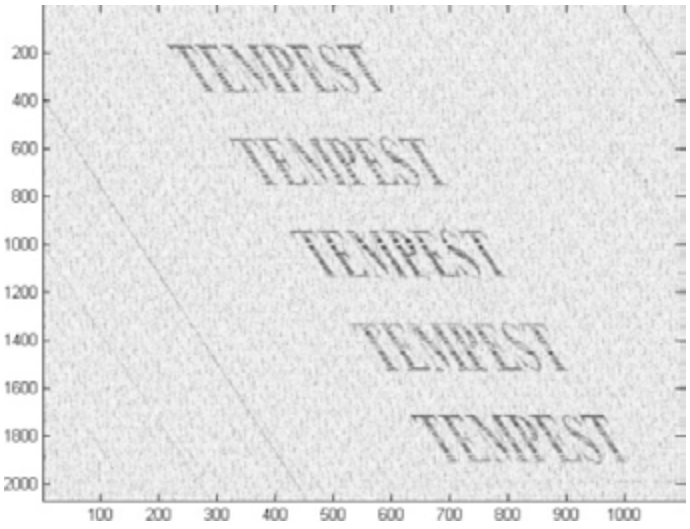


Fig. 10. Reconstructed image from the emission.

V. CONCLUSION

In this study, the CE of a laser printer are analyzed in different media. The emissions obtained from the electric radiation, power cable and signal line conductors like USB cable is investigated, and a measurement method to reveal the possible information leakages is proposed. Finally, the procedure of the image reconstruction of CE of the laser printer data is explained in detail. The experimental results show the vulnerability of the commercial laser printers in terms of emission security.

REFERENCES

[1] H. Highland and V. Fåk, "Electromagnetic radiation revisited, part II," Computers & Security, vol. 5, pp. 181-184, 1986.

[2] K. Beckman, "Leaking Computers – information on compromising emanations," presented at the National Council for Crime Prevention, Stockholm, Sweden, 1984.

[3] W. Van Eck, "Electromagnetic radiation from video display units: An eavesdropping risk?," Computers & Security, vol. 4, pp. 269-286, 1985.

[4] H. Fang, "Electromagnetic Information Leakage and its Protection of Computer," presented at the Science Press, Beijing, 1993.

[5] M. Vuagnoux and S. Pasini, "Compromising Electromagnetic Emanations of Wired and Wireless Keyboards," in USENIX Security Symposium, 2009, pp. 1-16.

[6] J. Zhang, "Information Recover Based on Compromising Electromagnetic Emanations of Keyboard," Beijing: Desertation of Beijing Jiaotong University, 2010.

[7] M. G. Kuhn, "Compromising emanations: eavesdropping risks of computer displays," University of Cambridge Computer Laboratory, Technical Report, UCAM-CL-TR-577, 2003.

[8] M. G. Kuhn, "Electromagnetic eavesdropping risks of flat-panel displays," in Privacy Enhancing Technologies, 2005, pp. 88-107.

[9] M. G. Kuhn, "Eavesdropping attacks on computer displays," presented at the Information Security Summit, 2006.

[10] R. Briol, "How to keep your data confidential," in Electromagnetic Security for Information Protection, 1991.

[11] M. Backes, M. Dürmuth, S. Gerling, M. Pinkal, and C. Sporleder, "Acoustic side-channel attacks on printers," presented at the Proceedings of the 19th USENIX conference on Security, Washington, DC, 2010.

[12] T. Tosaka, K. Taira, Y. Yamanaka, A. Nishikata, and M. Hattori, "Feasibility study for reconstruction of information from near field observations of the magnetic field of laser printer," in Electromagnetic Compatibility, 2006. EMC-Zurich 2006. 17th International Zurich Symposium on, 2006, pp. 630-633.

[13] R. Przesmycki, "Measurement and Analysis of Compromising Emanation for Laser Printer," in PIERS Proceedings, 2014.

SOSYAL MEDYA SİTELERİNİN KULLANDIKLARI ŞİFRE PAKETLERİNE GÖRE SINIFLANDIRILMASI

Mirsat Yeşiltepe, Muhammet Kurulay

Abstract — Today use rate of social media is a rapidly increasing. They were categorized due to the variety of tasks undertaken by these sites. Ranking among other sites of social media sites aim in this study and are classified according to their security cipher suite used by various parameters. Thus, the decision will be given to the security mechanisms at various levels should carry the newly created social media sites.

Index Terms — By rating , encryption, protocol, social media sites.

Özet — Günümüzde sosyal medya sitelerinin kullanım oranı artan bir hızla artmaktadır. Bu sitelerin üstlendikleri görevlerin çeşitliliği sebebiyle kategorize edilmişlerdir. Bu çalışmada amaç sosyal medya sitelerinin diğer siteler arasındaki sıralaması ve kendi kullandıkları güvenlik şifre paketlerine göre çeşitli parametrelerle sınıflandırılmaktadır. Böylelikle yeni oluşturulacak sosyal medya sitelerinin taşıması gereken güvenlik mekanizmalarına çeşitli düzeylerde karar verilmeye çalışılacaktır.

Anahtar Terimler — Protokol, reyting, sosyal medya siteleri, şifreleme.

I. GİRİŞ

Bu çalışmada çoğu web uygulamaları tarafından çeşitli nedenlerle kullanılan ve kullanımı artan sosyal medya sitelerinin[1] sahip olduğu güvenlik mekanizmaları incelenmiştir. Bu çalışmada günümüzde en çok kullanıcı sayısına sahip on beş site üzerinden inceleme yapılacaktır. Esas amaç sosyal medya sitelerinin isimleri üzerinden sitelerin karşılaştırılması olmadığından sitelerin isimlerine yer verilmemiştir. Sonraki bölümlerde sosyal medya sitelerinin tüm siteler içindeki sıralaması incelenmiş sonra güvenlik mekanizmaları özet ve özellikleri biçiminde incelenecektir. Sonuç bölümünde ise yeni oluşturulmak istenen sitelerin taşıması gereken güvenlik mekanizmaları parametrelerine karar verilecektir. Tüm test edilen ortamın test tarihi 15 Haziran 2015'tir.

II. SOSYAL MEDYA SİTELERİNİN REYTİNGLERİ

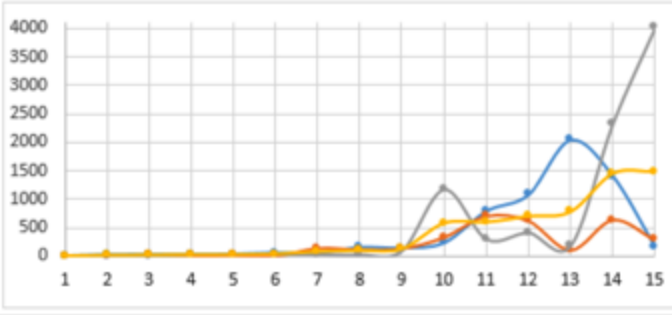
Tablo I'de sosyal medya sitelerinin tüm siteler arasındaki sıralaması üç adet site sıralayıcısı tarafından alınan bilgilere göre elde edilmiştir. Siteler sıralanırken kullanıcı sayısı dikkate alınılmıştır[2].

Site Sırası	Kullanıcı Sayıları	Sıralayıcı 1	Sıralayıcı 2	Sıralayıcı 3
1.	900.000.000	3	3	2
2.	310.000.000	21	8	8
3.	255.000.000	25	19	9
4.	250.000.000	27	13	26
5.	120.000.000	32	28	
6.	110.000.000	55	13	34
7.	100.000.000	49	145	36
8.	80.000.000	150	120	21
9.	65.000.000	138	139	91
10.	42.000.000	231	335	1172
11.	40.000.000	791	701	296
12.	38.000.000	1.082	615	408
13.	37.000.000	2.046	113	179
14.	15.500.000	1407	635	2328
15.	15.000.000	153	285	4022

Tablo I - Sosyal medya sitelerinin tüm siteler arasındaki sıralaması ve kullanıcı sayıları

Tablo I incelendiğinde sosyal medya sitelerinin kullanıcı sayıları ile tüm siteler arasında bir ilişki olduğu fakat bunu kuvvetli bir biçimde olmadığı gözlemlenmiştir. Kullanıcı sayısı azaldıkça sitelerin sıralaması kurallılıktan uzaklaşmaya başlamıştır. Buradan sitelerin sıralaması ile siteyi kullanan kullanıcı sayısı ilk başta sıralamasını belirleyen önemli bir faktör olarak görülürken, sıralama düştükçe bu bağ zayıflamış ve bazende ilişki ortadan kalmıştır. Bundan sonraki bölümlerde bu sıralamada güvenlik mekanizmalarının etkisinin olup olmadığı tartışılacaktır.

Sıralayıcı 1, sıralayıcı 2 ve sıralayıcı 3'ün sosyal medya sitelerini tüm siteler içinde sıralarken hangi parametreleri kullandıklarından çok sitelerin kullandıkları şifreleme paketlerinin güvenlik düzeylerinin hangi sıralayanın daha önem vermiş olabileceği fikri üzerinde durulmuştur.

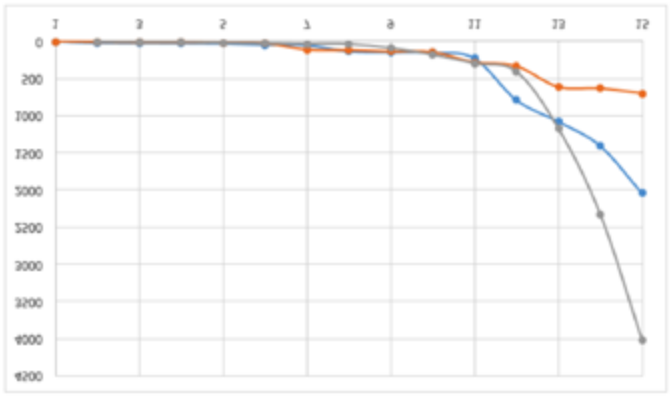


Tablo II - Sosyal medya sitelerinin tüm siteler arasındaki sitelerin sırası

Tablo II'de yatay eksen sitelerin sırasını düşey eksen ilgili sitenin sıralayıcılar tarafından belirlenen sırasını göstermiştir. Mavi renk 1. sıralayıcıyı, turuncu renk ikinci sıralayıcı, gri renk 3. Sıralayıcıyı, sarı renk ise sıralayıcıların ortalamasını göstermiştir.

Tablo II incelendiğinde site sıralayıcıları arasında birinci ve üçüncüsü birbirine uyumlu iken ikinci sıralayıcı diğerlerinden farklı bir eğilim göstermiştir. Fakat sıralayıcıların ortalaması

alındığında ise sitelerin sıraları ile kullanıcı sayıları uyumlu olduğu gözlemlenmiştir.



Tablo III - Sosyal medya sitelerinin tüm siteler arasındaki sıralanışından bağımsız olarak sıralanması

Tablo III'de yatay eksen sitelerin sırasını, düşey eksen ilgili sitenin sıralayıcılar tarafından belirlenen sıralaması gösterilmiştir. Burada amaç site sıralayıcılarının kendi içlerindeki uyumunun gösterilmek istenmesidir. Sitelerin reytingleri artan olarak (sitelerin sıralanışından bağımsız olarak) gösterilmiştir. 1. ve 2. site sıralayıcısının 3.'süne göre daha uyumlu oldukları gözlemlenmiştir.

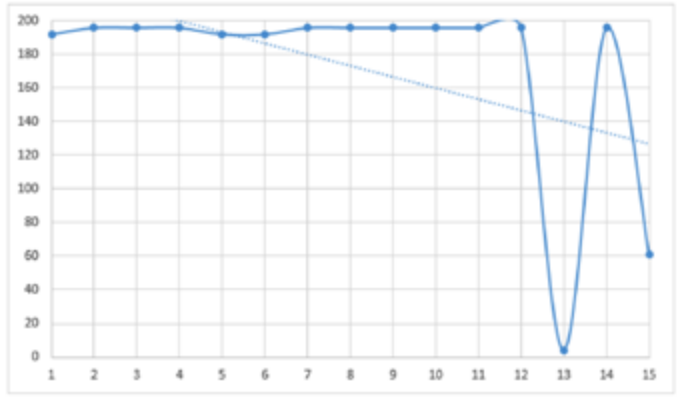
III. SOSYAL MEDYA SİTELERİNİN GÜVENLİK MEKANİZMALARI ÖZETİ

Şifre paketi kavramı kimlik doğrulama, şifreleme, mesaj kimlik doğrulama kodu (MAC), anahtar değişim algoritmalarının Transfer Seviye Güvenlik (TLS), Güvenlik Soket Katmanı (SSL) protokollerinden biri kullanılarak çeşitli güvenlik belirtimlerinin bir bütün olarak açıklamasıdır. Bu sebeple bu kavramı bir bütün olarak düşünülmesi gerekir. Şifre paketleri seviyelerine göre sıralanmıştır[3][6].

PAKET TANIMLAYICISI	ŞİFRE PAKETİ İSMİ
0x00C02B	TLS_ECDHE_ECDSA_WI TH_AES_128_GCM_SHA256
0x00C02B	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x00C02F	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x00C02F	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x00C02B	TLS_ECDHE_ECDSA_WITH_AES_128_GCM_SHA256
0x00C02B	TLS_ECDHE_ECDSA_WITH_AES_128_GCM_SHA256
0x00C02F	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x00C02F	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x00C02F	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x00C02F	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x00C02F	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x00C02F	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x000004	SSL_RSA_WITH_RC4_128_MD5
0x00C02F	TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256
0x00003C	TLS_RSA_WITH_AES_128_CBC_SHA256

Tablo IV - Sosyal medya sitelerinin şifre paketleri

Tablo IV'de test edilen sosyal medya sitelerinin kullandıkları güvenlik mekanizmalarının kodları ve isimleri verilmiştir[4].



Tablo V - Sosyal medya sitelerinin şifre paketleri güvenlik düzeyleri

Tablo V'de yatay eksen sitelerin sırasını düşey eksen ilgili sitenin şifre paketinin güvenlik düzeyini gösterilmiştir. Güvenlik düzeyi ile anlatılmak istenen TLS ve SSL protokolünün ilgili şifreleme paketi türünü kaçıncı sırada belirttiğidir. Şifreleme paketleri türleri ilgili protokollerde güvenlik açısında artan düzende sıralanmıştır. Yani şifreleme paketi türü sonra belirtilen (paket tanımlayıcısı yüksek olan) şifreleme paketi bir bütün olarak daha güvenlidir. Sıralama için paket tanımlayıcısının belirttiği binary kod yerine paket sıralayıcı en az olana bir değeri verilerek sıralanmıştır. Test edilen sosyal medya sitelerinin kullandıkları şifre paketlerinin türleri seviyelerine göre listelenmiştir. Paket tanımlayıcıları bir bütün olarak sıralanmıştır. Yani paketin tanımlayıcısının değerinin büyük olması şifreleme paketinin daha güçlü olduğunu gösterir. Tablo incelendiğinde 13. ve 15. site haricinde diğer sitelerin kullandıkları şifre paketlerinin güvenlik düzeylerinin yakın olduğu gözükmüştür. Genel olarak bakıldığında ise eğim çizgisinin aşağı yönlü olduğundan sitelerin kullanıcı sayıları ile şifre paketlerinin genel olarak uyumlu olduğu sonucu çıkarılabilir. 13. sosyal medya sitesinin özel durumu için birinci ve üçüncü sıralayıcının yukardaki duruma uygun sonuçlar içerdiği ikinci sıralayıcının ise bu durumu göz ardı ettiği görülmüştür. Son olarak çıkarılacak durum site reytinginin şifre paketleri ile ilişkisinin olduğudur.

IV. SOSYAL MEDYA SİTELERİNİN GÜVENLİK MEKANİZMALARININ ÖZELLİKLERİ

Bu bölümde isimleri veriler şifre paketlerinin özelliklerine değinilip sitelerin kullanıcı sayılarıyla ilişkili olup olmadığı incelenecektir.

Tablo VI ve tablo VII'da şifreleme paketlerinin içerdiği bilgiler iki guruba ayrılmıştır. Burada amaç şifreleme paketlerinin parçalarının sıralamaya olan etkisini incelemektir. Örneğin 13. Site anahtar değişim algoritması olarak ECDHE, kimlik doğrulama olarak RSA kullanmış, 14. Site ise her iki güvenlik belirtecinde RSA kullanmıştır. Eğer ki sadece şifreleme paketi düzeyinden (genel olarak) incelendiğinde 13. Sitenin güvenlik düzeyi düşük çıkacaktır. Fakat şifreleme paketi bu iki güvenlik düzeyi belirtecine göre incelenirse 13. Site daha yüksek çıkacaktır. Burada amaç güvenlik düzeyi düşük çıkanların kendilerinden düzeltmesi gereken yönleri belirlemeye çalışmaktır. Güvenlik düzeyinin artmasında şifreleme algoritmaların anahtar uzunlukları önemli olmakla birlikte tek başına yeterli değildir. Güvenlik bir bütündür.

	PROTOKOL	ANAHTAR DEĞİŞİM ALGORİTMASI	KİMLİK DOĞRULAMA ALGORİTMASI
1.	TLS	ECDHE	ECDSA
2.	TLS	ECDHE	RSA
3.	TLS	ECDHE	RSA
4.	TLS	ECDHE	RSA
5.	TLS	ECDHE	ECDSA
6.	TLS	ECDHE	RSA
7.	TLS	ECDHE	RSA
8.	TLS	ECDHE	RSA
9.	TLS	ECDHE	RSA
10	TLS	ECDHE	RSA
11.	TLS	ECDHE	RSA
12.	SSL	RSA	RSA
13.	TLS	ECDHE	RSA
14	TLS	RSA	RSA
15.	SSL	RSA	RSA

Tablo VI - Sosyal medya sitelerinin şifre paketlerinin özellikleri

Tablo VI'da test edilen sosyal medya sitelerinin protokol olarak çoğunlukla TLS kullandıkları gözlemlenmiştir. SSL'in iletişimde sertifika uyarı mesajı kullanmaması, sertifika doğrulama mesajı oluşturulabilmesinin mümkün fakat zor bir süreç olması gibi nedenlerden dolayı günümüzde TLS'in kullanımının artması[5] ve bu durumun test edilen sitelerde görülmesi normaldir. Test edilen 12. sosyal medya sitesinin SSL kullanma durumu birinci ve üçüncü site sıralayıcısında önem arz etmekte iken, 15. sosyal medya sitesinin SSL kullanması sadece üçüncü site sıralayıcısında önem arz etmiştir. Genel olarak üçüncü site sıralayıcısı site sıralarının belirlerken kullanılan protokole daha çok ilgilendiği sonucu çıkarılabilir.

Anahtar değişimi algoritmalarının kullanımlarındaki esas amaç iki kullanıcının bir anahtarı güvenli bir şekilde birbirlerine iletmeleri ve daha sonrasında da bu anahtar yardımı ile şifreli mesajları birbirlerine gönderebilmelerini sağlamaktır. Test edilen sosyal medya sitelerinin anahtar değişim algoritması olarak çoğunlukla ECDHE kullandıkları, 12. 14. ve 15. Sosyal medya sitelerinin ise RSA kullandıkları gözlemlenmiştir. Test edilen çoğu sitenin ECDHE kullanmasının nedeni bu şifreleme türünün kullanılan şifreleme bit sayısı attığında RSA'ya göre hızının artmasıdır. Diğer bir neden sitelerin yeni teknolojilerle uyumlu olma istediğidir[6]. Bulut ile iletişimde hızın önemi ortada olması ECDHE'nin bir başka tercih nedenidir. RSA kullanan sitelerin bu durumlarının sıralayıcılarında etken olarak en iyi gören üçüncü sıralayıcıdır.

Sunucu kendisindeki bilgi veya siteye erişimi sağlayanın tam olarak kim olduğunu bilmesi gerektiğinde kimlik doğrulama kullanılır. Kimlik doğrulamasında, kullanıcı veya bilgisayar / sunucu veya istemci kimliğini karşı tarafa kanıtlamak zorundadır. Genellikle, bir sunucu tarafından kimlik doğrulama, kullanıcı adı ve parola kullanımını gerektirir. Test edilen sosyal medya sitelerinin kimlik doğrulaması olarak genellikle RSA kullandıkları, 1. ve 5. sosyal medya sitesinin ise ECDSA kullandığı gözlemlenmiştir. Fakat bu durum site sıralayıcıları için özel bir durum oluşturmamıştır. Anahtar değişimde RSA az kullanılırken kimlik doğrulamada RSA'nın daha yaygın kullanılmasının en önemli nedeni kimlik doğrulama mekanizmasının oluşturulmasındaki zorluktur[9].

	SİMETRİK ŞİFRELEME ALGORİTMASI	SİMETRİK ŞİFRELEME ANAHTAR UZUNLUĞU	HASH ALGORİTMASI
1.	AES_128_GCM	128	SHA256
2.	AES_128_GCM	128	SHA256
3.	AES_128_GCM	128	SHA256
4.	AES_128_GCM	128	SHA256
5.	AES_128_GCM	128	SHA256
6.	AES_128_GCM	128	SHA256
7.	AES_128_GCM	128	SHA256
8.	AES_128_GCM	128	SHA256
9.	AES_128_GCM	128	SHA256
10	AES_128_GCM	128	SHA256
11.	AES_128_GCM	128	SHA256
12.	AES_128_GCM	128	SHA256
13.	AES_128_GCM	128	SHA256
14	AES_128_GCM	128	SHA256
15.	AES_128_GCM	128	SHA256

Tablo VII - Sosyal medya sitelerinin şifre paketlerinin özellikleri (devam)

Test edilen sitelerin simetrik şifreleme algoritması, simetrik şifreleme anahtar uzunluğu ve hash algoritması parametrelerinin aynı olduğu gözlemlenmiştir.

Hash fonksiyonu, değişken uzunluklu veri kümelerini, sabit uzunluklu veri kümelerine haritalayan algoritma veya alt programdır. Genelde SHA ve MD5 türleri kullanılır. SHA'nın kullanımındaki amaç tek seferde daha çok biti özümseyebilmesidir (hash) [10].

V. SONUÇ

Sosyal medya sitelerinin reytinglerinin belirlenmesinde kullanıcı sayılarının ve şifre paketlerinin ilişkisi vardır. Kullanıcı sayısının fazla olduğu siteler daha güçlü şifreleme paketi tercih etmişlerdir. Fakat sitelerin genel olarak belirli bir seviyedeki şifre paketleriyle çalıştığı gözlemlenmiştir. Elbette ki siteler sıralanırken ortamdaki birçok parametre kullanılmıştır. Fakat şifre paketlerindeki farklılıklardan sitelerin sıralamaması etkilenmiştir. Sitelerin verimli olabilmesi için kendilerine en uygun şifre paketlerinin belirlenip kullanılması gerekir.

Test edilen sosyal medya sitleri genellikle protokol olarak TLS, anahtar değişim algoritması olarak ECDHE, kimlik doğrulama algoritması olarak RSA, simetrik şifreleme algoritması olarak AES_128_GCM, simetrik şifreleme anahtar uzunluğu olarak 128 değerini ve hash algoritması olarak SHA256 kullanılmıştır.

Günümüzde oluşturulmak istenen sosyal medya sitelerinin taşıması gereken en düşük seviyeli şifre paketi TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256 iken yapılabilinecek saldırılardan kurtulmak için şu anki sosyal medya

sitelerinden daha çok güvenlik mekanizması taşınılması gerektiği düşünülerek ya hash algoritması olarak SHA256 yerine SHA384 kullanarak TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384 veya anahtar değişim algoritması olarak ECDH kullanarak TLS_ECDH_RSA_WITH_AES_128_GCM_SHA256 şifre paketi kullanılmaları tavsiye edilir. Eğer uzun dönemde kurulacak sitelerin güvenlik mekanizmaları ile mekanizmaların maliyeleri düşünüldüğünde yine bu şifre paketlerinin şu an için kullanılması tavsiye edilebilir. Ama uzun dönemde SSL2 daha uzun dönemde PCT protokollerinin kullanımları düşünülebilir. Bu protokollere uyumlu güvenlik ortamlarına uygun ortamlar yedekte hazır bulundurulabilir. Çünkü günümüzde yapı bilinecek saldırılara karşı sitelerin mümkün olan en kısa sürede yeni ortama adapte olması gerekir. Olmaması durumunda yeni ortam bu site için kalitenin düştüğü bir ortam olacaktır.

KAYNAKLAR

[1] Boyd, Danah. "Why youth (heart) social network sites: The role of networked publics in teenage social life." MacArthur foundation series on digital learning–Youth, identity, and digital media volume (2007): 119-142.

[2] Huyensau, "Top 5 Trendiest Social Networking Sites in 2015", http://www.toplisttips.com/top-5-trendiest-social-networking-sites-update-january-2015/ , 14 Haziran 2015.

[3] Ristic, Ivan. "SSL/TLS Deployment Best Practices." URL https://www. ssllabs. com/downloads/SSL_TLS_Deployment_Best_ Practices_1. 0. pdf (2013).

[4] The Sprawl, "Researchtls and SSL Cipher Suites - Known cipher suites", https://www.thesprawl.org/research/tls-and-ssl-cipher-suites/ , 16 Haziaran 2015.

[5] Rescorla, Eric. SSL and TLS: designing and building secure systems. Vol. 1. Reading: Addison-Wesley, 2001.

[6] Mishra, Vivek. "Cassandra Data Security." Beginning Apache Cassandra Development. Apress, 2014. 61-78.

[7] Nagaraju, Mr S., and Mr B. Latha Parthiban. "An Enhanced Symmetric Role-Based Access Control Using Fingerprint Biometrics for Cloud Governace.".

[8] Çeviri:Mesut Timur -OWASP GUIDE 2.0.1, 2007.

[9] Fu, David E., and Jerome A. Solinas. "IKE and IKEv2 authentication using the elliptic curve digital signature algorithm (ECDSA)." (2007).

[10] Krawczyk, Hugo, Ran Canetti, and Mihir Bellare. "HMAC: Keyed-hashing for message authentication." (1997).

ENERJİ SEKTÖRÜNDE BİLGİ GÜVENLİĞİNİN YÖNETİLMESİ: MEVZUAT VE STANDARTLAR

Fikret Ottekin[1] – Orhan Çalık[2]
[1] Cterra A.Ş. Genel Müdür Danışmanı fikret.ottekin@icterra.com
[2] TÜBİTAK BİLGEM Siber Güvenlik Enstitüsü, Uzman Araştırmacı orhan.calik@tubitak.gov.tr

Özetçe — Kritik altyapıları barındıran en önemli sektörlerden biri enerji sektörüdür. Enerji altyapılarının kurumsal bilişim sistemlerine ilave olarak endüstriyel kontrol sistemleri dolayısı ile de bilgi sistemlerine bağımlılığı tartışmasız durumdadır. Bu bilgi sistemlerine kurum içinden veya dışından çeşitli saldırıların yapıldığı, saldırılar sonucunda çeşitli ölçekte zararların oluştuğu bilinmektedir.

Aralık 2014 tarihinde EPDK bilgi güvenliğine dönük gereksinimleri göz önünde bulundurarak konu ile ilgili üç yönetmeliği güncellemiş ve lisans sahiplerine bilişim sistemleri güvenliğini sağlama doğrultusunda yükümlülükler getirmiştir. Ancak enerji sektöründe yer alan kuruluşların önemli bölümünde konu ile ilgili farkındalık eksikliği bulunmaktadır. Makalenin amacı bu eksiği bir nebze olsun gidermek, bilgi güvenliği kapsamında enerji sektöründe faaliyet gösteren kuruluşların uygulaması gereken önlemleri ortaya koymaktır.

Anahtar Kelimeler — Kritik altyapılar, bilgi güvenliği, enerji sektörü,risk yönetimi,

Abstract — Energy sector is by far one of the most important sectors that utilize critical infrastructures. Energy infrastructures' dependance on information systems is indisputable due to the presence of Industrial Control Systems as well as corporate information systems. It is common knowledge that these information systems are targeted by various attacks from inside and outside the organizations that cause consequences of various degrees.

In December 2014, Energy Market Regulatory Authority has revised three directives and brought the obligation of assuring the security of their information systems to license holders. However, awareness regarding information security is insufficient in most of the institutions operating in the energy sector. The objective of this article is improving the information security awareness and defining the controls that should be applied by the institutions operating in the energy sector.

Index Terms — Critical infrastructure, information security, energy sector, risk management.

I. GİRİŞ

Bu makalenin konusu, EPDK'nın kritik enerji altyapılarında önemi giderek artan bilgi güvenliği ihtiyacına dayanarak kuruluşlara uyum zorunluluğu getirdiği bilgi güvenliği standartlarıdır. Enerji sektöründe bilgi güvenliği konusunda yeterli miktarda uzman bulunmadığı göz önünde bulundurularak bilgi güvenliği standartlarına ve standartlarda gözden kaçabilecek hususlara dikkat çekilmekte, standartların birleştiği ve ayrıldığı noktalar tablo ve şekillerle vurgulanarak bir bakışta anlaşılacak şekilde gözler önüne serilmektedir. Böylece bilgi güvenliğine dönük önlemler kurumlarda uygulanırken konu ile ilgili standartların ve standartların önemli bölümlerinin gözden kaçırılmadan eksiksiz şekilde uygulanması ve ulusal kritik altyapı güvenliğinin en kritik bileşenlerinden kritik enerji altyapıları güvenliğine katkı yapılması hedeflenmektedir.

II. DÜNYA'DA KRİTİK ENERJİ ALTYAPILARI VE GÜVENLİK

20 Haziran 2013'te yayınlanarak yürürlüğe giren "Ulusal Siber Güvenlik Stratejisi ve 2013-2014 Eylem Planı" belgesinde kritik altyapı,

"İşlediği bilginin gizliliği, bütünlüğü veya erişilebilirliği bozulduğunda,

- Can kaybına,
- Büyük ölçekli ekonomik zarara,
- Ulusal güvenlik açıklarına veya kamu düzeninin bozulmasına yol açabilecek bilişim sistemlerini barındıran altyapıları ifade eder"

şeklinde tanımlanmıştır [1].

Bu tanım göz önünde bulundurulduğunda, kritik altyapıları barındıran ilk sektörlerden biri olarak akla enerji sektörü gelmektedir.

AB ve ABD'de kritik altyapı güvenliği konusunda yapılan çalışmalar da enerji sektörünün önemine işaret etmektedir. 23 Aralık 2008'de AB Resmi Gazetesi'nde yayınlanan "Avrupa Kritik Altyapılarının Belirlenmesi ve Güvenliklerinin İyileştirilme Gereksiniminin Değerlendirilmesine İlişkin 2008/114/AB Konsey Yönergesi" dokümanının 5. maddesinde yönergenin enerji ve ulaştırma sektörlerine odaklandığı, ilave olarak bilgi ve iletişim sektörlerinin de gözden geçirilebileceği belirtilmektedir [2].

ABD Anayurt Güvenliği Bakanlığının, resmi web sitesinde ilan ettiği "Kritik Altyapı Sektörleri" arasında enerji sektörü de yer almaktadır [3].

Enerji altyapılarının, kurumsal bilişim sistemlerine ilave olarak endüstriyel kontrol sistemleri dolayısı ile de bilgi sistemlerine bağımlılığı tartışmasız durumdadır. Bu bilgi sistemlerine kurum içinden veya dışından çeşitli saldırıların yapıldığı, saldırılar sonucunda çeşitli ölçekte zararların oluştuğu bilinmektedir. ABD'de faaliyet göstermekte olan EKS-BOME (Endüstriyel Kontrol Sistemleri - Bilgisayar Olaylarına Müdahale Ekibi) 2013 yılı içinde kendilerine işletmeci kurumlar tarafından 257 olay bildirildiğini, saldırıya uğrayan kurumların %57'sinin enerji sektöründe bulunduğunu bildirmiştir [4].

Tüm bu bilgiler, hem kurumların ve tüketicilerinin ekonomik zarardan korunması, hem de vatandaşın can güvenliğinin ve kamu düzeninin muhafaza edilmesi için kritik enerji altyapılarında bilişim sistemlerinin güvenliğinin sağlanmasına yönelik yatırım yapılması gerekliliğini gözler önüne sermiştir

III. EPDK'NIN GETİRDİĞİ YÜKÜMLÜLÜKLER

4628 sayılı Enerji Piyasası Düzenleme Kurumunun Teşkilat ve Görevleri Hakkında Kanun, Enerji Piyasası Düzenleme Kurulu'na,

- "Tüketicilere güvenilir, kaliteli, kesintisiz ve düşük maliyetli elektrik enerjisi hizmeti verilmesini teminen gerekli düzenlemeleri yapmak" (Madde 5, c bendi).
- "Üretim, iletim ve dağıtım şirketleri ile otoprodüktör ve otoprodüktör grubu tesisleri için güvenlik standartları ve şartlarını tespit etmek ve bunların uygulanmasını sağlamak" (Madde 5, e bendi).

görevlerini vermiştir [5].

EPDK, Aralık 2014 tarihinde bilgi güvenliğine dönük gereksinimleri göz önünde bulundurarak aşağıda belirtilen üç yönetmeliği güncellemiş ve lisans sahiplerine bilişim sistemleri güvenliğini sağlama doğrultusunda yükümlülükler vermiştir. Bu kapsamda,

1. Elektrik Piyasası Lisans Yönetmeliğinde Değişiklik Yapılmasına Dair Yönetmelik
2. Doğal Gaz Piyasası Lisans Yönetmeliğinde Değişiklik Yapılmasına Dair Yönetmelik ve
3. Petrol Piyasası Lisans Yönetmeliğinde Değişiklik Yapılmasına Dair Yönetmelik,

26 Aralık 2014 tarihinde Resmi Gazete'de yayınlanarak yürürlüğe girmiştir. [6, 7, 8] Yönetmeliklerle bilgi güvenliği kapsamında yükümlülüğe tabi olan kurumlar şunlardır:

1. Elektrik piyasasında,
   a. OSB (Organize Sanayi Bölgesi) üretim lisansı sahipleri hariç olmak üzere, kurulu gücü 100MW ve üzerinde olan ve geçici kabulü yapılmış bütün üretim tesisleri,
   b. İletim lisansı sahibi,
   c. Piyasa işletim lisansı sahibi,
   d. OSB dağıtım lisansı sahipleri hariç olmak üzere dağıtım lisansı sahipleri (elektrik dağıtım şirketleri) [6]
2. Doğal gaz piyasasında,
   a. İletim lisansı sahibi şirketler,
   b. Sevkiyat kontrol merkezi kurmakla yükümlü dağıtım lisansı sahibi şirketler [7]
3. 3. Petrol piyasasında,
   a. Rafinerici lisansı sahipleri [8].

Tüm bu kurumlara, aşağıdaki yükümlülük getirilmiştir:

"Kurumsal bilişim sistemi ile endüstriyel kontrol sistemlerini TS ISO/IEC 27001 Bilgi Güvenliği Yönetim Sistemi standardına uygun bir şekilde işletmek, TS ISO/IEC 27001 standardına uygun faaliyet gösterdiğini Türk Akreditasyon Kurumuna akredite olmuş bir belgelendirme kurumuna ispat ederek sistemlerini belgelendirmek ve söz konusu belgelerin geçerliliğini sağlamak,"

Yükümlülükler 1/3/2016 tarihinde yürürlüğe girecektir.

IV. KURUMLARIN UYUM SAĞLAMASI GEREKEN BİLGİ GÜVENLİĞİ STANDARTLARI

Bilgi Güvenliği konusunda yönetmelikte belirtilen TS ISO/IEC 27001 Bilgi Güvenliği Yönetim Sistemi Gereksinimleri standardının son sürümü 2013'te yayınlanmış olup, uzun süreden beri yürürlükte olan 2005 sürümü ile arasında dikkate değer farklar vardır [9].

ISO 27001 standardı, bilgi güvenliğini sağlamaya çalışan kurumun hangi sektörde yer aldığı ile ilgilenmez. Yıl boyunca yapılması gereken risk analizi, eğitim, iç tetkik ve iyileştirme gibi başlıca faaliyetleri ve yönetimin sorumluluklarını, yani ana hatları ile yönetim sürecini tanımlar [10]. Somut güvenlik önlemleri ise ISO 27002 standardında yer almaktadır [11]. 27002 standardının 27001:2005'te "vazgeçilmez" olduğu belirtilmektedir. 27001:2013'te bu ifade bulunmasa da, 6.1.3 Bilgi Güvenliği Risk Tedavisi başlığı altında 27002:2013'ün dikkatle gözden geçirilmesi, uygulanmayan güvenlik önlemlerinin neden uygulanmadığının kaydedilmesi gerektiği belirtilmektedir.

Bu makalenin özelinde incelenecek olan ISO/IEC 27019 ve ISO/IEC 27011 standartları da, sırasıyla bilgi güvenliğinin enerji ve iletişim sektörlerinde sağlanmasına yönelik olarak yayınlanmış standartlardır [12, 13]. Bu makalenin yayınlanma tarihi itibarı ile standartların son sürümleri 27019:2013 ve 27011:2008 olup, bu standartlar 27002:2005 standardına yapılan ilavelerle oluşturulmuştur.

27002:2005 standardında bulunan önlemlere iki tip ilave yapılmıştır.

a. 27002'de hiç bulunmayan, "Sektöre Özel Önlemler" tanımlanmıştır.
b. 27002'de zaten mevcut olan bazı önlemlerde "Sektöre Özel Uygulama Kılavuzu" başlığı altında birkaç paragraf eklenerek ek tavsiyeler yapılmıştır.

Enerji sektöründe bulunan ve bilgi güvenliğini tesis etmeye çalışan kurumların, 27011 ve 27019 standartlarına uyum yükümlülüğü bulunmamakla birlikte bu ilavelerden fikir alabilecekleri ve kurumsal bilgi güvenliğine katkı sağlayabilecekleri söylenebilir.

Bu aşamada enerji sektöründe bilgi güvenliğinin sağlanması ile iletişim sektörüne özel bilgi güvenliği standardının ne ilgisi olduğu sorusu akla gelebilir. Özellikle enerji iletim ve dağıtımı ile uğraşan kuruluşlar, geniş alanlara yayılmış durumda bulunan sistemlerin kontrolünü sağlarken çeşitli iletişim sistemlerine bağımlı duruma gelmektedir. Bu nedenle, 27019 standardının birinci bölümünde, "Süreç kontrol sistemlerini destekleyen iletişim sistemlerinde ve bileşenlerinde 27011 standardında yer alan önlemlerin uygulanması" tavsiye edilmektedir.

Yönetmeliklerde yer alan "Kurumsal Bilişim Sistemi ve Endüstriyel Kontrol Sistemleri" kapsamı bilgi güvenliği standartları ile birlikte değerlendirildiğinde, 27001'in kurumun genel bilgi güvenliği yönetim sürecini tanımladığı, Kurumsal Bilişim Sistemlerinde ve insan kaynağı da dahil olmak üzere diğer varlıklarda alınacak önlemlerle ilgili olarak 27002, Endüstriyel Kontrol Sistemlerinde alınacak önlemlerle ilgili olarak ise 27002, 27011 ve 27019 standartlarının göz önünde bulundurulması gerektiği söylenebilir. Bu durum aşağıdaki şekille de ifade edilebilir.

Şekil-1. Kurumsal Bilgi Sistemleri ve ilgili standartlar

V. ENERJİ VE İLETİŞİM SEKTÖRÜNE ÖZEL STANDARTLARIN GETİRDİKLERİ

Aşağıdaki tabloda, 27011 ve 27019 standartlarında yer alan ilave bilginin 27002 ile eşleştirilerek gözler önüne serilmesi amaçlanmıştır. Tabloda yer alan kontroller, yani güvenlik önlemleri, üç kategoride toplanabilir:

1. Tablonun 27011 ve 27019 sütunlarında '►' ve '◄' sembollerinin bulunduğu satırlar 27002'de tanımlanan önlemin 27011 ve/veya 27019'da da aynen mevcut olduğu, herhangi bir ek yapılmadığı anlamına gelmektedir.
2. Bazı önlemlerde, 27002'deki tanıma ilave olarak 27011 ve/veya 27019'da "Sektöre özel uygulama kılavuzları" bulunmaktadır.
3. 27002'de yer almayan, iletişim sektörüne veya enerji sektörüne özel güvenlik önlemleri de mevcuttur. 27011 veya 27019 standardında bu önlemlerin nasıl uygulanacağı açıklanmaktadır. Kırmızı yazı karakteri kullanılarak belirtilen önlemler 27019 (Enerji), Lacivert yazı karakteri kullanılarak belirtilen önlemler ise 27011 (İletişim) sektörüne özel olarak tanımlanmış ve standartlara eklenmiş olan kontrollerdir. Bu önlemler Standartların EK-A bölümlerinde yer almaktadır.

27011:2008 (İletişim)	27002:2005	27019:2013 (Enerji)
	5.1 Bilgi Güvenlik Politikası	
►	5.1.1, 5.1.2	◄
	6.1 Kurum İçi Organizasyonu	
►	6.1.1, 6.1.2, 6.1.3, 6.1.4	◄
Sektöre özel uygulama kılavuzu vardır.	6.1.5 Gizlilik anlaşmaları	◄
Sektöre özel uygulama kılavuzu vardır.	6.1.6 Otoritelerle iletişim	Sektöre özel uygulama kılavuzu vardır.
►	6.1.7 Uzmanlık grupları ile iletişim	Sektöre özel uygulama kılavuzu vardır.
►	6.1.8	◄
►	6.2 Üçüncü Taraf Erişiminin Güvenliği	
►	6.2.1 Üçüncü taraf erişiminde risklerin tanımlanması	Sektöre özel uygulama kılavuzu vardır.
Sektöre özel uygulama kılavuzu vardır.	6.2.2. Müşterilerle çalışırken güvenlik	Sektöre özel uygulama kılavuzu vardır.
Sektöre özel uygulama kılavuzu vardır.	6.2.3 Üçüncü taraf sözleşmelerinde güvenlik gerekleri	Sektöre özel uygulama kılavuzu vardır.
	7.1 Varlıklarla ilgili sorumluluklar	
Sektöre özel uygulama kılavuzu vardır.	7.1.1 Varlık Envanteri	Sektöre özel uygulama kılavuzu vardır.
►	7.1.2 Varlıkların sahipleri	Sektöre özel uygulama kılavuzu vardır.
►	7.1.3	◄
	7.2 Bilgi Sınıflandırması	
Sektöre özel uygulama kılavuzu vardır.	7.2.1 Sınıflandırma rehberleri	Sektöre özel uygulama kılavuzu vardır.
►	7.2.2	◄
	8.1 İşe almadan önce	
Sektöre özel uygulama kılavuzu vardır.	8.1.1 Roller ve sorumluluklar	Sektöre özel uygulama kılavuzu vardır.
Sektöre özel uygulama kılavuzu vardır.	8.1.2 Personel gözetleme	Sektöre özel uygulama kılavuzu vardır.
Sektöre özel uygulama kılavuzu vardır.	8.1.3 İşe alınmanın şartları	Sektöre özel uygulama kılavuzu vardır.
	8.2 Çalışma Sırasında	
►	8.2.1, 8.2.2, 8.2.3	◄
	8.3 Görev değişikliği veya işten ayrılma	
►	8.3.1, 8.3.2, 8.3.3	◄

27011:2008 (İletişim)	27002:2005	27019:2013 (Enerji)
	9.1 Güvenlik Alanı	
Sektöre özel uygulama kılavuzu vardır.	9.1.1 Fiziksel güvenlik sınırı	Sektöre özel uygulama kılavuzu vardır.
Sektöre özel uygulama kılavuzu vardır.	9.1.2 Fiziksel giriş kontrolleri	Sektöre özel uygulama kılavuzu vardır.
►	9.1.3, 9.1.4, 9.1.5, 9.1.6	◄
9.1.7 İletişim merkezlerinin güvenliği (Securing communication centres)	X	X
9.1.8 İletişim teçhizat odalarının güvenliği (Securing telecommunications equipment room)	X	X
9.1.9 Fiziksel olarak izole edilmiş operasyon alanlarının güvenliği (Securing physically isolated operation areas)	X	X
X	X	9.1.7 Kontrol merkezlerinin güvenliği (Securing control centers)
X	X	9.1.8 Teçhizat odalarının güvenliği (Securing equipment rooms)
X	X	9.1.9 Çevresel mekânların güvenliği (Securing peripheral sites)
	9.2 Ekipman Güvenliği	
Sektöre özel uygulama kılavuzu vardır.	9.2.1 Ekipman yerleşimi ve koruması	Sektöre özel uygulama kılavuzu vardır.
Sektöre özel uygulama kılavuzu vardır.	9.2.2 Destek hizmetleri	Sektöre özel uygulama kılavuzu vardır.
►	9.2.3 Kablolama güvenliği	Sektöre özel uygulama kılavuzu vardır.
►	9.2.4, 9.2.5, 9.2.6, 9.2.7	◄
9.3 Diğer tarafın sağladığı güvenlik (Security under the control of other party)	X	X
9.3.1 Diğer taşıyıcının tesisindeki cihazların güvenliği (Equipment sited in other carrier's premises)	X	X
9.3.2 Müşterinin tesisindeki cihazların güvenliği (Equipment sited in user premises)	X	X
9.3.3 Birbirine bağlı iletişim servislerinin güvenliği (Interconnected telecommunications services)	X	X
X	X	9.3 Üçüncü tarafların tesislerinde güvenlik (Security in premises of 3rd parties)
X	X	9.3.1 Diğer enerji kuruluşlarının tesislerinde bulunan cihazlar (Equipment sited on the premises of other energy utility organizations)
X	X	9.3.2 Müşterinin tesislerinde bulunan cihazlar (Equipment sited on customer's premises)

27011:2008 (İletişim)	27002:2005	27019:2013 (Enerji)
X	X	9.3.3 Birbirine bağlı kontrol ve iletişim sistemleri (Interconnected control and communication systems)
	10.1 İşletme Prosedürleri ve Sorumluluklar	
Sektöre özel uygulama kılavuzu vardır.	10.1.1 Belgelenmiş işletme prosedürleri	Sektöre özel uygulama kılavuzu vardır.
Sektöre özel uygulama kılavuzu vardır.	10.1.2 Değişim kontrolü	◄
►	10.1.3 Görevler ayrılığı	◄
Sektöre özel uygulama kılavuzu vardır.	10.1.4 Geliştirme sistemi, test sistemi ve aktif sistemlerin ayrılması	Sektöre özel uygulama kılavuzu vardır.
	10.2 Üçüncü taraflardan alınan hizmetin yönetilmesi	
►	10.2.1, 10.2.2, 10.2.3	◄
	10.3 Sistem Planlama ve Kabul Etme	
►	10.3.1, 10.3.2	◄
	10.4 Kötü Niyetli Yazılımlara Karşı Korunma	
►	10.4.1 Kötü niyetli yazılımlara karşı kontroller	Sektöre özel uygulama kılavuzu vardır.
Sektöre özel uygulama kılavuzu vardır.	10.4.2 Mobil yazılımlarla ilgili denetimler	Sektöre özel uygulama kılavuzu vardır.
	10.5 Yedekleme	
►	10.5.1	◄
	10.6 Ağ Güvenliğinin Yönetilmesi	
►	10.6.1	◄
Sektöre özel uygulama kılavuzu vardır.	10.6.2 Ağ hizmetlerinin güvenliği	◄
10.6.3 Verilen iletişim hizmetlerinin güvenliğinin yönetimi (Security management of telecommunications services delivery)	X	X
10.6.4 İstenmeyen e-Posta'ya müdahale (Response to spam)	X	X
10.6.5 DDoS saldırılarına müdahale (Response to DoS/DDoS attacks)	X	X
X	X	10.6.3 Süreç kontrol verilerinin iletişim güvenliği (Securing process control data communication)
	10.7 Bilgi ortamı yönetimi ve güvenlik	
►	10.7.1, 10.7.2, 10.7.3, 10.7.4	◄
	10.8 Bilgi ve Yazılım Değiş Tokuşu	
►	10.8.1, 10.8.2, 10.8.3, 10.8.4, 10.8.5	◄
	10.9 Elektronik Ticaret Hizmetleri	
►	10.9.1, 10.9.2, 10.9.3	◄
	10.10 Sistem Erişiminin Gözlenmesi ve Kullanımı	
Sektöre özel uygulama kılavuzu vardır.	10.10.1 Olay kayıtlarının tutulması	Sektöre özel uygulama kılavuzu vardır.
►	10.10.2, 10.10.3, 10.10.4, 10.10.5	◄
►	10.10.6 Saat senkronizasyonu	Sektöre özel uygulama kılavuzu vardır.

27011:2008 (İletişim)	27002:2005	27019:2013 (Enerji)
X	X	10.11 Eski sistemler (Legacy systems)
X	X	10.11.1 Eski sistemlere müdahale (Treatment of legacy systems)
X	X	10.12 Güvenlik işlevleri (Safety functions)
X	X	10.12.1 Güvenlik işlevlerinin bütünlüğü ve erişilebilirliği (Integrity and availability of safety functions)
	11.1 Erişim Denetimi Gereksinimleri	
Sektöre özel uygulama kılavuzu vardır.	11.1.1 Erişim denetimi politikası	Sektöre özel uygulama kılavuzu vardır.
	11.2 Kullanıcı Erişiminin Yönetilmesi	
▶	11.2.1, 11.2.2, 11.2.3, 11.2.4	◀
	11.3 Kullanıcı Sorumlulukları	
▶	11.3.1 Parola kullanımı	Sektöre özel uygulama kılavuzu vardır.
▶	11.3.2, 11.3.3	◀
	11.4 Ağ Erişimi Denetimi	
▶	11.4.1, 11.4.2, 11.4.3, 11.4.4	◀
▶	11.4.5 Ağlardaki ayrım	Sektöre özel uygulama kılavuzu vardır.
▶	11.4.6, 11.4.7	◀
11.4.8 Kullanıcıların taşıyıcı sistemlerin kimliğini doğrulaması (Telecommunications carrier identification and authentication by users)	X	X
X	X	11.4.8 Harici süreç kontrol sistemlerinin mantıksal bağlantısı (Logical coupling of external process control systems)
	11.5 İşletim Sistemi Erişim Denetimi	
▶	11.5.1	◀
▶	11.5.2 Kullanıcı tanımlaması ve doğrulaması	Sektöre özel uygulama kılavuzu vardır.
▶	11.5.3, 11.5.4	◀
▶	11.5.5 Oturum zaman aşımı	Sektöre özel uygulama kılavuzu vardır.
	11.6 Uygulama Erişimi Denetimi	
▶	11.6.1, 11.6.2	◀
	11.7 Mobil Bilgi İşlem ve Uzaktan Çalışma	
▶	11.7.1, 11.7.2	◀
	12.1 Bilgi Sistemlerinin Güvenlik Gereksinimleri	
▶	12.1.1 Güvenlik gereksinimlerinin analizi ve özelleştirilmesi	Sektöre özel uygulama kılavuzu vardır.
	12.2 Uygulamaların Doğru Çalışması	
▶	12.2.1, 12.2.2, 12.2.3, 12.2.4	◀
	12.3 Kriptografik Kontroller	
▶	12.3.1, 12.3.2	◀

27011:2008 (İletişim)	27002:2005	27019:2013 (Enerji)
	12.4 Sistem Dosyalarının Güvenliği	
Sektöre özel uygulama kılavuzu vardır.	12.4.1 Çalışmakta olan sistem yazılımın denetimi	Sektöre özel uygulama kılavuzu vardır.
►	12.4.2, 12.4.3	◄
	12.5 Geliştirme ve Destek Süreçlerinde Güvenlik	
►	12.5.1, 12.5.2, 12.5.3, 12.5.4, 12.5.5	◄
	12.6 Teknik Açıklık Yönetimi	
►	12.6.1	◄
Sektöre özel uygulama kılavuzu vardır.	13.1.1 Bilgi güvenliği olaylarının rapor edilmesi	◄
►	13.1.2 Bilgi güvenliği zafiyetlerinin rapor edilmesi	◄
	13.2 Bilgi Güvenliği Olaylarının Yönetimi ve İyileştirmeler	
Sektöre özel uygulama kılavuzu vardır.	13.2.1 Sorumluluklar ve prosedürler	◄
Sektöre özel uygulama kılavuzu vardır.	13.2.2 Bilgi güvenliği olaylarından deneyim edinme	◄
	14.1 İş Sürekliliği Yönetiminin Bilgi Güvenliği Boyutu	
Sektöre özel uygulama kılavuzu vardır.	14.1.1 İş sürekliliği yönetim sürecinin bilgi güvenliğini içermesi	Sektöre özel uygulama kılavuzu vardır.
►	14.1.2	◄
Sektöre özel uygulama kılavuzu vardır.	14.1.3 Bilgi güvenliğini içeren iş sürekliliği planlarının geliştirilmesi ve uygulanması	◄
►	14.1.4, 14.1.5	◄
X	X	14.2 Başlıca acil hizmetler (Essential emergency services)
X	X	14.2.1 Acil durum iletişimi (Emergency communication)
	15.1 Yasal Gereklere Uyumluluk	
►	15.1.1 İlgili yasaların belirlenmesi	Sektöre özel uygulama kılavuzu vardır.
►	15.1.2, 15.1.3, 15.1.4, 15.1.5, 15.1.6	◄
15.1.7 İletişimin Gizliliği (Non-disclosure of communciations)	X	X
15.1.8 Temel haberleşme (Essential communications)	X	X
15.1.9 Acil eylemlerin yasallığı (Legality of emergency actions)	X	X
	15.2 Güvenlik Politikası ile Uyum ve Teknik Uyum	
►	15.2.1, 15.2.2	◄
	15.3 Bilgi Sistemi Denetimi İle İlgili Hususlar	
►	15.3.1, 15.3.2	◄
13 sektöre özel önlem 26 sektöre özel uygulama kılavuzu	**133 Kontrol**	**11 sektöre özel önlem 31 sektöre özel uygulama kılavuzu**

Tablo-1. 27002, 27011 ve 27019 standartlarındaki önlemlerin karşılaştırılması

Tablonun içeriğini aşağıdaki şekille özetlemek de mümkündür:

Şekil-2. 27002, 27011 ve 27019 standartlarındaki önlemlerin karşılaştırılması

27002 standardında bulunmayıp enerji sektörüne özel uygulama kılavuzu 27019'da yer alan önlemler gözden geçirildiğinde, şu konularda hassasiyet gösterildiği gözlenmektedir:

a. Sistem kontrol merkezlerinde yaşanabilecek aksaklıkların neden olabileceği geniş kapsamlı etkiler göz önünde bulundurularak, kontrol merkezlerinin, merkezlerde kritik cihazların bulunduğu odaların ve kontrol merkezlerine ev sahipliği yapan tesislerin fiziksel ve çevresel güvenliğinin sağlanması.
b. Enerji sektörünün (üretim, iletim, dağıtım vb.) çok katmanlı, kurumlararası etkileşimli yapısı göz önünde bulundurularak, paydaş kurumların ve müşterilerin tesislerinde yer alan sistem bileşenlerinin güvenliğinin sağlanması, kontrol ve iletişim sistemleri arasındaki bağlantıların yönetilmesi, izlenmesi ve gerektiğinde paydaşların sistemlerinden ayrılmak üzere tedbirler alınması.
c. Özellikle geniş alanlara yayılan dağıtım ve iletim sistemlerinde söz konusu olabilecek riskler göz önünde bulundurularak, süreç kontrol verisinin gizlilik, bütünlük ve sürekliliğinin güvence altına alınması.
d. Kurumsal bilişim sistemlerinden çok daha uzun süre hizmet veren ve güvenlik işlevlerinden yoksun olabilen Endüstriyel Kontrol Sistemlerin'den kaynaklanan risklerden korunma.
e. Kurum içinden ve paydaş kurumlardan afet ve acil durumlarda iletişim halinde kalınması gereken personel ile ve vazgeçilmez kontrol sistemleri ile muhaberenin sürdürülmesini ve olağanüstü durumun atlatılmasını güvence altına alacak planlamanın yapılması, önlemlerin alınması.

27019 standardında yukardaki konuların herbiri ile ilgili olarak son derece somut öneriler bulunmaktadır.

27011 standardının da benzer başlıklara yoğunlaştığı görülmektedir. İlave olarak SPAM (yığın E-posta) ve DoS (servis dışı bırakma) saldırılarına dikkat çekilmekte ve bu bağlamda alınabilecek önlemlerden bahsedilmektedir.

Standartlarda dile getirilen risklerin tamamı, Türkiye için de söz konusu olan risklerdir. Şöyle ki, kritik enerji altyapıları her tür fiziksel ve çevresel riskle karşı karşıyadır. Türkiye, geniş bir coğrafyaya yayılmış olması dolayısı ile enerji altyapıları geniş alan ağları ile haberleşmektedir. Enerji altyapılarında miyadı dolmuş kontrol sistemleri ile karşılaşmak sürpriz olmamaktadır. Afet ve acil durumlar Türkiye'de gündelik yaşamın ayrılmaz, nerede ise kanıksanmış parçası haline gelmiştir. Her tür bilgi sistemine SPAM ve Dos saldırıları yapılmaya devam etmektedir. Dolayısı ile standartlarda dile getirilen önerilerin yurdumuz için de gerekli ve geçerli olduğu, Kritik Enerji Altyapısı işleten kuruluşlar tarafından gözden geçirilmelerinin son derece faydalı olacağı kesindir.

VI. SONUÇ

EPDK, elektrik, petrol ve doğal gaz piyasalarında lisanslı faaliyet gösteren kurumların bazılarına TS ISO/IEC 27001 standardına uyum mecburiyeti getirmiştir. 27001 standardına uyum kapsamında, kurumda mevcut risklerin işlenmesi için 27002 standardında yer alan önlemlerin bir bölümünün uygulanması da gerekir. Bu aşamada, enerji ve iletişim sistemlerine özel 27011 ve 27019 standartlarının da faydalı olacağı unutulmamalıdır. Yürürlüğe girme tarihi olarak belirtilen 1 Mart 2016 ile birlikte bilgi güvenliği yönetim sistemini bir kurumda kurmak ve çalıştırmak için bir yılı aşkın süre gerektiği de kurumlar tarafından göz önünde bulundurulmalıdır.

REFERANSLAR

[1] 4890 sayılı Bakanlar Kurulu Kararı-Ulusal Siber Güvenlik Stratejisi ve 2013-2014 Eylem Planı, T.C. Resmi Gazete, 20 Haziran 2013

[2] "COUNCIL DIRECTIVE 2008/114/EC on the identification and designation of European critical infrastructures and the assessment of the need to improve their protection", Official Journal of the European Union, 23.12.2008.

[3] Kritik Altyapı Sektörleri, ABD Anayurt Güvenliği Bakanlığı resmi web sitesi, http://www.dhs.gov/critical-infrastructure-sectors 10 Ocak 2015'de erişildi.

[4] EKS-BOME 2013 Yılsonu Raporu, (ICS-CERT Year in Review, Industrial Control Systems Cyber Emergency Response Team, 2013), https://ics-cert.us-cert.gov/sites/default/files/documents/Year_In_Review_FY2013_Final.pdf , 10 Ocak 2015'de erişildi.

[5] 4628 sayılı Enerji Piyasası Düzenleme Kurumunun Teşkilat ve Görevleri Hakkında Kanun, EPDK resmi web sitesi, http://www.epdk.gov.tr/index.php/epdk-hakkinda , 10 Ocak 2015'de erişildi.

[6] Elektrik Piyasası Lisans Yönetmeliğinde Değişiklik Yapılmasına Dair Yönetmelik, EPDK resmi web sitesi, www.epdk.gov.tr/documents/elektrik/mevzuat/yonetmelik/elektrik/lisans/Epdk_Ynt_Deg_EPLY_26122014_29217.doc , 10 Ocak 2015'de erişildi.

[7] Doğal Gaz Piyasası Lisans Yönetmeliğinde Değişiklik Yapılmasına Dair Yönetmelik, EPDK resmi web sitesi, www.epdk.gov.tr/documents/dogalgaz/mevzuat/yonetmelik/dogalgaz/lisans/Ddp_Ynt_Deg_Lisans_26122014.docx , 10 Ocak 2015'de erişildi.

[8] Petrol Piyasası Lisans Yönetmeliğinde Değişiklik Yapılmasına Dair Yönetmelik, EPDK resmi web sitesi, www.epdk.gov.tr/documents/petrol/mevzuat/yonetmelik/petrol/lisans/Ppd_Ynt_Deg_Lisans_26122014.docx , 10 Ocak 2015'de erişildi.

[9] Orhan Çalık, "ISO 27001:2013 Bilgi Güvenliği Yönetim Sistemi Standardındaki Değişiklikler ve Yenilikler", Ulusal Bilgi Güvenliği Kapısı. http://www.bilgiguvenligi.gov.tr/bt-guv.-standartlari/iso-27001-2013-bilgi-guvenligi-yonetim-sistemi-standardindaki-degisiklikler-ve-yenilikler.html , 10 Ocak 2015'de erişildi.

[10] Fikret Ottekin, "Çok Katmanlı ISO 27001 Süreci", Ulusal Bilgi Güvenliği Kapısı. http://www.bilgiguvenligi.gov.tr/bt-guv.-standartlari/cok-katmanli-iso-27001-sureci.html , 11 Ocak 2015'de erişildi.

[11] Fikret Ottekin, "Bilgi Güvenliğinde ISO 27000 Standartlarının Yeri ve Öncelikli 27002 Kontrolleri", Ulusal Bilgi Güvenliği Kapısı. http://www.bilgiguvenligi.gov.tr/bt-guv.-standartlari/bilgi-guvenliginde-iso-27000-standartlarinin-yeri-ve-oncelikli-iso-27002-kontrolleri.html , 11 Ocak 2015'de erişildi.

[12] ISO/IEC TR 27019:2013 (en) Information technology - Security techniques - Information security management guidelines based on ISO/IEC 27002 for process control systems specific to the energy utility industry. https://www.iso.org/obp/ui/#iso:std:iso-iec:tr:27019:ed-1:v1:en , 10 Ocak 2015'de erişildi.

[13] ISO/IEC 27011:2008 (en) Information technology - Security techniques - Information security management guidelines for telecommunications organizations based on ISO/IEC 27002 https://www.iso.org/obp/ui/#iso:std:iso-iec:27011:ed-1:v1:en , 10 Ocak 2015'de erişildi.

Fikret Ottekin, 1990 yılında Orta Doğu Teknik Üniversitesi Elektrik–Elektronik Mühendisliği Bölümünden mezun olmuştur. 1992 yılında aynı bölümde yüksek lisans çalışmasını tamamlamıştır.

1992-2007 tarihleri arasında ASELSAN Haberleşme Cihazları Grubunda çeşitli sivil ve askeri projelerde yazılım mühendisi ve tasarım lideri olarak çalışmış, yurt içinde ve yurt dışında görevlerde bulunmuştur.

2007-2015 yılları arasında TÜBİTAK Siber Güvenlik Enstitüsü'nde başuzman araştırmacı olarak görev yapmıştır. Bilişim teknolojileri ürünlerinin ortak kriterler uyarınca belgelendirilmesi, siber güvenlik ve ISO 27001 tabanlı bilgi güvenliği yönetim sistemleri konularında danışman olarak çalışmıştır. 2012-2013 yıllarında "Kritik Altyapılarda Bilgi Güvenliği Yönetimi" projesinin yöneticiliğini yapmıştır. Siber Güvenlik Kurulu Sekreterya'sına Ulusal Siber Güvenlik Yönetimi konusunda danışmanlık yapmış, Ulusal Siber Güvenlik Stratejisi ve Eylem Planlarının hazırlanmasına katkı sunmuştur.

Fikret Ottekin Eylül 2015 tarihinden itibaren ICterra A.Ş.'de Genel Müdür Danışmanı olarak görev yapmaktadır.

Orhan Çalık, 2011 yılında Berlin Teknik Üniversitesi Bilişim (İnformatik) Bölümünden lisans ve yüksek lisans derecesini alarak mezun olmuştur.

Şu anda Tübitak Bilgem Siber Güvenlik Enstitüsü'nde Siber Güvenlik Hizmetleri Birimi'nde Uzman Araştırmacı olarak çalışmaktadır.

ISO 27001 tabanlı bilgi güvenliği yönetim sistemleri konularında çeşitli kamu kurum ve kuruluşlarında danışman olarak çalışmıştır.

Kritik altyapılarda bilgi güvenliği, ülkelerin siber güvenlik stratejileri, kurumsal bilgi güvenliği yönetimi, risk analizi ve yönetimi konularında çalışmaya devam etmektedir.

VERİTABANI GÜVENLİĞİNDE SALDIRI TAHMİNİ VE TESPİTİ İÇİN KULLANICILARIN SINIFLANDIRILMASI

Çiğdem Bakır, Veli Hakkoymaz

Çiğdem Bakır, Yıldız Teknik Üniversitesi, Elektrik-Elektronik Fakültesi, Bilgisayar Mühendisliği, İstanbul, Türkiye, cigdem@ce.yildiz.edu.tr

Veli Hakkoymaz, Yıldız Teknik Üniversitesi, Elektrik-Elektronik Fakültesi, Bilgisayar Mühendisliği, İstanbul, Türkiye, veli@ce.yildiz.edu.tr

Özet — Kullanıcı ve uygulamaların Internet üzerindeki web sitelerini kullanarak bilgi edinmesi, belge ve bilgi paylaşımı, bankacılık ve diğer işletimsel işlemleri gerçekleştirmesi her geçen gün daha da artmaktadır. Ancak son zamanlarda Internet kullanımının yaygınlaşmasıyla birlikte yetkisiz erişim, veri tutarlılığı, veri bütünlüğü ve veri gizliliği gibi birtakım güvenlik problemleri de ortaya çıkmıştır. Bu durum kişisel ve kamusal alanlarda kullanılan bilgilerin korunmasını zorunlu kılmaktadır. Bu çalışmada, kullanıcıların yaptıkları işlem sayısı, IP adresleri, kullandıkları veri miktarı, yaptıkları işlem türü ve üstlendikleri roller gibi kriterler dikkate alınarak kullanıcı saldırılarının tespit edilmesi amacıyla ortak bir model oluşturulmuştur. Böylelikle risk teşkil eden kullanıcı grupları önceden farkedilerek bilgi güvenliğini sağlamak amaçlanmıştır.

Anahtar Kelimeler — Veritabanı Güvenliği, Saldırı Tespit Sistemleri, Log Kayıtları, Risk Analizi, Saldırı Tahmin Sistemleri.

Abstract — Nowadays, user and applications are increasingly getting used to sharing documents and information, performing banking and other business operations by using the web sites on the Internet. However, together with the proliferation of the Internet use, some security problems such as unauthorized access, data coherency, data integrity and data confidentiality have resulted in. This necessitates that information used in personal and public fields to be protected. In this study, a common model is formed in order to detect the user attacks by means of the criteria such as the number of operations performed by the users, IP addresses, the amount of data used, types of operations performed to the database and user roles. Thus, in order to provide database security, it is our aim to recognize in advance and categorize the user groups that may pose a risk.

Index Terms — Database Security, Intrusion Detection Systems, Log Records, Risk Analysis, Intrusion Prediction Systems.

I. GİRİŞ

Teknolojinin hızla gelişmesiyle birlikte hem kişisel amaçlar hem de bankacılık, bilgi paylaşımı, e-ticaret, iletişim gibi birçok alanda Internet vazgeçilmez bir ihtiyaç olmuştur. Ancak Internetle birlikte bilgi sızması, bilginin yetkisiz kişilerce ele geçirilmesi, değiştirilmesi, bilgi gizliliğinin sağlanamaması vb. ortaya çıkan sorunların çözülmesi amacıyla birtakım çalışmalar yapılmıştır[13,14,15,16]. Bu çalışmalar genellikle oluşan saldırı sonrasında yapılacak işlemlerle ilgilidir. Bu çalışma ise, farklı olarak log kayıtları ile olası kullanıcı saldırılarını önceden tahmin etmeyi ve ileride risk oluşturabilecek saldırıları oluşmadan önlemeyi amaçlamıştır. Yani, sistemi tehdit eden kullanıcılar belirlenmiş ve risk analizi yapılmıştır.

Saldırısızlık garantisi olmadığı (bilgi/kaynak paylaşımının vazgeçilmezliği ve sadece iyiliklerle dolu bir dünyada yaşamadığımız) için veritabanı güvenliği problemini ele almak gerekir. Veritabanı sistemlerinde veri/bilgi güvenliği problemi birkaç aşamada ele alınabilir.

Bilgi güvenliği problemi oluşmadan önce: Potansiyel saldırganı (uygulama/kullanıcı) saldırı oluşmadan önce, çeşitli sisteme erişim biçimi ya da işlemler gibi davranışlara bakarak tahmin etme ve sistem yönetici ve sorumlularını bu potansiyel saldırı ihtimaline göre uyanık olmalarını sağlamak amaçlanır.

Bilgi güvenliği problemi oluştuktan sonra: Bilgi güvenliği problemi olmayacak varsayımı ile hiçbir önlem almama ve problemin oluşmasını bekleme durumudur. Ancak problem oluştuktan sonra sistemin durumunu inceleyerek bir saldırı ya da bilgi güvenliği problemi oluştuğunu tespit etmek ve bu sorunu ortadan kaldırmak için yapılacak çalışmalar.

Eğer bir sistem, bilgi güvenliği problemi olmayacağını garanti etmiyor ve saldırı tespit/kurtarma mekanizması sunmuyorsa, sisteme bir saldırı gerçekleşebilir ancak kullanıcılar ne olduğunu ayırt edemez. Bu durumda güvenli olmayan sistem, kullanıcılarca kullanılmaya devam edecek ve istenmeyen durumlar oluşacaktır. Bu, kabul edilemez bir durum gibi gözükse de eğer saldırı olma olasılığı çok düşükse ve sistem kritik bilgi içermiyorsa tolere edilebilir bir durumdur. Ancak tersi durumda, bilgi güvenliğini garanti eden ve saldırı potansiyellerini ortadan kaldıran önlemleri maliyetine bakmaksızın değerlendirmek gerekir. Bu çalışma, bilgi güvenliği problemi oluşmadan önce alınacak önlemler kapsamında değerlendirilmelidir.

Yapılan çalışmada örnek log kayıtları analiz edilerek veritabanına aktarılmıştır. Buradan, kullanıcıların davranışlarını modelleyebilecek çeşitli kıstaslar kullanarak kurallar oluşturulmuştur. Diğer çalışmalardan farklı olarak bilgi sızmalarına karşı bir risk analizi gerçekleştirilmiştir. Ayrıca oluşturulan modele göre en riskli, riskli, orta riskli, düşük riskli ve en düşük riskli olmak üzere kullanıcılar risk durumuna göre beş grupta kümelenmiştir.

Yapılan çalışmanın ilk aşamasında düzensiz, dağınık halde bulunan log kayıtları toplanmış ve filtreden geçirilerek gereksiz, üzerinde işlem yapılmayacak alanlar temizlenmiştir[3]. Böylelikle log kayıtları analiz edilerek içerdiği veriler anlamlı bilgi haline getirilmiştir. İkinci aşamada, normalize edilmiş veriler veritabanına aktarılarak kullanıcı kullanım ve operasyonlarını takip etmek için çeşitli kurallar oluşturulmuştur. Üçüncü aşamada, oluşturulan kuralları kullanarak kullanıcı saldırılarını belirleyebilmek amacıyla, ortak bir model gerçekleştirilmiştir. Son aşamada ise oluşturulan modelle hesaplanan risk oranına bağlı olarak kullanıcılar risk durumlarına göre gruplandırılmış ve

veritabanına aktarılmıştır. Ek olarak, risk grubu en yüksek düzeyde olan kullanıcılar ve IP adresleri raporlanmıştır.

Makalenin kalan kısmı şu şekilde organize edilmiştir: İlgili ve benzer çalışmalar II. kısımda ele alınacaktır. III. kısımda log kayıtları ve bilgi güvenliği anlatılırken, IV. kısımda yapılan çalışma ve tasarımı anlatılacaktır. Son kısımda (V), çalışmanın katkıları özetlenecek ve gelecek çalışmaların ne yönde olacağı tartışılacaktır.

II. İLGİLİ ÇALIŞMALAR

Bu çalışmada veritabanı güvenliği özellikle sisteme tanımlı olan kullanıcılar tarafından yapılabilecek saldırıların tahmin edilmesi amaçlanmıştır. Sistem kullanıcılarının risk bazlı sınıflandırılmaları için ortak bir model geliştirilerek kullanıcıların risk oranları hesaplanmıştır. Virüs, solucan, Truva atı, casus yazılım gibi kötümcül yazılımlar ya da sahte IP adresi ile sisteme girme, port tarayıcı saldırıları, kullanıcıların yapmış oldukları hatalı işlemler ve yetkisiz kullanıcıların veri üzerinde okuma ve yazma yaparak sisteme zarar vermesi veritabanı güvenliğini olumsuz etkilemekte, bu durum bu tür ölçümlerle tahmin edilmektedir[17,19].

Veritabanı saldırı tespit sistemleri imza tabanlı ve imza tabanlı olmayan sistemler olmak üzere ikiye ayrılır. İmza tabanlı olanlar, daha önceden gerçekleşmiş, bilinen saldırıların sayısal imzalarının kaydedildiği bir veritabanından yararlanırlar[4]. Bu saldırılar genelde virüs, solucan, Truva atı ve casus yazılım gibi kötücül yazılımlar aracılığıyla yapılmaktadır[8]. Bunların dezavantajları, imza bilgilerinin güncel tutulmasının maliyetli olması, yani karşılaşılan saldırıların imzası henüz bilinmediğinden yeni saldırı türlerini tespit edememesidir.

İmza tabanlı olmayan saldırı tespit sistemleri, sistemde gerçekleşen anormal durumların tespitinde kullanılmaktadır. Bunlar kullanıcı erişimi, rol bazlı erişim gibi iç kaynaklı saldırıları tespit etmek içindir. Saldırıların çoğu, bazı kullanıcıların donanım ve yazılım açıklarını kullanarak istedikleri bilgilere ulaşması ile oluşur. Genelde SQL enjeksiyon saldırılarını içerir[4].

Veritabanı saldırılarını önceden tahmin etme ve güvenlik açıklarını saldırı oluşmadan farketmek amacıyla çok etmenli istatiksel bir tahmin sistemi Quickprop sinir ağları geliştirilmiştir[19]. Bu çalışmada gizli katmanlarını hesaplayabilmek için Pearson korelasyon katsayısı kullanılmış ve bir banka verisi üzerinde yetkisi olmayan kullanıcılar belirlenmeye çalışılmıştır. Kısa vadede gerçekleşen anormal ve hatalı kullanıcı davranışları bulunmuştur. Ancak uzun vadede gerçekleşen potansiyel riskli kullanıcılar ve kontrolsüz kullanıcı işlemleri bir günlük log kayıtları kullanıldığından tam olarak belirlenememiştir. Yani, uzun vadede gerçekleştirilecek saldırılar için bir risk analizi içermemektedir.

Hatalı kullanıcı davranışlarını çözebilmek amacıyla yapılan bir başka çalışma da, genetik algoritma kullanımıdır[20]. Genetik algoritma, sinir ağlarına dayalı olarak ağ özelliklerinden çeşitli kurallar oluşturarak elde edilen kurallar ile sınıflandırma yapar. Bu çalışmanın sonuçları diğer çalışmalarla karşılaştırılmalı olarak verilmiştir. Ancak bu çalışma da öncekinde olduğu gibi kısa vadede gerçekleşebilecek saldırılar için bir çözüm önerisi getirmiştir. Saklı Markov Modeli kullanarak saldırıyı tahmin ve önleme çalışması, diğer bir ilgili çalışmadır[21]. Saklı Markov Modeli, verilen durumlardan yola çıkarak gizli durumları bulmak için gerçekleştirilen bir sınıflandırma algoritmasıdır. Dağıtık veriler birbirleriyle çok büyük ağlar üzerinde haberleşir ve bu sebeple ciddi saldırılara açıktır. Bu çalışmada fuzzy tekniği kullanılarak risk analizi yapılmış, tehlikeli olarak giden paket oranı tespit edilmeye çalışılmıştır. Ayrıca dağıtık çevreler için risk oluşturacak saldırılar belirlenmeye çalışılmıştır.

Rol tabanlı erişim kontrol modeli kullanan saldırı tespit sisteminin ilk aşamasında veritabanı günlüğü incelenir. Geçmiş hareketler ve dahil oldukları rollere göre sınıflandırma modeli (Naive Bayes) oluşturulur. Kullanıcı hareketleri bu modele göre sınıflandırılır. Sonuçta bulunan rol veritabanı günlüğündeki kullanıcılarla karşılaştırılır. Buna göre, eğer bulunan rol ilgili kullanıcıyla tanımlı ise bir rol olarak kabul edilir, değilse alarm verilir. Ancak bu çalışma, sadece kullanıcıların rollerine göre saldırı tespiti yapmakta, kullanıcıların bireysel olarak yapmış oldukları hareketleri göz önüne almamaktadır[13].

Veritabanında kullanıcı erişiminden kaynaklanan saldırı tespit sistemine örnek Detection of Misuse in Database Systems (DEMIDS) gösterilebilir. Bu çalışma, ilişkisel veritabanları için iç kaynaklı saldırıların (hatalı davranışlar) tespit edilmesi içindir. Bu sistem denetleyici, veri işleyici, profil düzenleyici ve algılayıcı olmak üzere dört ana bileşenden oluşur. Denetleyici verileri toplar ve denetim günlüğüne kaydeder. Veri işleyici, verileri istenen yapı ve türlere dönüştürür. Profil düzenleyici, öğrenme ile herbir kullanıcı için bir profil oluşturur. Denetim aşamasında, kullanıcı aktivitelerinin şüpheli olup olmadıkları sık kullanılan kullanıcı profilleri ile karşılaştırılarak hesaplanmıştır. Önceki çalışmadan farklı olarak, sadece sistemdeki kullanıcı hareketleriyle değil, veritabanında olmayan kullanıcı için de bir profil belirlemektedir[14].

Çeşitli veri madenciliği teknikleri kullanan veritabanı saldırı tespit sistemleri veri nesneleri arasındaki okuma ve yazma bağımlılıklarını analiz eder[15]. Öncelikle saldırı içermeyen veritabanı hareketleri (günlüklerde) analiz edilerek okuma ve yazma bağımlılıklarını ifade eden kurallar belirlenir. Bu kurallar sistem için öğrenilen bir modeli oluşturur. Daha sonra gelen yeni hareketlerin, bu kurallara uyup uymadığına bakılarak saldırılar tespit edilir. Bu çalışma gerçek veriler için yapılmamış olup, büyük sentetik veriler için veritabanı sadece okuma ve yazma hareketleri için yapılmıştır.

Gerçek zamanlı veritabanı sistemlerinde, hareketler için belirli zaman kısıtlamaları (deadline) vardır. Bu veritabanları, değerleri zaman içerisinde değişen ve periyodik olarak güncellenen zaman boyutlu veri nesneleri için tasarlanmıştır. Hareketler belirli zaman sınırları içerisinde tanımlanmıştır. Gerçek zamanlı veritabanı sistemleri için Lee tarafından yapılan bir çalışmada, hareketlerin zaman imzaları kullanılmış ve zaman boyutlu verileri güncelleyen hareketler için ortaya çıkarmıştır[16]. Farklı olarak, gerçek zamanlı sistemlerde sadece yazma hareketi için güvenlik uyarısı verir.

Yapılan çalışmada gerçekleştirilen model ile kullanıcı erişiminden kaynaklanan saldırıların olup olmamasına bakılmaksızın kullanıcılar risk durumlarına göre sınıflandırılmıştır. Sistemdeki tüm kullanıcıların hareketleri

dikkate alınmıştır ve riskli kullanıcılar belirlenmeye çalışılmıştır. Böylelikle bir veritabanı saldırı tahmin sistemi geliştirilmeye çalışılmıştır.

III. BİLGİ GÜVENLİĞİ İÇİN LOG KAYITLARI

Bilgi güvenliği, bilgilerin izinsiz olarak yetkisiz kişilerin kullanımından ve değiştirilmesinden korunmasıdır. Bütünlük, gizlilik ve erişebilirlik olmak üzere üç temel unsurdan meydana gelmektedir. Bütünlük bilginin yetkisiz kişilerce yapılan rastgele değişiklerden korunmasını; gizlilik bilginin yetkisiz kişilerin erişimine izin verilmemesini; erişebilirlik ise bilginin yetkili kişiler tarafından ulaşılabilir olmasını ifade eder[9,11]. Bilgi güvenliğiyle sistemi tehdit eden riskler belirlenir. Gizlilik, erişebilirlik ve bütünlük sağlanarak iş sürekliliği artar[10].

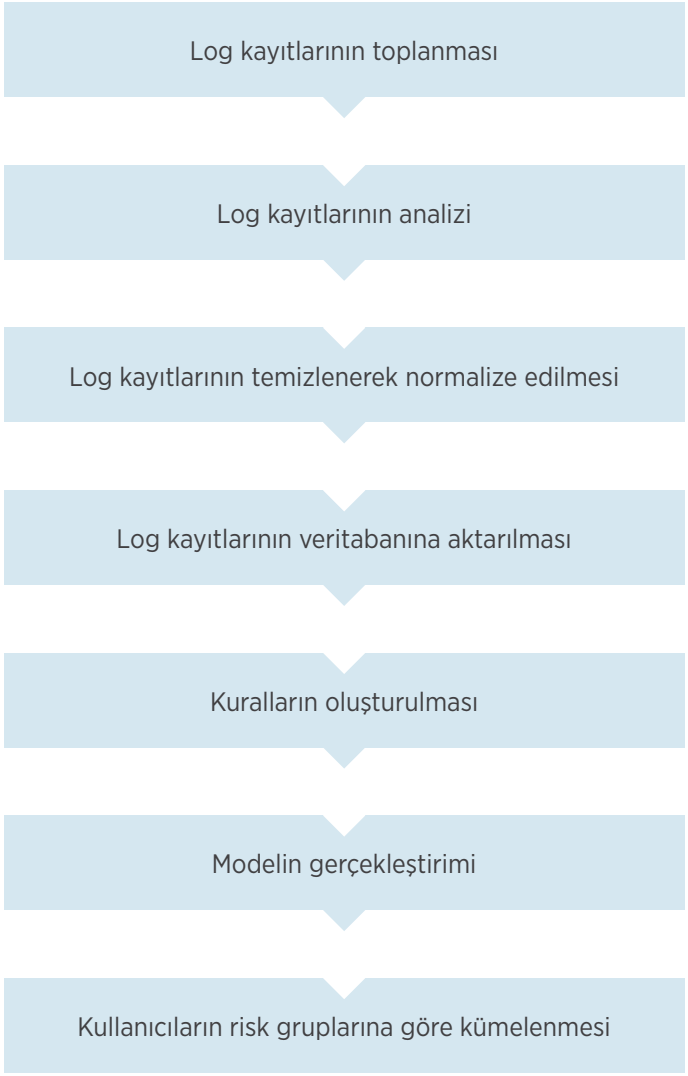
Günümüzde bilgi teknolojilerinin yaygınlaşmasıyla özellikle uygulamaların Internet üzerine taşınması ve Internet ortamında birçok bilginin paylaşılması ve hemen hemen tüm işlemlerin internet üzerinden gerçekleşmesi kötü niyetli veya yetkisiz kişilerin sistem bütünlüğüne zarar vermesine yol açmaktadır[2,12]. Dolayısıyla bir güvenlik sorunu ortaya çıkmaktadır. Veritabanında log yönetimi değişen bilgilerin izlenmesini olanaklı kıldığı gibi, bilgi güvenliğinin de en temel yapısını oluşturur. Bilgi güvenliği ihlallerinin yaşanmaması için kesintisiz bir log yönetimi gerekir. Log yönetimiyle birlikte sistemdeki tüm kullanıcılar, işlemler, işlemlerin zamanı, IP adresleri, başarılı ya da başarısız olma durumu izlenir ve tehlikeli bir durumla karşılaşıldığında log kayıtlarına bakılarak nedeni belirlenir. Bilgi güvenliği ve risklerin ortadan kaldırılması bakımından log yönetimi büyük önem taşır[1].

Log kayıtları incelenerek, veritabanında hatalı kullanım ve saldırılarının tespiti yapılmaktadır. Bunun için veritabanı nesneleri ve bu nesnelere yapılan işlemler değerlendirilir. Özellikle veritabanı sistem nesnelerine erişim işlemleri izlenmelidir. İşlemler ise özellikle sadece okunma amaçlı tasarlanmış bir tabloda değişiklik yapmaya çalışan bir işlem zararlı olarak görülebilir ve tespiti önemlidir.

Log yönetimi farklı kaynaklardan toplanan log verilerinin tek bir merkezde tutulması ile kolayca gerçekleştirilebilir[6]. Log yönetimiyle sadece sistemdeki anormal durumlar belirlenmez. Aynı zamanda ileride oluşabilecek güvenlik sorunlarına karşı da kurumlarda çalışan personeller için farkındalık oluşturur. Cobit ve ISO27001 gibi uluslararası standartlar log yönetimini desteklemektedir[7]. Log kayıtları tek bir merkezde toplanır, depolanır ve istenildiğinde hızlı bir şekilde erişilir. Ayrıca log kayıtları istenildiği zaman, geri getirilme özelliğine sahiptir. Veri bütünlüğü, tutarlılığı ve veri gizliliğinin sağlanması için veriler ile ilgili kullanıcıların yetkileri ve ilişkiler düzeyindeki erişimi belirlenir. OSSIM, Manage Engine Event Log Analyzer, Swatch ve File System Auditor gibi hazır yazılımlar log analizini gerçekleştirmek için kullanılan açık kaynak kodlu ücretsiz yazılımlardır[7].

IV. ÇALIŞMA ADIMLARI VE SİSTEM TASARIMI

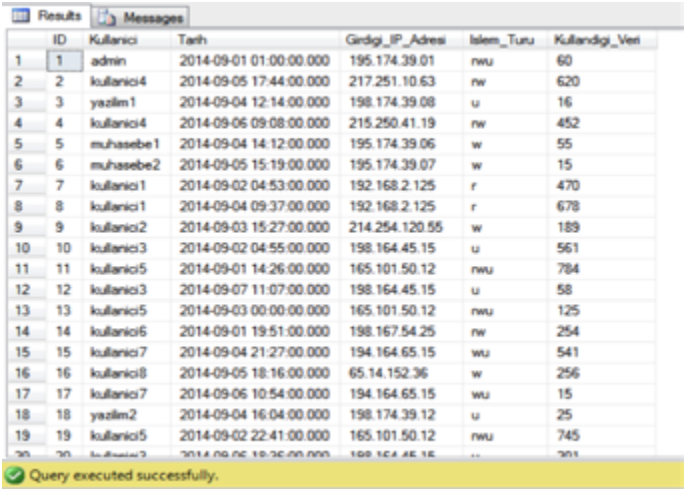
Log kayıtlarından veri güvenliğine tehdit oluşturabilecek kullanıcıların belirlenmesi çalışması birden fazla adımı içermektedir. Çalışmanın genel adımları Şekil 1'de gösterilmiştir.



Şekil. 1. Log kayıtları ile risk gruplarını belirleme adımları

A. Log Kayıtlarının Toplanması ve Normalize Edinmesi

Örnek çalışmada, 1 ay içinde sisteme giren 2673 kullanıcı ve 10872 işlem esas alınarak log veritabanı oluşturulmuştur. Ancak toplanan log kayıtları (veri) karmaşık ve düzensiz bir haldedir. Toplanan log kayıtlarını anlamlı hale getirmek ve çalışma ile amaçlanan hedefe ulaşabilmek için çeşitli web madenciliği ve veri madenciliği teknikleri ile veri normalize edilmiş ve gerekli önişlem adımları yapılmıştır[1]. Normalize edilecek veri, log kayıtlarının içerdiği bilgiler, sunucu tarafından toplanan veriler, kullanıcılar ve kullanıcıların gerçekleştirdiği işlemlerden oluşur. Tüm gereksiz ve kullanılamayacak olan veriler log kayıtlarında temizlendikten sonra benzer işlemi gerçekleştiren kullanıcılar, yaptıkları işlemler, işlem miktarları ve kullanıkları veri miktarları belirlenerek veriler düzenli hale getirilmiştir. Normalizasyon, log dosyasındaki verilerden anlamlı bilgi oluşturabilmek için yapılan veri temizleme, filtreleme, raporlama ve indexleme işidir. Yani, log kayıtlarının istenilen formatta kullanılmasını sağlamak için gereksiz alanlar çıkarılarak normalizasyon işlemi yapılmıştır. Şekil 2'de normalize edilen bir kısım log kayıtlarının biçimi gösterilmiştir.



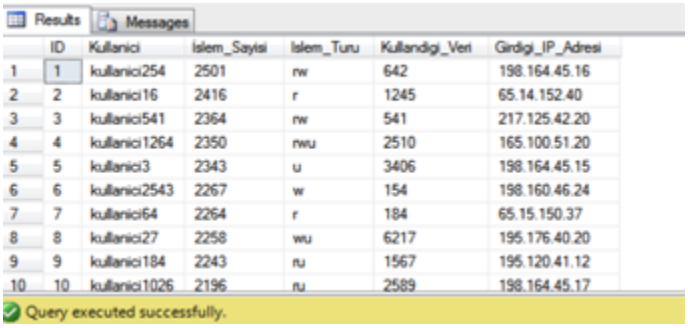
Results | Messages

	ID	Kullanıcı	Tarih	Girdigi_IP_Adresi	Islem_Turu	Kullandigi_Veri
1	1	admin	2014-09-01 01:00:00.000	195.174.39.01	rwu	60
2	2	kullanici4	2014-09-05 17:44:00.000	217.251.10.63	rw	620
3	3	yazilim1	2014-09-04 12:14:00.000	198.174.39.08	u	16
4	4	kullanici4	2014-09-06 09:08:00.000	215.250.41.19	rw	452
5	5	muhasebe1	2014-09-04 14:12:00.000	195.174.39.06	w	55
6	6	muhasebe2	2014-09-05 15:19:00.000	195.174.39.07	w	15
7	7	kullanici1	2014-09-02 04:53:00.000	192.168.2.125	r	470
8	8	kullanici1	2014-09-04 09:37:00.000	192.168.2.125	r	678
9	9	kullanici2	2014-09-03 15:27:00.000	214.254.120.55	w	189
10	10	kullanici3	2014-09-02 04:55:00.000	198.164.45.15	u	561
11	11	kullanici5	2014-09-01 14:26:00.000	165.101.50.12	rwu	784
12	12	kullanici3	2014-09-07 11:07:00.000	198.164.45.15	u	58
13	13	kullanici5	2014-09-03 00:00:00.000	165.101.50.12	rwu	125
14	14	kullanici6	2014-09-01 19:51:00.000	198.167.54.25	rw	254
15	15	kullanici7	2014-09-04 21:27:00.000	194.164.65.15	wu	541
16	16	kullanici8	2014-09-05 18:16:00.000	65.14.152.36	w	256
17	17	kullanici7	2014-09-06 10:54:00.000	194.164.65.15	wu	15
18	18	yazilim2	2014-09-04 16:04:00.000	198.174.39.12	u	25
19	19	kullanici5	2014-09-02 22:41:00.000	165.101.50.12	rwu	745

Query executed successfully.

Şekil. 2. Normalize edilmiş log kayıtları

B. Log Kayıtlarının Veritabanına Aktarılması



Results | Messages

	ID	Kullanici	Islem_Sayisi	Islem_Turu	Kullandigi_Veri	Girdigi_IP_Adresi
1	1	kullanici254	2501	rw	642	198.164.45.16
2	2	kullanici16	2416	r	1245	65.14.152.40
3	3	kullanici541	2364	rw	541	217.125.42.20
4	4	kullanici1264	2350	rwu	2510	165.100.51.20
5	5	kullanici3	2343	u	3406	198.164.45.15
6	6	kullanici2543	2267	w	154	198.160.46.24
7	7	kullanici64	2264	r	184	65.15.150.37
8	8	kullanici27	2258	wu	6217	195.176.40.20
9	9	kullanici184	2243	ru	1567	195.120.41.12
10	10	kullanici1026	2196	ru	2589	198.164.45.17

Query executed successfully.

Şekil. 3. Log Kayıtlarında kullanacak kıstasların veritabanına aktarılması

Şekil 2'de gösterilen normalize edilen kısmi log kayıtları oluşturulacak modeldeki kıstaslarda kullanılan işlem miktarı, işlem türü, kullandığı veri miktarı ve IP adreslerini gösterecek biçimde Şekil 3'de verilmiştir.

C. Modelin Gerçekleştirilmesi

Saldırı tahmin ve tespit sistemlerinde, özellik seçimi, modeli oluşturmakta oldukça önemlidir[18]. Bu çalışmada sistem kullanımı, kullanılan veri miktarı, kullanıcı yetkileri ve IP adresi ve sistemi yöneten ve gerçekleştiren kullanıcıların üstlendikleri roller model oluşturmak için temel kriterler olarak belirlenmiştir. Yapılan diğer çalışmalarda tek ya da iki kriter kullanılarak model gerçekleştirimi sağlanmıştır[17,18]. Gerçekleştirilen çalışmada ise birden fazla ve en çok kullanılan kriterler seçilerek sistemin güvenliğini tehdit edebilecek kötü niyetli kullanıcıların risk durumları analiz edilmiştir. Ayrıca sistem kullanıcılarının hepsini kapsayan ortak bir model gerçekleştirilmesi amaçlanmıştır. Sistem kullanımı ve sistemde kalma süresi çeşitli saldırıları ayırt etmeyi sağlar. Kullanıcıya verilen yetkiler, veritabanında meydana gelen hatalı işlemleri ya da yetkisi olmayan kullanıcıların sisteme girip girmediklerini belirlemede ana göstergelerdendir[4]. Kullanıcıların, okuma ve yazma işlemlerine yetkili ya da yetkisiz olup olmadığını yansıtır. Ayrıca port numarası, kaynak ve hedef IP adresi bilgileri SQL komutlarının gruplandırılmasında kullanılır. Veritabanı haberleşmesi farklı port üzerinde gerçekleştirilir. Port numarası, oturumda paketleri düzenlemek, paketlerin hatalı gidip gitmediğini, verinin eksik bir şekilde sunucuya gönderilip gönderilmediğini denetleyebilmek için kullanılabilir. Sahte IP adresi ile sisteme grime ve port tarayıcı saldırılarını tespit amacıyla IP adresi ve veri miktarı model oluşturmak için seçilen kriterlerdendir[17,19].

D. Kuralların Oluşturulması

Bizim çalışmamızda, kullanıcıların sistemdeki kullanım ve operasyonlarını dikkate alarak bazı kıstaslar tanımlanmış ve bu kıstaslara göre risk oranını belirleyebilmek için kurallar geliştirilmiştir (bkz. Şekil 3). Risk oranının hesaplanması için öngörülen kurallar aşağıdaki gibi belirlenmiştir:

1) Sistem Kullanımı : Sistemde kalma süresi, sisteme giriş sayısı ve sisteme ne kadar sıklıkla girdiği kullanıcının sistemle ne kadar alakalı olduğunu gösterir. Kullanıcının sistemde kalma miktarı arttıkça sistemle ilgili bilgi alma, bilgiyi değiştirme ya da bilgi ekleme gibi temel hareket (transaction) işlemlerini gerçekleştirme olasılığı artar. Bu sebeple kullanıcının sistem kullanım durumu bu çalışmada kıstas olarak kullanılmıştır. Sisteme girme sıklığına göre aşağıdaki kurallar belirlenmiştir: Sisteme,

- 1-10 kez girenlerin ağırlıkları 0,
- 11-20 kez girenlerin ağırlıkları 1,
- 21-50 kez girenlerin ağırlıkları 2,
- 51-100 kez girenlerin ağırlıkları 3,
- 101-250 kez girenlerin ağırlıkları 4,
- En az 251 kez girenlerin ağırlıkları 5,

olarak belirlenmiştir. Bu kurallar örnek veritabanındaki kullanıcıların kalma miktarına göre büyükten küçüğe doğru sıralandığında, tüm örnek log kayıtlarının dengeli bir şekilde dağılması amacıyla bu biçimde oluşturulmuştur.

2) Kullanılan veri miktarı : Bir kullanıcının sistemde kullandığı veri miktarı kullanıcılar arasında risk analizi yapabilmek için önemlidir. Çünkü kullanılan veri miktarı arttıkça kullanıcılar sistemde daha fazla bilgi edinebilir ve daha fazla bilgiyi kullanarak değiştirme imkanı bulabilir. Sistemde kullanılan veri miktarına göre;

- 0-50 KB arası kullananların ağırlıkları 0,
- 51-100 KB arası kullananların ağırlıkları 1,
- 101-150 KB arası kullananların ağırlıkları 2,
- 151-500 KB arası kullananların ağırlıkları 3,
- En az 501 KB arası kullananların ağırlıkları 4,

olarak belirlenmiştir.

3) Kullanıcı Yetkileri : Bir sistemde kullanıcı yetkileri read, write ve update işlemleri için belirlenir[5]. Sistem üzerinde okuma ve yazma yetkisi olmayan bir kullanıcının gizli bilgiler içeren veriyi okuması, veri üzerinde değişiklik yapması ve veri üzerine yazması veri gizliliği, veri bütünlüğü ve veri güvenirliği açısından risk taşır.

Sistemde kullanıcı yetkilerine göre;
- Read işlemi yapanların ağırlıkları 1,
- Write işlemi yapanların ağırlıkları 2,
- Update işlemi yapanların ağırlıkları 3,

olarak belirlenmiştir.

4) IP Adres Sıklığı: En çok kullanılan IP adresleri veritabanı güvenliği için riskli olarak kabul edilmiştir. Sistemde en çok kullanılan ilk 100 IP adresi belirlenmiştir.

- Sık kullanılmayan IP adreslerinden bağlananların ağırlıkları 0,
- Sık kullanılan 100 IP adreslerinden bağlananların ağırlıkları 1,

olarak belirlenmiştir.

5) Sistemi yöneten ve gerçekleyen kullanıcıların üstlendikleri roller : Sistemi gerçekleştiren kullanıcılar içinde yetkisi dışında başka kullanıcıların hareketlerini (transaction) görüntülemesi veri gizliliği, veri bütünlüğü ve veri güvenirliği açısından risk taşır. Bu kullanıcıların üstlendikleri rollere bakarak bir risk ve güven analizi yapılmaktadır[13]. Sistemi yöneten ve gerçekleştiren kullanıcıların üstlendikleri rollere göre;

- Yapılan işlemleri sadece görüntüleyenlerin ağırlıkları 0,
- Yapılan işlemleri görüntüleme ve yazma yapanların ağırlıkları 1 ,

olarak belirlenmiştir.

E. Kullanıcıların Kümelenmesi

Kurallar oluşturulduktan sonra risk oranının hesaplanabilmesi için eşitlik 1 kullanılmıştır:

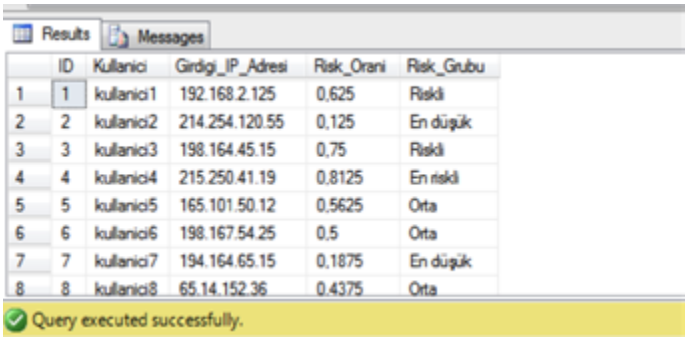
$$\text{Risk Oranı} = \sum_{i=1}^{n} (w_{a,i}, w_{b,i}, w_{c,i}, w_{d,i}, w_{e,i})/T \qquad (1)$$

Bu hesaplamada, $w_{a,i}$ **a** özelliğinin (sisteme giriş sayısı), $w_{b,i}$ **b** özelliğinin (kullanılan veri miktarı), $w_{c,i}$ **c** özelliğinin (kullanıcılara verilen yetki), $w_{d,i}$ **d** özelliğinin (IP adresi), $w_{e,i}$ ise **e** özelliğinin (sistemi yöneten ve gerçekleyen kullanıcıların rolleri) ağırlığını gösterir. **n** sistemde bulunan kullanıcı sayısını gösterirken, **T** ise herbir kriter içerisinde en güçlü ağırlığa sahip kuralların oluşturduğu toplam ağırlığı ifade eder. Ancak kullanıcı yetkileri kriteri için bir kullanıcı tüm yetkileri (read, write, update) gerçekleştirebileceğinden dolayı risk oranı hesaplanırken bu kriterin tüm ağırlıkları **T** değişkenine aktarılmıştır. Yukarıdaki formül dikkatle incelendiğinde risk oranının 0-1 arasında değerler alacağı gözlenecektir. Buna göre risk sınıflandırılması yaparsak;

Risk Oranı =
- 0-0.2 ise en düşük riskli
- 0.21-0.4 ise düşük riskli
- 0.41-0.6 ise orta riskli
- 0.61-0.8 ise riskli
- 0.81-1.0 ise en riskli

olmak üzere sistemi tehdit eden 5 grupta kümelenir. Tablo 1 de bazı kullanıcıların geliştirilen modele göre risk oranları ve risk grupları gösterilmektedir. Ayrıca bu bilgiler riskli kullanıcıların bulunmasını kolaylaştırmak için ilgili personel tarafından kullanılmak üzere veritabanında kaydedilmiştir (bkz.Şekil 4).

ID	Kullanıcı	Risk Oranı	Risk Grubu
1	kullanici1264	1	En riskli
2	kullanici10	0.75	Riskli
3	kullanici124	0.25	Düşük
4	kullanici26	0.18	En düşük
5	kullanici2350	0.68	Riskli
6	kullanici1557	0.12	En düşük
7	kullanici65	0.5	Orta
8	kullanici546	0.56	Orta
9	kullanici42	0.37	Düşük
10	kullanici9	0.06	En düşük



Şekil. 4. Kullanıcıların risk oranına göre gruplandırılması

V. DEĞERLENDİRME VE SONUÇ

Bu çalışma bir risk analizi yaparak, güvenlik açıklarının azaltılmasına yardımcı olmayı hedeflemektedir. Riskli kullanıcılar belirlenerek güvenlik analizleri yapılmış, sistemi kullanan personele gereken önlemlerin alınması önerilmiştir. Çeşitli kurallar kullanılarak geliştirilen modele göre sistemi tehdit edenlerin risk oranları hesaplanmış ve risk durumları sınıflandırılmıştır. Log kayıtlarına bakılarak olası risk grubu oluşturacak kullanıcı ve IP adresleri belirlenmiş ve bunlar log dosyasının tutulduğu veritabanına aktarılmıştır. Kümeleme işlemi kullanıcıların sisteme girme sıklığına, kullandıkları veri miktarına, kullanıcılara verilen yetkilere ve IP adreslerine göre yapılmıştır.

Bu çalışma özellikle kurumda güvenlik konusunda çalışan personele yol gösterici olacaktır. Sistemi tehdit edebilecek kullanıcılar önceden belirlenerek muhtemel saldırı senaryolarının önlenmesi amaçlanmıştır. İleride yapılacak çalışmalar için, örnek bir çalışma olarak değerlendirilebilir. Devamında, bu çalışmanın yeni kıstaslar kullanarak genişletilmesi hedeflenmektedir.

KAYNAKLAR

[1] E.Sahinaslan, A.Kanturk etc., "Kurumlarda Log Yönetiminin Gerekliliği", Akademik Bilişim Konferansları, 2013.

[2] T.Ozseven, M.Dugenci, "Log Analiz: Erişim Kayıt Dosyaları Analiz Yazılımı ve GOP Üniversitesi Uygulaması", Bilişim Teknolojileri Dergisi, pp.55-66, 2011.

[3] T.Aye, " Web log cleaning for mining of web usage

patterns", 3rd International Conference on volüme 2, pp.490-494, 2011.

[4] Y.Zhang, X.Ye etc, "A practical database intrusion detection system framework", In Computer and Information Technology, vol.1, pp.342-347, 2009.

[5] C.Mohan, B.Linday and R.Obermarck, "Transaction management in the R distributed management system", ACM Transactions on Database Systems, vol.11, issue 4, pp.378-396, 1986.

[6] I.Ray, K.Belyaev atc., "Secure logging as a service-delagating log management to cloud", IEEE Journal Systems, vol.7, issue.2, pp.323-334, 2013.

[7] K.Kent, M.Souppaya, "Guide to Computer Security Log Management, National Institute of Standarts and Technology, 2006.

[8] C.Pfleeger, S.Pfleeger, "Security in Computing, 3rd Prentice Hall Professional Technical Reference, 2002.

[9] J.Andress, "The Basics of Information Security", Understanding the Fundamentals of InfoSec in Theory and Practice", 2014.

[10] M.Tekerek, "Bilgi Güvenliği Yönetimi", KSU Journal of Science Engineering, pp.132-137, 2008.

[11] E.Yildiz, "Gerçek Zamanlı Bir Saldırı Tespit Sistemi Tasarımı ve Gerçekleştirimi", Journal of New World Sciences, vol.5, no 2, pp.143-159, 2010.

[12] M.Ogun ve A.Kaya, "Siber Güvenliğin Milli Güvenlik Açısından Önemi ve Alınabilecek Tedbirler", Journal of Security Strategies, issue 18, pp.145-181, 2013.

[13] E.Bertino, E.Terzi etc., "Intrusion detection in RBAC-administered databases", 21st Annual in Computer Security Application Conference, pp.10- 182, 2005.

[14] C.Chung, Y.Gertz etc., "Demids:A missue detection system for database systems", Integrity and Internal Control in Information Systems, vol.37, pp.159-178, 2000.

[15] Y.Hu, B.Pand, "A data mining approach for database intrusion detection", Proceedings of the 2004 ACM symposium on Applied computing, pp.711-716, 2004.

[16] V.C.Lee, J.A.Stankovic, "Intrusion detection in real-time database systems via time signatures", Real-Time Technology and Applications Symposium (RTAS), pp.124-133, 2000.

[17] Bridges S.M, Vaughnn R.B, "Fuzzy Data Mining and Genetic Algorithms Applied to Intrusion Detection", 23rd National Information Systems Security Conference, 2000.

[18] Jemili F., Zaghdounal M. Etc, "Hybrid Intrusion Detection and Prediction multiAgent System, HIDPAS", International Journal of Computer Science and Information Security, vol.5, no.1, 2005.

[19] Ramasubramanian P. And Kannan A., "Multi-Agent based Quickprop Neural Network Short-term Forecasting Framework for Database Intrusion Prediction System", CiteSeerX, 2014.

[20] Romasubramanian P., Kannan A., "A genetic-algorithm based neural network short-term farecasting framework for database intrusion prediction system", Soft Computing, vol.10, issue 8, pp.699-714, 2006.

[21] Haslum K., Abrater A. Etc, "Disp: A framework for distributed intrusion prediction and prevention using hidden markov models and online fuzzy risk assasment", 3rd International Symposium on Information Assurance and Security, pp.183-190, 2007.

Çiğdem Bakır Sakarya Üniversite Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümünden 2010 yılında mezun oldu. 2013 yılında Yıldız Teknik Üniversitesi Bilgisayar Mühendisliği Bölümünde Araştırma Görevlisi olarak göreve başladı. 2014 yılında aynı üniversitenin Bilgisayar Mühendisliği Ana Bilim dalında Yüksek Lisansını tamamlayarak Doktora eğitimine başladı. İlgi alanları; veri madenciliği, bilgi güvenliği, biyomedikal görüntü ve işaret işleme konularıdır.

Veli Hakkoymaz lisans derecesini Hacettepe Üniversitesi Bilgisayar Bilimleri Mühendisliği Bölümünden 1987'de, yüksek lisans derecesini University of Pittsburgh (PA) Bilgisayar Bilimlerinden 1992'de, Doktora derecesini CWRU (OH) Bilgisayar Bilimleri Mühendisliğin'de 1997'de tamamladı. 2011'de Doçent ünvanını aldı. İlgi alanları; veritabanı yönetim sistemleri, bilgisayar mimarisi, işletim sistemleri ve dağıtık sistemlerdir. Halen Yıldız Teknik Üniversitesi Bilgisayar Mühendisliği Bölümünde Öğretim Üyesi olarak görevini sürdürmektedir.

APPLICATIONS AND DESIGN FOR A CLOUD OF VIRTUAL SENSORS

Ammar Jameel Hussein, Ammar Riadh, Mohammed Alsultan, and Abd Al-razak Tareq

Manuscript received August 23, 2015.
Ammar Jameel Hussein Al Bayati,
University of Technology, Baghdad, Iraq, ammar.jameel.ict@gmail.com
Ammar Riadh Kairaldeen,
Baghdad Unvercity, Baghdad, Iraq, eng_ammar81@yahoo.com
Mohammed Alsultan,
Çankaya University, Turkey, Ankara, mohammed.altaliby@gmail.com
Abd Al-razak Tareq Rahem,
University of Technology, Baghdad, Iraq, abdtareq@yahoo.com

Abstract — the use of sensors in our daily lives is a growing demand with the large number of electronic devices around us. These sensors will be included in our daily life requirements soon and they will affect our lives in both positive and negative ways. In this paper, we discuss the manner, applications and design issues for a cloud of virtual sensors, and we introduce a distributed system design to deal with physical sensors that reside in diverse locations and operate in different environments. This design operates in a cloud computing vision and can make virtual sensors in upper of physical one available from anywhere using ICT structure. Then, we negotiated the future of this technology, i.e., the Internet of Things (Io-T). Additionally, we go over the strengths and weaknesses of using this technology. Our test lab shows high performance and good total cost of ownership and effective response time.

Index Terms — Cloud Virtual Sensors, Internet of Things (Io-T), Sensor Cloud, Virtual Sensor, weir less sensor network

I. INTRODUCTION

PHYSICAL sensors are used all around the world in numerous applications [1]. For the most part, sensors are regularly used by their own applications since each application retrieves data entirely with the cooperation of physical sensors and their sensor statistics. Additionally, vendor requests cannot be customized to the physical sensors in a diverse event [2]. Perhaps this is one of the main reasons that give us a new concept of the need for virtual sensors [3] and a sensor cloud [4]. A sensor cloud can be defined as a collection of virtual sensors comprised of physical sensors. Consumers inevitably and dynamically can establish or deliver on the basis of application demands. This method has a number of advantages.

Firstly, this improves sensor administration capability. Consumers can use devices regarding their view of wireless sensor networks (WSN), typical tasks for a variety of factors include area of interest, security and latency.

Secondly, statistics attain by WSN can be public among many consumers, which can reduce the total cost of data gathering for both an organization and the customer.

A sensor-cloud virtualizes the physical sensor by way of putting them on the cloud dynamically, Grouping these sensors in virtual manner and putting them in cloud computing can be available on demand when other applications need them, and from this concept, a new term is found: “Internet of Things” (Io-T), which proposes the potential of assimilating the digital domain of the Internet with the physical domain in which we breathe [5]. In order to realize this proposal, we need to demand a systematic method for assimilating sensors, the operator and the information on which they operate on the Internet we see nowadays. In this paper, we will discuss the virtual sensor, sensor clouds and the Internet of Things. We will review issues and applications of a cloud of virtual sensors, introduce a design for a virtual cloud sensor and finally overview the pros and cons of this technology.

II. VIRTUAL SENSOR

A virtual sensor is the emulation of a physical sensor that obtains its own data from underlying physical sensors. Virtual sensors provide a customized view to users using distribution and location transparency [6]. Virtual sensors contain meta-data about the physical sensors. The required physical sensors should be dynamically organized in the following order: virtualization, standardization, automation, monitoring and grouping in the service model.

Implementation of virtual sensors is carried out in four different configurations: one-to-many, many-to-one, many-to-many, and derived configurations [7].

In the following parts, there are brief reviews of each structure:

1) One-to-Many Structure
This structure deals with one physical sensor link to several virtual sensors.

2) Many-to-One Structure
In this structure, the topographical areas are allocated into zones and each zone can have one or more physical sensors and sensor networks.

3) Many-to-Many Structure
This configuration is a combination of the one-to-many and many-to-one configurations. A physical sensor can correspond with many virtual sensors and also be a part of a network that provides aggregate data for a single virtual sensor.

4) Derived Structure
A derived configuration refers to a versatile configuration of virtual sensors derived from a combination of multiple physical sensors. In the derived configuration, the virtual sensor communicates with multiple sensor types while the virtual sensor communicates with the same type of physical sensor in the other three configurations. Fig. 1 shows the different structure schema.

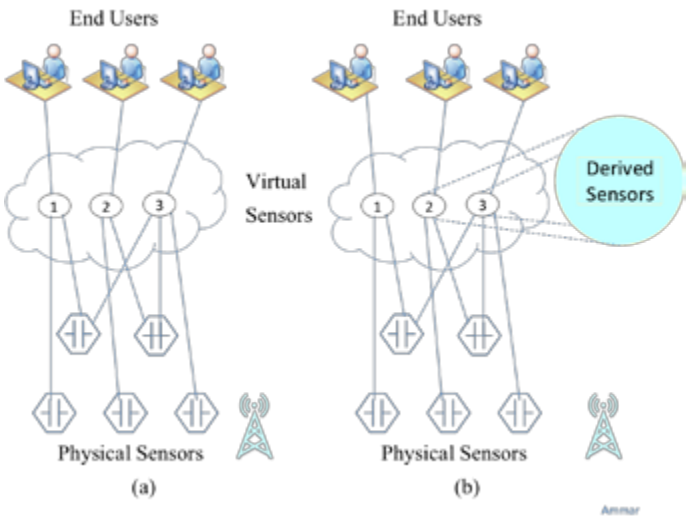


Fig. 1. Virtual Sensors Structure Schema

III. SENSOR CLOUD

A sensor cloud can be derived from the following definition: a structure that permits real universal calculation of data using virtual sensors as an edge among physical sensors using an Internet cloud network [8]. Statistics are calculated through servers to cluster infrastructure with the cyber network as the communication medium. These methods will enable consumers effortlessly to access, handle, visualize and evaluate, in addition to load, allocate and examine huge numeral data from a sensor. Data is gathered from more than a few types of applications, and this large sum of data are visualized by expending the IT and storage resources in cloud computing.

The idea of a virtual sensor cloud is a model that combines the idea of a virtual sensor and cloud-computing. Physical sensors (WSN) gather statistics and conduct whole sensor data into a cloud-computing frame. Cloud-sensors can grab sensor data resourcefully and use this data to monitor numerous applications. The cloud service structure is used to distribute the facilities of shared virtual network services in which consumers/end user's benefit by using these services. They are not worried about how they are detailed to implement the service. This is referred to as transparency and scalability.

IV. INTERNET OF THINGS (IO-T)

In order to access object or things from anywhere, it is a different idea from the concept of cloud sensors. Access these virtual sensors via the cloud service in our proposed design; it is called a cloud sensor. In fact, there is another concept that is nested within our subject, namely the "Internet of Things" (Io-T). It is stated that if objects, individuals or things provide an exclusive ID, they will have the capability robotically to transmit statistics through networks without needing a human or non-human/computer interface. Io-T has grown from the union of (WSN) technology and micro electromechanical systems (MEMS) by using the Internet [9][10].

Furthermore, the term "thing" in this sense may mean someone with an implant heart sensor, animals, plants, etc. This may refer to any component that has been integrated by sensors and making the driver be aware of changes in speed or any other expected measurement. Additionally, it may be any items with the capability of allocating IP addresses and delivering statistics through a network. Figure (2) symbolizes the "Io-T."

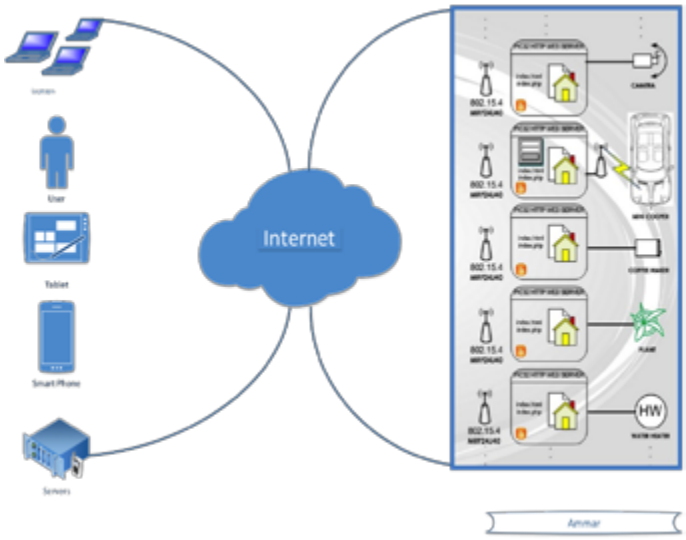


Fig. 2. Io-T

The idea behind the Internet of Things, is all about embedding microprocessors in objects, hence, they can communicate with each other. The information will lead us in the future to a new term called the Internet of everything.

V. SENSOR CLOUD APPLICATION

There are many applications that use the concept of cloud sensors. The four main categories include the following [11]:

1) Health Care
A cloud of virtual sensors can be used in the health care sector. In some new hospitals, physical and virtual sensor networks are commonly used to monitor patients' biological information, to switch drugs and to track and monitor patients and doctors within or outside a hospital.

2) Transportation Monitoring
A cloud of virtual sensors can be used also in transport monitoring systems by using basic administration systems such as traffic control, celestial navigation, car plate number deduction, emergency alarms, etc.

3) Military purposes
A cloud of virtual sensors can be used in many military applications such as following up friendly forces movement, action surveillance, exploration of enemy forces, determination of enemy pointing, war assessment and nuclea effects, anticipating and assessing biological and chemical attacks, etc.

4) Weather Prediction
The potential applications are very useful here to predict weather conditions and disasters such as tsunamis and earthquakes, volcanoes in addition to activity surveillance and expected effects, etc.

VI. RELATED WORK

Javier Miranda, et al [12], proposed a smart architecture that is based on smart-phones as a way to interact with people who are involved in Internet of Things applications. The new things in this paradigm are the consideration of interacting and the adaptively between peoples and smart things in every day live by context of internet of things, This is an important idea that extends the use of Internet of Things applications and makes them smarter in people's everyday life activities. Moreover, they discuss the socially related issues of the impact on people to accommodate this transformation, i.e., from real life to smart life. Finally, they design middleware architecture that depends on this discussion and considers People as a Service (PeaaS) [13] and Social Devices. This layer has many components, including an action repository, application repository, a device registry and an application manager. This model gives the user the ability to build a social profile on their own devices and share this profile with the middleware layer, thereby enabling the adaptive reaction between things. Some weaknesses in this project include discussing issues out of the scope of the technology framework and assuming end-user interference as a part of this model. Another study done by Sanjay Madria et al [14] proposed a new architecture for building a virtual sensor on top of the physical one. They discuss many components of this design. These architectures contain an intermediate layer between a sensor's device in the real world and consumers. The designed architecture includes three layers: a sensor-centric layer to deal with physical sensors; a middleware layer, intermediate layers; and a client-centric layer that handles the applications. In this design, it is not clearly shown how these layers can build a standard virtual sensor template on top of the physical one to handle different sensor types coming from different vendors and work using diverse technology. While Hoon-Ki Lee et al [15] proposed a new paradigm that enable the concept of the Social Web of Things (SoT), the paradigm was based on machine-to-machine talking in inspire the Web of Things. They implement a social sensor network that enables information associations in the context of web and social networks. The main component of this model includes the service domain, social relationships and user information. The main objective benefits of this model were finding a relationship between users, things and social networks and providing a dynamic service that has the ability to be reconfigured according to user needs and activities in the social network world. On the other hand, no security or privacy issues were discussed as a consequence of this wide sharing of information related to sensitive data, such as sensor networks.

VII. PROPOSED DESIGN

Our proposed design for a sensor cloud includes three main layers, each of which has a specific role and serves the up down layer. These layers can be classified thus:

1) Layer_1

This layer contracts with the preparation of the service template construction and provision standard definition in addition to defining the physical sensors as XML, web services or HTTP enabled. This will allow the service provider to access these sensors and develop them on several platforms without concern for the integration of a variance number of applications platforms.

2) Layer_2

This layer communicates with many groupings of physical sensors and attempts to place them into one classified group. In addition, this layer is the more important layer in our proposed design. The layer allows sensor service providers and other IT resources to be managed remotely without concern for the location of the real sensor sites. This layer can be considered the most important layer in our design, which includes servers, storage and networks devices. In this layer we use open source servers and applications and apply the concept of virtual servers to reduce the total cost of ownership.

3) Layer_3

This layer corresponds with consumers/end users and their applicable requests. Numerous consumers need to contact the valued data sensor from many kinds of operating system platforms using different types of application.

From the above, we can say that we have many types of actor (sensor owner, sensor-cloud administrator and end users) and many components in the cloud sensor (client, e-portal server, provision server and resources manager server, virtual sensor group, monitoring server and physical sensors). This proposes a schema which provides the transparency and scalability for end users to connect physical sensors. Fig. 3 demonstrates the three-layered structure while Fig. 4 shows actors and components in our proposed design.

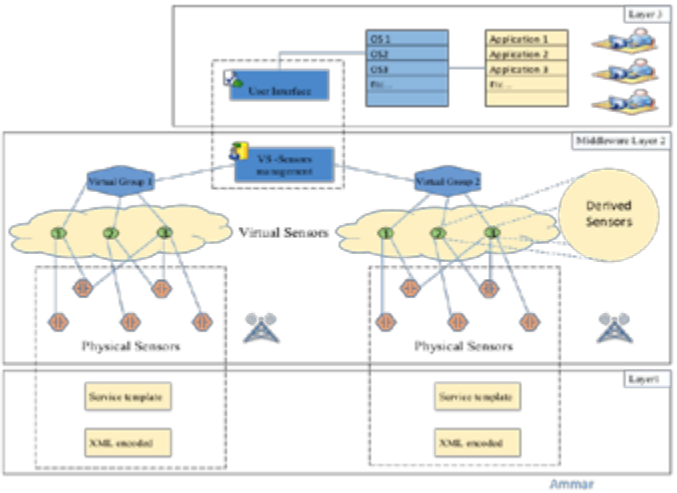


Fig. 3. Three Layered Structure

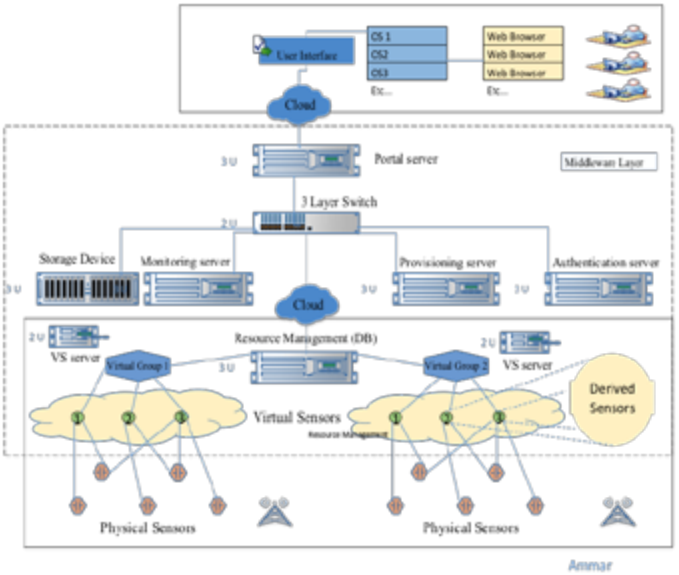


Fig. 4. Actors and Components in the Proposed Design

VIII. ISSUES IN THE SENSOR CLOUD DESIGN

There are many issues regarding the design of sensor clouds. Moreover, there are no modern concepts for applications and implementation from previous proposed structures. Therefore, to come out, there are many issues that should be considered while working with sensor cloud design, which includes but is not limited to cycle, as shown in Fig. 5:

Fig. 5. Design Issues Cycle

Sensor networks Security usually have several restrictions similar to other network types. Therefore, it is not logical to implement a conventional security policy such as the traditional security steps, consequently, to build a security operational platform for the sensor cloud, we need first to understand the nature of these restrictions on the form of the network. Some sensor network restrictions are, Unreliable Communication, Limited Resources and Unattended Operations.

IX. PROS AND CONS OF THE PROPOSED DESIGN

Following are some Pros. and Cons. Of our proposed design:

A. Pros

1. Transparency: The consumer does not need to worry about the details.
2. Scalability: The Sensor Cloud offers ease of management to the end consumer.
3. Reliability: The consumer can follow up the status of his own virtual sensors from anywhere.
4. Flexibility: The consumer can rapidly start to use the physical sensors by using virtual sensors remotely.
5. The consumer can make his group of sensors depend in his need by consuming virtual sensor groups.
6. The owner of the physical sensors can track the usage of the sensors.

B. Cons

1. ICT resources need for a sensor-cloud infrastructure should be well configured to serve this design purpose.
2. Each physical sensor needs templates for virtual sensors to be joined.
3. Bandwidth and connectivity types between the consumer and cloud-sensor server may be a factor of weakness.
4. The possibility of shearing data from some of the physical sensors gives the possibility of loss of precision data in real time.

X. LAB TEST

In our lab, we used one Windows Server 2012 and three Red Hat Linux servers to accomplish our proposed design. We also used Oracle Virtual Box as the virtual environment to host all our servers. Each virtual machine had 1 CPU 2.1HZ, Memory 2 GB and HD 15 GB. Our test lab showed high performance and a good total cost of ownership and effective response time. We applied a stress load (150,200 request and each user will run 100 threads simultaneously) to our design and gathered the results of system performance. Tables I and II show the static results obtained respectively.

Label	Samples	Av. (ms)	Min	Max	Std. Dev.	Through.	Kb/sec	Avg. Byt.
Login Request	150	1.044	32	2490	783.77	14.4	351.6	26424
Logout Request	150	0.738	34	2007	614.12	12.3	130.54	1164
HTTP Request	15000	0.158	4	2120	183.71	49.4	2499	47589

Table I - 150 Threads run 100 time

Label	Samples	Av. (ms)	Min	Max	Std. Dev.	Through.	Kb/sec	Avg. Byt.
Login Request	200	1.021	140	2516	347.6	18.6	422.45	26422.4
Logout Request	200	0.545	41	1553	339.11	17.3	19.04	1334
HTTP Request	20000	0.273	5	2254	322.1	56.5	2483.39	48489

Table II - 200 Threads run 100 time

XI. CONCLUSION

In this paper, we present a sensor cloud structure which enables the virtualization of physical sensors according to on-demand consumers' requirements without worrying about the details of how to implement virtual sensors. Our design provides transparency and flexibility to end users to host their own sensors. Moreover, our results show high system performance when applying the stress load test and the lowest total cost of ownership. On the other hand, using a communication line among the cloud sensor nodes is a formidable task, since the sensor cloud has many issues, such as security and integrity. Addressing these issues and attempting to develop them along with working in developing a new design of virtual environment will contribute to increasing the applications based on this type of sensor cloud architecture. Our proposed design is a big step towards the rapid progress of the new technology term "Internet of Things" which will be implemented in the future. Future work may focus on developing heterogeneous

distributed system designs and developing protocols to deal with physical sensors in standard ways, security issues for communication lines and allowing people to contribute to management design and allowing them to be part of the sensor cloud model by using their own sensors.

ACKNOWLEDGMENTS

The authors A. J. and A. T. thank the Iraqi Board of Supreme Audit Iraq/Baghdad (Government Body) which contributed effectively to give us this opportunity for publication. We would like to extend our sincere appreciation to Dr. Reza Hasnboor for his support in conducting the survey.

REFERENCES

[1] Akshay, N., Kumar, M.P., Harish, B., Dhanorkar, S., "An efficient approach for sensor deployments in wireless sensor network", IEEE Emerging Trends in Robotics and Communication Technologies (INTERACT), International Conference, pp 350 - 355, 2010.

[2] Madoka Yuriyama Takayuki Kushida "Sensor-Cloud Infrastructure - Physical Sensor Management with Virtualized Sensors on Cloud Computing", IBM Research – Tokyo, 13th International Conference on Network-Based Information Systems, pp 1-3, 2010.

[3] Thomas Franklin Litant, "The Fusion and Integration of Virtual Sensors", College of William and Mary, 2002.
[4] Pethuru Raj "Cloud Enterprise Architecture", CRC Press, pp. 312-320, 2012.

[5] Ovidiu Vermesan & Peter Friess, "Internet of Things - Global Technological and Societal Trends From Smart Environments and Spaces to Green ICT" , River Publishers, 2011

[6] Shashi Phoha, Thomas F. La Porta, Christopher Griffin, "Sensor Network Operations" John Wiley & Sons, pp 104-116, 2006.

[7] Sanjay Madria, Vimal Kumar, and Rashmi Dalvi, "Sensor Cloud: A Cloud of Virtual Sensors" IEEE next-generation mobile computing, Volume: 31, 2014.

[8] Yuriyama, M.; Kushida, T., "Sensor-Cloud Infrastructure - Physical Sensor Management with Virtualized Sensors on Cloud Computing", IEEE, Network-Based Information Systems (NBiS), 13th International Conference, pp. 1 – 8, 2010.

[9] Cuno Pfister "Getting Started with the Internet of Things: Connecting Sensors and Microcontrollers to the Cloud" O'Reilly Media, Inc., ISBN1449393578, 9781449393571, Ch4, 2011.

[10] Honbo Zhou "The Internet of Things in the Cloud: A Middleware Perspective" CRC Press, Mar 21, 2013.

[11] C.O. Rolim, F.L. Koch, C.B. Westphall, J.Werner, A. Fracalossi, G.S. Salvador,," A Cloud Computing Solution for Patient's Data Collection in Health Care Institutions", 2nd Intl Conference on eHealth, Telemedicine, and social medicine, 2010.

[12] Miranda, J., Makitalo, N., Garcia-Alonso, J., Berrocal, J.; Mikkonen, T., Canal, C., Murillo, and J.M.,, "From the Internet of Things to the Internet of People",IEEE Journals & Magazines Vol.: 19, Iss.: 2,pp. 40 – 47,2015.

[13] Guillen, J., Miranda, J., Berrocal, J., Garcia-Alonso, J.,Murillo, J.M., and Canal, C.,"People as a Service: A Mobile-centric Model for Providing Collective Sociological Profiles", IEEE Journals & Magazines, Vol.: 31, Iss.: 2,pp. 48 – 53,2014.
[14] Sanjay M., Vimal K., and Rashmi D.," Sensor Cloud: A Cloud of Virtual Sensors", IEEE Journals & Magazines Vol.: 31, Iss.2, pp. 70 – 77, 2014.

[15] Hoon-Ki L., Jong-Hyun J.,and Hyeon-Soo K., "Provision of the Social web of Things", Consumer Electronics ,Berlin (ICCE-Berlin), IEEE Fourth International Conference ,pp. 404 – 407,2014.

Ammar Jameel Hussein He received a MS in Computer Engineering at Çankaya University, Ankara, Turkey. His research areas are the design issues in the cloud of virtual sensors and Internet of Things (IoT). He is also working as a network administrator at the Iraqi Board of Supreme Audit. He received his B.Sc. and PGD in Computer Engineering at the University of Technology, Baghdad, Iraq. He is a member of the Federation of Arab engineers and was editor for DG Pioneer Magazine in Iraq (2005-2007).

Ammar Riadh Kairaldeen, Is currently lecturer Assistant in Baghdad University in Iraq,. He received the B.S Computer Engineering from Electric and Electronic Technical College, Baghdad, Iraq, in 2005, and the Master of Computer Engineering from Cankaya University, Ankara, Turkey, in 2014. His current research interests include Artificial Intelligent and Natural Language Processing.

Mohammed Alsultan, is currently a Ph.D candidate at Gaziantep University, he graduated as a M.Sc. in computer Eng. From Çankaya University. He received his B.Sc. degree in computer technology engineering from Mosul Technical College. His principal research interests lie in the field of the topology control of wireless sensor networks. He is also have research experience and interested in other fields such as data mining and Image processing & computer vision.

Abd Al-razak Tareq Rahem, is currently a Ph.D candidate at Universiti Kebangsaan Malaysia (UKM), the Department of Electrical, Electronics and Systems Engineering. He received the B.S Computer Engineering and Information Technology from University of Technology, Baghdad, Iraq, in 2002, and the Master of Technology degree in Information Technology college of Engineering from BVDU, Pune, India, in 2012. His current research interests include wireless networking and mobile ad hoc network. Routing Protocol. Network Performances. He was consulting in DG Pioneer Magazine in Iraq (2005-2008).

TEKNOLOJİNİN CASUSLUKTA KULLANILMASI VE KARŞI ÖNLEMLER

Samet OĞUZ, Eyüp Burak CEYHAN, Şeref SAĞIROĞLU

Özet — Günümüzde teknoloji ile beraber istihbarat ve casusluk sınır tanımaz hale gelmiştir. Kurum ve kuruluşların özellikle de devletlerin siber uzaya bağımlı hale geldiği günümüzde siber istihbarat ve siber casusluk önemini arttırmıştır. Bu çalışmada siber istihbarat/casusluk faaliyetlerinin nerelerde yoğunlaştığı ve nasıl kullanıldığına değinilerek, bu faaliyetlere karşı alınması gereken önlemlerin neler olduğu ortaya konmaya çalışılmış, özellikle de kişi ve kurumların yapması gerekenler ortaya konulmuştur.

Anahtar Kelimeler — Siber istihbarat, Siber Casusluk, Sosyal Mühendislik, Sosyal Ağlar, Casus Yazılımlar, Arama Motoru.

Abstract — Today, through the technology, intelligence and espionage does not recognize borders. The characteristics of institutions and organizations, especially states, have become dependent on the cyberspace therefore importance of cyber intelligence and cyber espionage has increased. In this study, where the focus of cyber intelligence/espionage activities and how to use these activities are mentioned. Also what measures need to be taken against these activities, particularly the measures to be taken by the individuals and organizations are presented.

Index Terms — Cyber intelligence, Cyber espionage, Social Engineering, Social Networks, Spyware, Search Engine

I. GİRİŞ

Faydalı bilgiler toplayıp, değerlendirmeler yaparak karar vericilerin yoluna ışık tutmak anlamına gelen istihbarat ve casusluk, aynı zamanda karar vericilerin tespit ettiği uygulamalar doğrultusunda psikolojik eylem ve propaganda gibi vasıtalarla toplumların algılarına yön vermek anlamını da ihtiva etmektedir. Ortalama olarak yarım asrı geride bırakan, askeri açıdan bir disiplin haline gelen istihbarat ve casusluğun yönetimi, hedefleri, çalışma metotları ve kullandığı araçlar teknolojinin değişim ve gelişimi ile beraber sürekli aşama kaydetmiştir [1].

Bilişim dünyasındaki gelişimin ivmesi, bize önümüzdeki zaman dilimlerinde siber istihbarat ve siber casusluk faaliyetlerinin ulusların güvenliğinin temel taşı haline geleceğini göstermektedir. Öyle ki günümüzde sanal hamleler sonucu çok fazla iş gücünün atıl bırakılması ile büyük kayıplara sebep olunabilmektedir. Bilişim sistemlerinin ve sanal âlemin toplumun bütün birimlerinde kullanılıyor olmasına ve birçok kamu kurumunda bu teknolojilerin yaygınlaşmasına paralel olarak, istihbarat birimleri de bu bilişim teknolojisinden yoğun olarak faydalanmaktadırlar. Ülkeler teknolojik ilerlemeler doğrultusunda uygulamalarını siber uzaya transfer etmek durumunda kalmışlardır. Bu durumun doğal sonucu olarak ülkelerin sahip oldukları tüm faydalı bilgiler, veriler siber uzayın bir parçasını oluşturmaktadır. Ne yazık ki bu eksende ülkelerin ulusal güvenlikle ilgili risk unsuru olabilecek bilgileri de sanal dünyada yer edinmektedir. Yarış içerisinde olan veya birbirlerine düşmanlık yapan devletler, değişik topluluklar bu bilgilere ulaşmak maksadıyla siber istihbarat ve casusluk çalışmaları yapmaktadırlar.

Bu çalışmada; siber istihbarat ve siber casusluğun ne olduğu ve hangi yöntemleri kullandığına değinilmiştir. Ayrıca kamu kurum ve kuruluşlarının gelişen teknolojiye bağlı olarak maruz kalabilecekleri siber casusluk faaliyetlerine nasıl önlemler alması gerektiği ortaya koyulmuştur.

II. SİBER İSTİHBARAT KAVRAMI VE YÖNTEMLERİ

A. Siber İstihbarat

İstihbarat, Arapça kökenli bir kelime olup; Türk Dil Kurumu tarafından yeni öğrenilen ve elde edilen bilgi olarak ifade edilmektedir [2]. İşlevsel olarak istihbarat olanak ve araçlar vasıtasıyla ulaşılması gereken bir konuda bilgi edinimi ve edinilen bilgilerin sadelikten çıkarılarak işlenmesi, anlamlandırılması ve çeşitli boyutlar kazandırılarak buradan bir sonuç çıkarılmasıyla ilgili işlevsel çabalar bütünüdür [3].

Arapça kökenli "tecessüs" kelimesinin karşılığı olan casusluk, bilgisi dışındaki işlemleri öğrenme arzusu gizli şeyleri merak etme anlamına gelmektedir. Casusluk faaliyeti ve espiyonaj kelimeleri aynı doğrultuda kullanılmak üzere yabancı dillerden alınarak casusluk manasında kullanılmaktadır [3].

Toplumların yaşamlarını devam ettirdikleri her yerde, her zaman varlığını sürdürmüş olan casusluk faaliyetleri dünyanın en eski iş sahalarından biridir. İstihbarat ve casusluk, yıllardan beri süre gelen devletlerin bekalarını sağlamak ve üstünlük yarışlarında ön sıralarda yer alabilmek için yürüttükleri faaliyetlerdir. Bu iki terim ehemmiyetlerini hiç kaybetmemiştir [4].

İnternet alt yapılı bilişim teknolojilerinin neredeyse bütün endüstri dallarında temel taşı oluşturduğu görülmektedir. İnternette veri iletimi için kullanılagelen e-postanın ardından, e-reklam, e-ticaret, e-devlet, e-oylama vb. pek çok terim yeni teknoloji ile birlikte her geçen gün daha fazla gündelik yaşamımızın birer parçası haline gelmektedir. Bunun yanında internet bağlantılı cep telefonları, diz üstü, kişisel ve tablet bilgisayarlar gibi üretilen her yeni sistem de bilgisayar tanımının sınırlarını zorlamaktadır [5].

Yeniliklerle büyüyen teknolojinin, ulusların güvenliğini ve özgürlüğünü doğrudan etkilediği, dünyada gücü tayin eden en önemli ilke haline geldiği çağımızda, teknolojik ve ekonomik istihbarat çok daha ehemmiyetli duruma gelmiştir. Casusluk faaliyetlerinin eski zamanlarda sadece askeri ve siyasi alanlara yönelirken, günümüzde telekomünikasyondan bilgisayar teknolojisine, taşımacılıktan tekstil endüstrisine, nano teknoloji ile optik alanındaki araştırmalara kadar her alanda etkisini hissettirerek dünyayı kocaman bir şehir haline getirmeye

devam etmektedir. Teknolojik yenileşmeler bilgi birikimini ve bilgiye ulaşmayı kolaylaştırırken aynı zamanda casusluk faaliyetlerini de kolaylaştırmaktadır [6].

Teknolojik yeniliklerin bilgiye ulaşımı kolaylaştırmasından ziyade asıl olarak istihbarat faaliyetlerini kolaylaştırdığı ortaya çıkmıştır . Hayatın olağan akışı içinde etkilerini çoğu zaman unuttuğumuz internet üzerinden kişilerin şahsi bilgilerinin, özel şirketlere ve kamu kurumlarına ait verilerin, kullanıcı hesaplarının ele geçirilerek kötü amaçlı kullanılması gibi olaylar sıkça yaşanmaktadır. İstihbaratın çalışma alanı; devletin kontrol görevini yerine getirebilmesi için, tehdit unsuru olabilecek konularda önem seviyesine göre karar mercilerine gerekli olan bilgi hakkında gerekli desteği sağlamak ve ayrıca propaganda, psikolojik eylemler gibi faaliyetler ile düşman istihbarat ve diğer faaliyetlerini engellemek olduğu dikkate alındığında, siber uzayda bu gayeyi taşıyan çalışmaların tamamı "siber istihbarat" olarak kavramsallaştırılabilir [1].

İstihbarat alanında ağırlıklı olarak bilişim teknolojisinin kullanımını kapsayan ve ayrıca içerisinde uzay araçları, uydular ve hava araçları ile icra edilen istihbarat faaliyeti olarak düşünülen siber istihbarat [7]; şahsi, sosyal, siyasal veya askeri avantaj sağlamak için, bilişim sistemlerine veya bilgisayarlara kanunsuz bir şeklide sızarak, kişilerden, rakiplerden, şirketlerden, devlet kurumlarından veya bankalardan, onların izni olmadan sırlarını elde etme eylemidir [2]. Ayrıca bilişim ortamındaki tehlike ve kötü amaçlı faaliyetlerin izlenmesi, değerlendirilmesi ve tedbirlerin alınması sürecidir. Siber istihbarat, bilişim teknolojisinin ve casusluk faaliyetlerinin birlikte kullanılarak savunmanın güçlendirilmiş halidir. Siber istihbarat faaliyetleri elektronik ortamda özellikle de sanal âlemde riskli verilerin siber teröristler tarafından ele geçirilmesini önlemek için en etkili yöntemleri kapsamaktadır [7].

Siber istihbaratı diğer istihbarat yöntemlerinden ayıran ve onlardan daha avantajlı konuma getiren şey kullanılan araçlardır. Bu materyaller ileri seviyede teknolojik ürünlerdir; faydalı ve nitelikli stratejik verilere sahip olurken, zaman ve ekonomik tasarruf sağlarlar. Bunun yanı sıra, bu araçlar daha az iş gücü ve sermaye ile daha kesin bilgi sağladığından dolayı alışılagelmiş istihbarat düşüncesini geliştirmektedir. Birçok uzmana göre siber istihbarat; düşmanın bizim hakkımızda bilgi sahibi olmasını engellerken, onun hakkındaki her şeyi öğrenmektir [5].

Bilişim sistemlerine karşı gerçekleştirilecek saldırılar, büyük tehditler ancak gelişmiş bir siber istihbarat alt yapısı ile fark edilip, önüne geçilebilir. Kişisel veya kurumsal güvenlik alanı ne kadar önemli olsa da ülke çapındaki güvenliğin önemi ile kıyaslandığı zaman kişilerin güvenliği bir parça daha geri planda kalmaktadır. Çünkü burada mevzu bahis olan, kişisel bir banka hesabı veya küçük çaplı bir bilişim sistemi değildir. Burada büyük bir devletin; istihbaratının, gizli servislerinin, sırlarının, ekonomisinin, vatandaş bilgilerinin, teknolojisinin vb. şeylerinin ele geçirilebilmesi gibi bir tehlike söz konusudur. Siber savaşların her geçen gün fazlalaşması, devletlerin bilinçli ve yoğun bir şekilde siber istihbarata yatırımda bulunmalarını gerektirmektedir [8].

B. Siber İstihbarat Yöntemleri

Sanal dünyada işe yarayan veya yaramayan birçok veri mevcuttur. Bu verileri ele geçirip istihbarat oluşturmak için çok çeşitli yöntemlerden yararlanılmaktadır. Bu yöntemleri Şekil 1'deki gibi sınıflandırmak mümkündür.

Şekil 1. Siber istihbarat yöntemleri.

Sosyal Mühendislik
(Sosyal Ağlara Dayalı Siber İstihbarat)

Güvenliğin en zayıf halkasının insan olduğu varsayımına dayanan sosyal mühendislik, ilk olarak insanların birbirleri ile olan iletişim ve ilişkilerini veya kişilerin farkında olmadan yaptıkları hataları kullanarak hedef kişi, kurum veya kuruluş hakkında bilgi toplamak olarak açıklanabilir. Sosyal mühendislik, bilgiye ulaşmak için kişilerden yararlanılması, etkileme ve ikna yöntemlerinin kullanılmasıdır. Sosyal mühendislik, normal şartlarda insanların tanımadıkları kimseler için yapmayı göze almayacakları şeylerin yapılabilirliğini arttırma becerisi olarak ifade edilebilir [5].

İnsan ilişkilerinden, dikkatsizliklerinden veya eğilimlerinden faydalanarak gizli bilgilere erişme çabası sosyal mühendislik olarak adlandırılmıştır [1]. Amaç; hedef alınan şahıs veya şirket yapısı, kurumsal ağ yapısı, iş verenler ya da işçilerin kendilerine ait bilgileri, şifreleri ve saldırıda kullanılabilecek her türlü materyalin toplanmasıdır [9]. Bu teknikte uzman analizci kişiler, sanal alemde paylaşılan herkesin kullanımına açık kişiye ait verileri, herhangi bir gizliliğe sahip olmamalarına rağmen, diğer kaynaklardan temin edilen başka bilgilerle uzmanlık gerektirecek bir biçimde birleştirmekte ve gizliliğe sahip veya önem arz eden bilgi haline getirmektedir [1].

Sosyal mühendisliğin faaliyet gösterdiği alanlardan biri de herkesin kullanabileceği karşılıklı iletişim imkânı sunan sosyal ağlardır. En çok kullanılan sosyal ağ sitelerinden biri olan Facebook'un farklı insanlar tanımak, kendi düşüncelerini anlatmak ve diğer insanların hangi fikir ve düşüncelere sahip olduklarını anlamak üzere üç temel fonksiyonu bulunmaktadır. Facebook sosyal ağ sitesinin global eksende bir milyara yakın internet kullanıcısı bulunmaktadır. Bu durum söz konusu sosyal ağ sitelerinin dünya çapında ne kadar yaygın olarak kullanıldığının ve ne kadar etkili olduğunun bir ispatıdır [10].

Sosyal mühendisliğe iyi bir örnek olması açısından Thomas Ryan tarafından yapılan Robin Sage testini gösterebiliriz. Robin Sage adını kullanarak Facebook, Twitter, LinkedIn

gibi sosyal ağlarda birçok profil oluşturulmuştur. 2 aylık test sonucunda hayali bir kişilik olmasına rağmen Google ve Lockheed Martin gibi firmalardan iş teklifi almış, erkeklerden akşam yemeği teklifi almış, FBI ve CIA hariç birçok istihbaratçı ve askeri personel kendisi ile arkadaş olmuş ve böylelikle bir çok ulaşılması zor olan bilgiye ulaşılmıştır [6].

Sosyal Ağlar

Kurum ve kuruluşlar, gruplar ve kişiler arasında eş zamanlı bilgi paylaşımı imkânı sağlayan, internet üzerinde birbirleriyle yaptığı diyaloglar ve paylaşımların bütünüdür [4]. İnternet blogları, internet günlükleri, video, resim gibi veri paylaşım siteleri, kişisel forumlar, arkadaşlık siteleri, haber paylaşım siteleri vb. internet hizmetleri sosyal ağ grubunda sayılabilir. Facebook, twitter, youtube, flickr, mylife, myspace, raptr ve linkedln örnek olarak verilebilir. Sosyal ağlar, istihbarat toplayan birimler ve istihbarat örgütleri için büyük fırsatlar sunmaktadır. Sosyal ağları kullanarak şahıslar, devletler, kamu kurumları hakkında bilgi sahibi olunabilir [5].

Wikileaks'i yapan Julian Assange, Russia Today'e verdiği demeçte Facebook'un kişilerin ad ve şahsi bilgileri hakkında büyük bir havuz olduğunu ve kullanıcıları tarafından isteyerek kullanılsa da, ABD istihbaratının kullanması için geliştirildiğini iddia etti. Facebook'u şu ana kadar yapılmış en korkutucu "casus makinesi" olarak adlandıran Assange, "Herkes şunu anlamalı ki, arkadaşlarını Facebook'a ekleyerek ABD istihbarat servisleri için bedavaya çalışıyorlar ve onlar için bu veritabanını oluşturuyorlar." demiştir. Assange'ın bu iddiası da sanal iletişim ortamları ile istihbarat örgütlerinin ne kadar çok iç içe geçtiğini göstermektedir. İstihbarat örgütleri sanal iletişim ağlarını kullanarak ulusal güvenliği tehdit edecek konuma gelebilmektedir. Günümüzde bunu en etkin kullanan örgütlerden biri de İŞİD terör örgütüdür. Bu siteler sayesinde bulduğu kişileri sosyal mühendislik metodunu kullanarak ikna etmekte ve kendisine militan yapmaktadır [5].

Rusya, sosyal ağlarda fotoğraf ve bilgi paylaşımı sağlayan (LiveJournal, vkontakte.ru vb.) paylaşım sitelerine askeri personelin üye olmasını yasaklamıştır. Sosyal ağlar istihbarat birimlerinin işine yaradığı gibi, suçluları yakalamada veya personel seçiminde de kullanılabilmektedir [4].

Casus Yazılımlar

Casus yazılım, kişilere ait önemli bilgilerin ve kişilerin yaptığı işlemlerin, kişilerin bilgisi dışında kopyalanmasını ve bu bilgilerin kendi çıkarları için kullanan kişilere transfer edilmesini sağlayan yazılım olarak ifade edilebilir [11]. Casus yazılımlar; Truva atı, rootkit, klavye dinleyici gibi, kullanıcıdan habersiz olarak bilgisayarlarda çalışan ve bilgisayarlardaki verileri belirli sunuculara gönderen yazılımlardır [4].

Casus yazılım veya spyware (İngilizce spy ve software sözcüklerinden türetilmiştir), başlıca kötücül yazılım (malware) türlerinden biridir. Geniş bant kullanan bilgisayarların yaklaşık olarak %90'ına yakınında casus yazılım bulunduğu düşünülmektedir. ABD'de Gartner Group tarafından Eylül 2004'de yapılan bir araştırmada 3 milyon işletme bilgisayarı gözden geçirilmiş ve bilgisayarlar üzerinde 83 milyon casus yazılım tespit edilmiştir [11].

Casus yazılımlar, hedef sisteme entegre olduktan sonra kendi kopyalarını oluşturmazlar. Casus yazılımın amacı önceden belirlenmiş bir sistem üzerinde gizli kalarak ulaşılmak istenen gizli bilgileri toplamaktır. Bu bilgi kimi zaman bir banka şifresi gibi önemli bir bilgi bile olabilir [11]. Bunun dışında, maddi amaç güden kuruluşlar internet üzerindeki kullanıcı alışkanlıklarını saptamak ve kullanıcıların ihtiyaç duyduğu emtiaları tespit ederek bu hizmet veya ürünlere ilişkin markaların reklamlarını kullanıcılara ulaştırmak gibi amaçlarla casus yazılımları internet üzerinde yayabilmektedirler.

Arama Motorları

Arama motoru, veri tabanında bulunan bilgileri aramak için kullanılan bir yazılımdır. Web robotu, arama indeksi ve kullanıcı arabirimi gibi üç bileşenden oluşmaktadır. Ancak arama sonuçları genellikle en çok tıklanan internet sayfalarından oluşan bir liste olarak belirlenmektedir [12]. Arama motorları birkaç yönden önemli istihbarat kaynağı olarak kullanılmaktadır.

Bunlardan bir tanesi, dünya üzerinde bulunan bütün sunuculardaki verileri depolamasıdır. Böylelikle her türlü bilgiye ulaşmış olmaktadır. Bu bilgiler arasından da veri madenciliği ile önemli bilgilere ulaşabilmektedir [4].

Bir diğer istihbarat elde etme yöntemi ise, arama motorunu kullanarak kimlerin neyi aradığı bilgisidir. Arama motoru firmaları (Google, yandex vb.) hangi ülkenin, hangi şehrinin, hangi kişilerin, hangi bilgileri aradığını bilmektedir. Örneğin; google firması hangi IP (internet protocol) adresinden hangi aramaların yapıldığını bilmektedir. Hangi IP adresini hangi şirketin veya devlet kurumunun kullandığını bulmak çok basittir. Bu şekilde hangi firmanın neleri araştırdığı veya hangi devlet kurumunun nelere ilgi duyduğu bilinebilir. Bu bilgiler genellikle ticari firmalar tarafından reklamcılık faaliyetleri için kullanılmaktadır [1].

III. SİBER İSTİHBARATA KARŞI KOYMA YÖNTEM VE TEKNİKLERİ

Siber istihbarat, siber saldırıların ve siber savaşın en önemli ve etkin unsuru yani olmazsa olmazıdır. İstihbaratsız savaş düşünülemeyeceği gibi siber istihbaratsız da siber saldırılar düşünülemez. Zaten incelendiği zaman bu iki terimin iç içe olduğu ve uygulama alanlarının ve karşı koyma yöntemlerinin neredeyse aynı olduğu ortaya çıkmaktadır. Siber istihbaratın önüne geçebilmek için öncelikle tam olarak siber güvenliğin sağlanması zarureti vardır.

Siber güvenlik tam olarak olgunlaşmamış bir disiplindir. Bu hususta güvenlik birimlerinin kabiliyetleri ve yetişmiş kaliteli kişi sayısı oldukça azdır. Buna sanal dünyada olan olaylara karşı gereken faaliyetlerin yapılmasının gerektirdiği uluslararası karakter de eklenince siber güvenlik ve siber savunma hususunda meydana getirilen oluşumların istenildiği kadar yeterli olmadığı anlaşılmaktadır. Siber ağların devletleri aşan sınırları düşünüldüğünde, siber güvenlik alanında faydalı önlem ve durumlar oluşturulabilmesi için uluslararası kurum ve kuruluşların icralarının ve devletlerin kendi aralarında oluşturdukları işbirliğinin önemi ortaya çıkmaktadır. Devletler açısından bakıldığında, internet üzerinden gelecek tehlikeler ve bunlara karşı uygulanacak önlemlerle ilgili farklı yöntemler

ve bakış açıları ortaya çıkmaktadır. Siber alemde kendilerine karşı yapılan saldırılara askeri karşılık verilmesi düşüncesine dayanan siber saldırıları savaş sebebi görebilecekleri gibi bir yaklaşımın yanında; siber uzaydan gelen tehditlerin aynı yerde karşılık bulması gerektiğini düşünen ve söyleyen yaklaşımlar da bulunmaktadır. Bu yaklaşımlar saldırıların nerden kaynaklandığı, ne amacı güttüğü ve orantılı güç kullanımı tartışmalarını da beraberinde getirmektedir [13].

Siber istihbarat alanında güvenlik çalışmalarının sonuç alabilmesi için alışılagelmiş tehditler ile siber tehditler arasındaki farklılıkları ortaya koyarak siber ortamın kendine has özelliklerine dikkat edilmesi gerekmektedir. Bu hususta ilgi çeken birinci nokta süreç içerisinde daha az bilgi birikimi ile daha karışık saldırıların gerçekleştirilebilir duruma gelmesidir. Üzerinde durulması gereken ikinci nokta ise siber saldırıları yapan şahısları ve bu saldırıların yapıldığı mekânları bulmaktır. Alışılagelmiş tehditlerin ve bu tehditlerin yapıldığı mekanların bulunması günümüzde göründüğü kadar zor değildir. Bu konuda dikkat edilmesi gereken üçüncü nokta ise siber saldırıların menzilinin ve gücünün artmasıdır. Alışılagelmiş saldırı araçlarının etkinlik alanının belli bir sığası, mesafesi vardır ve araçlar ancak bu mesafe içinde bir tehlike oluşturabilmektedir. Oysaki siber saldırı araçları günümüzde çok cüz'i bilgi ve para ile geliştirilebilmekte ve internet üzerinden dünyanın herhangi bir noktasına bu araçlar kullanılarak siber saldırılar ve siber casusluk faaliyetleri gerçekleştirilebilmektedir [14]. Bu saldırıları ve casusluk faaliyetlerini engelemek için; sistem güvenliğinin arttırılması, askeri ve sivil doktrinin geliştirilmesi, bunlarla ilgili cezaların belirlenmesi ve siber silahların ve sistemlerin uluslar arası düzeyde olacak şekilde kullanımının sınırlanması gerekmektedir [15]. Tabi ki bu önlemleri almak sadece tehlikeleri azaltmaktadır. Dünya'da mevcut olan veya örnek teşkil edebilecek sistemleri de incelememiz gerekmektedir. Örneğin Couldron adındaki yazılım gibi, sistemleri önceden denetleyen, açıkları bulan ve sonra da analiz yaparak bizlere alınması gereken önlemleri gösteren programlara sahip olmamız gerekmektedir. Fakat unutulmaması gereken önemli bir nokta ise bunların milli olması gerekliliğidir [16].

Dünya geneline bakıldığında, siber istihbarat ve siber savunma faaliyetlerine özel kurum ve kuruluşlardan ziyade en çok devletlerin orduları ve güvenlik güçleri tarafından başvurulduğu gözlemlenmektedir. Özellikle de çağı yakalamak amacıyla siber istihbarat faaliyetlerini aktif olarak kullanmaktadırlar.

Nisan 2015'de ABD Güvenlik Sekreteri tarafından açıklanan Yeni Güvenlik Stratejisi'nin (The DoD Cyber Strategy) bilgi paylaşımı ve kurumlar arası koordinasyonu, özel sektör ve müttefikler arasında gerekli irtibatların oluşturulmasını, koalisyon ve andlaşmaların yapılmasını içerdiği ve ayrıca David Kaye'nin (BM Özel Raportörü) 2015'de İnsan Hakları Komisyonu'nda ifade ettiği gibi dijital haberleşmede şifrelemenin kullanılmasının pozitif etkiye sahip olduğu bununla beraber internet güvenliği, bireysel gizlilik, özgür düşünce ve ekonomik büyümeye de etki edebildiği ortaya konulmuştur [17]. Hem ABD'nin Yeni Güvenlik Stratejisi'ne hem de David Kaye'nin açıklamalarına dikkat etmek ve bizlerin de aynı hassasiyetle faaliyetlerimizi (siber güvenlik alanında) yapıp yapmadığımızı tekrar gözden geçirmemiz gerekmektedir.

Dünyada olup biten bu siber casusluk olayları karşısında, siber casusluk ve siber saldırılara karşı koymak ve gücüne güç katmak amacıyla Türk Silahlı Kuvvetleri (TSK) de kendi bünyesinde gerekli önlemleri almaya başlamıştır. TSK yalnız kendi bünyesinde siber güvenlikle ilgili çalışmalar ve faaliyetler gerçekleştirmekle kalmayıp, diğer kamu kurum ve kuruluşları ile siber güvenlik alanında işbirliği yapmaya da başlamıştır. TÜBİTAK UEKAE bünyesinde 2001 yılında Bilişim Sistemleri Güvenliği Bölümü kurulmuştur. Ayrıca ilk çalışmalar TSK bünyesinde yapılmıştır. 2012 yılında TSK bünyesinde Siber Savunma Merkez Başkanlığı kurulmuş ve bu başkanlık diğer kurum ve kuruluşlar ile de koordineli olarak faaliyet göstermektedir. Siber güvenlik alanında yapılan çalışmalar planlanıp gerçekleştirilen sunumlar bu alanda tecrübe paylaşımı, yeni bakış açıları kazanma, ortak bilinç oluşturulması ve iş birliği konusunda çok özel bir yere sahiptir. Yapılmakta olan siber güvenlik tatbikatları kâğıt üzerinde sağlam görünen sistemlerin eksikliklerinin meydana çıkması ve böylelikle gereken önlemlerin alınması açısından çok elzemdir [13].

Özel veya kamu ayrımı yapılmaksızın birçok sektörde kurumların, bilişim teknolojilerine olan bağımlılığı ve ihtiyacı artmasıyla birlikte siber alanda yaşanan tehlikeler de artmaktadır. Siber casuslukta en etkili yöntemlerden biri olarak kullanılan sosyal mühendislik gibi en zayıf halka olan insandan kaynaklanan saldırıların riskini minimize etmek için personele siber güvenlik ile alakalı konularda daha fazla eğitimler verilip bilgilendirmeler yapılması gerekmektedir. Siber saldırıları engellemek için kurum ve kuruluşların dikkat etmesi gereken 10 husus aşağıdaki şekilde özetlenmektedir [18]:

1. Mobil cihazların kullanımında gerekli hassasiyet gösterilmeli: Kurum çalışanlarının her yerden kurum bilgilerine ulaşabilmesi ve herhangi bir önlemin alınmamış olması gözden kaçırılmamalıdır.
2. Kurum içi güvenlik politikası oluşturulmalı: Kurumdaki herkesin her yere erişim izninin olmaması ve yetkilendirilmelerin olması.
3. Sorumluluklar belirlenmeli: Herkesin sorumluluk alanlarının belirlenmiş olması ve işe alınmadan personele sorumlu olacağı alanların neler olacağı hakkında bilgi verilmesi.
4. Çalışanlara eğitim: Kurum çalışanlarına düzenli bir şekilde eğitim verilmeli.
5. IT ekibine eğitim: IT ekibinin de eğitimi aksatılmadan icra edilmeli ve sistemleri nasıl kullanacakları yönünde uzman olmaları sağlanmalı.
6. Güçlü şifreler kullanılmalı: Sistemde güçlü şifrelerin kullanılması sağlanmalı,alfa numerik ve üç ayda bir değiştirilen şifrelerin kullanılması sağlanmalı.
7. Envanter raporu tutulmalı: Envanter raporu düzenli olarak tutulmalı.
8. Yedekleme yapılmalı: Bilgi ve yazılımların yedeklenmesi yapılmalı ve ayrıca yedekler test edilmelidir.
9. İş sürekliliği yönetimi gerekli: Kurumun karşılaşabileceği riskleri, önemli iş süreçleri ile ilgili varlıklarını, bilgi güvenliği zafiyetlerinden dolayı oluşabilecek zararları, ekstra önlemlerin belirlenmesi ve icrasını, bilgi güvenliğini de kapsayan iş sürekliliği planlarının kararlaştırıldığı konuları içermelidir.
10. Güvenlik yazılımı olmalı: Güncel ve aktif olan güvenlik yazılımlarının kullanılması gerekmektedir.

IV. SONUÇ

Yaşadığımız çağ itibariyle ulusal güvenliğin sağlanması çok geniş çerçeveli olarak ele alınmalı ve çalışmalar bu doğrultuda gerçekleştirilmelidir. Sadece sınır güvenliğinin korunması için çaba gösterilmesinin günümüzde ulusal güvenliğin sağlanmasına yetmeyeceği aşikârdır. Teknolojinin eriştiği aşamalar ve teknolojik araçların öncelikle kamu kurumları olmak üzere hayatın her alanında ve aşamasında yer alması, siber güvenliği ve bu doğrultuda siber istihbaratı, ulusal güvenliğin ehemmiyeti yüksek olmazsa olmazı yapmıştır.

Gelecek günlerde ülkeler arasındaki savaşların sonuçlarını klasik cephelerdeki güç yerine daha karmaşık bir etki oluşturan ve savaşa yeni boyutlar kazandıran siber uzayda yaşanan muharebeler belirleyecektir. Siber yetenekler sayesinde, teknolojiye bağımlılığı her geçen gün artan ülkelerin teknolojik alt yapıları önemli hedefler haline gelecektir. Devletlerin ülkelerini koruyabilmek için, teknoloji ile donatılmış her türlü askeri imkanlarının yanında onlar kadar önemli siber saldırı ve siber istihbarat yeteneklerini de geliştirmeleri gerekecektir. Son zamanlarda dünyanın farklı yerlerinde meydana gelen siber saldırılar, devletlerin bu alanda kendi kabiliyetlerini geliştirmelerinin ve teşkilatlanmalarının zaruretini gözler önüne sermiştir. Bu doğrultuda siber tehlikelerin gün geçtikçe şekil değiştirerek kendini geliştirilmesi ve yeni tehdit türlerinin ne olduğunun tam olarak belirlenememesi, siber savunma alanında alınacak karşı önlemleri çıkmaza sokmaktadır. Tüm bu yenilikler, geleceğin muharebe alanında siber savaşların ehemmiyetini göstermekte ve siber savaşların dünya üzerindeki bütün devletlerin özelliklede bilişim teknolojisini her alanda kullanan devletlerin ulusal ve uluslar arası güvenlikleri için tehlike arz ettiğini gün yüzüne çıkarmaktadır.

Devletler ve kurumlar kendi güvenliklerine yönelik siber tehditlerle mücadele ederken, yeni teknolojik gelişmişlik seviyelerine ulaşabilmek açısından siber istihbarat karar vericiler için kilit rol oynamaktadır. Siber istihbarat sayesinde ülkeler hedef olabilecekleri siber tehditlere ve siber casusluk faaliyetlerine karşı gerekli önlemleri alabileceklerdir. Ayrıca devletler ve kurumlar, siber istihbarat sayesinde gelecekte karşılaşabilecekleri saldırı-casusluk-bilgi hırsızlığı gibi faaliyetlere karşı önceden önlem alabileceklerdir. Siber savaşların muhakkak olacağı gelecekte, devletlerin ve kurumların bünyesinde bulunan karar vericiler, siber istihbaratın önemini iyi benimsemeli ve bünyelerinde bulundurdukları istihbarat veya bilgi edinmek için kullandıkları teşkilatları, siber istihbaratı aktif bir biçimde kullanabilecek düzeye getirmelidirler. Bu teşkilatlar kanunların kendilerine verdiği yetki çerçevesinde yeri geldiği her alanda siber istihbarat faaliyetlerini aktif veya pasif bir şekilde kullanabilmelidirler.

Siber saldırılar ve siber casusluk konusundaki gelişmeler ve yöntemler çok hızlı bir şekilde değişmektedir. Siber saldırılar ve bu çerçevede siber casusluk faaliyetleri konusunda yasal düzenlemeler ihtiyaçlar doğrultusunda değiştirilmelidir. Siber savaş alanlarının olmazsa olmazı siber istihbarat yapılanmasının tek bir elden yönetilmesi gerekmektedir. Ayrıca kritik altyapı sistemlerinde kullanılan yazılım ve donanımlar mümkün olduğu kadar milli olmalıdır. Kritik altyapı sistemlerinde kullanılmak üzere yurt dışından alınan yazılım ve donanımlar beraberinde riskleri de getirmektedir. Özellikle yazılımların içerisine yerleştirilme ihtimali bulunan gizli kodlar, arka kapılar vb. yazılımlar, sistemi siber saldırılara açık hale getirmektedir. Bu nedenle, özellikle gelişmiş ülkeler kamu kurumlarındaki bilgisayarlarında kendi işletim sistemlerini kullanmaktadır. Kamu hizmeti ağları ve internet arasında bağlantı olmaması veya sınırlanması gerekmektedir. Özellikle kritik altyapıların güvenliğinden sorumlu personelin eğitimine gereken önem verilmeli ve bu doğrultuda yeterli mali kaynak ayrılmalıdır. Konu, insan ve olumsuz yönleri, düşkünlükleri ve zayıflıkları olunca çok dikkatli olunması gerekmektedir. Siber istihbarat ve casusluk bakımından en önemli unsur olan insanların, zayıflıklarını tamamen ortadan kaldırmak mümkün görünmediğine göre, siber istihbaratta ve casusluğa karşı başarıyı elde etmenin yolu insan etkeninin çok iyi değerlendirilmesinden yani eğitiminin en üst seviyeye çıkarılması ve durumsal farkındalığının arttırılmasından geçmektedir. Bu çalışmada açıklanan siber istihbarat yöntemlerine baktığımızda hepsinin insan odaklı olduğunu söyleyebiliriz. Ayrıca bir zincir halkasının kuvveti en zayıf halkası kadardır özdeyişinden yola çıkarak; kurumların insanlardan oluştuğu ve gerekli eğitim-farkındalık seviyesine çıkarılmaları gerektiği akıldan çıkarılmamalıdır. Eğitimli, nitelikli ve farkındalık seviyesi yüksek olan personelin her zaman anahtar role sahip olduğu ve olacağı dikkate alınarak insan faktörü üzerine odaklanmalı ve gerekli çalışmalar yapılmalıdır.

KAYNAKÇA

[1] Bayraktar, G. "Harbin Beşinci Boyutunun Yeni Gereksinimi: Siber İstihbarat", Güvenlik Stratejileri Dergisi,120-135, 2014.

[2] "İstihbarat Nedir?", www.tdk.gov.tr , Erişim Tarihi: 30 Mart 2015.

[3] Gültekin, A. "İstihbarat Teknikleri", Timaş Yayınları, 32-36, 2004.

[4] Çifci, H. "Her Yönüyle Siber Savaş", TÜBİTAK Popüler Bilim Kitapları, 289-302, 2013.

[5] Özçoban, C. "21.Yüzyılda Ulusal Güvenliğin Sağlanmasında Siber İstihbaratın Rolü", Harp Akademileri Stratejik Araştırmalar Enstitüsü Yüksek Lisans Tezi, 75-84, 2014.

[6] Yayla,M. "Hukuki Bir Terim Olarak Siber Savaş", TBB Dergisi, 104 : 194-198, 2013.

[7] Çetinkaya, Ş. "Siber Terör ve Siber İstihbarat", http://www.21yyte. org/tr/ arastirma /terorizm-ve-terorizmle-mucadele /2011 /09/23 /6 309 /siber-teror-ve-siber-istihbarat, Erişim Tarihi: 30 Mart 2015

[8] Akçadağ,E. "Sürekli Artan Önemi Işığında Siber Güvenlik", http://www.bilgesam.org/incele/1207/-surekli-artan-onemi-isiginda-siber-guvenlik/#.VUJy7pWJiP8, Erişim Tarihi: 08 Nisan 2015

[9] Şahin,M.Y. "Karşı istihbaratta insan boyutunun irdelenmesi:Gafil muhbirlik örneği", Harp Akademileri Stratejik Araştırmalar Enstitüsü Yüksek Lisan Tezi, 48-70, 2013.

[10] Yılmaz, S. "Batı İstihbaratı ve Sosyal Medya", http://usam.aydin.edu.tr/ analiz /guvenlik_isthbrt..pdf, Erişim Tarihi: 08 Nisan 2015.

[11] "Casus Yazılım", http://tr .wikipedia.org/wiki/Casus_yaz %C4%B1l%C4%B1m, Erişim Tarihi: 10 Nisan 2015.

[12] Yurdakul,N.,.Bat,M. "Şirketler İçin Rekabette Sanal Farkındalık: Arama Motoru Pazarlaması", Gümüşhane Üniversitesi İletişim Fakültesi Elektronik Dergisi, 45-60, 2011.

[13] Kaya,A., Öğün, M. "Siber Güvenliğin Milli Güvenlik Açısından Önemi ve Alınabilecek Tedbirler" Güvenlik Stratejileri Dergisi, 158-170, 2013.

[14] Canbay,C., Ünver,M. "Ulusal ve Uluslararası Boyutta Siber Güvenlik", Elektrik Mühendisliği Dergisi.(438), 94-103, 2010.

[15] Geers,K. "Strategic Cyber Security" CCD COE Publication, 132-139, 2011.

[16] Jajodia, S., Noel, S., Kalapa, P., Albanese, M., Williams, J. "Cauldron Mission-Centric Cyber Situational Awareness with Defense in Depth", IEEE Military Communications Conference (MILCOM), 1339-1344, 2011.

[17] Serrano, B. "Cyber Security and Cyber Espionage in International Relations", http://diplomacydata.com/cyber-security-and-cyber-espionage-in-international-relations/, Erişim. Tarihi: 23.09.2015.

[18] Ağaç,F. "Siber Güvenliğin Anahtarı Ulusal Çözümler", Bilişim Dergisi, 173, 88-91, 2015.

SOSYAL AĞLARDA GÜVENLİK FARKINDALIĞININ ARTTIRILMASI

Ebru Yeniman Yıldırım

Özet — Bilişim teknolojilerindeki gelişmeler, bilişim teknolojilerinin kullanımıyla hizmetlerin daha hızlı sunulması, yaygınlaştırılması, doğru ve yeterli bilgiye hızla ulaşma, iş ve zaman verimliliği gibi pek çok kolaylığı sunarken sanal ortamlarda güven sorununu da beraberinde getirmiştir. Son yıllarda sosyal ağlarda yaşanan olumsuzluklar, bu ortamları kullanmanın sosyalleşmeyi kolaylaştırdığı kadar, yeni tehdit ve tehlikeleri de beraberinde getirmiştir.

Bu çalışmada, sosyal ağ ortamlarında oluşabilecek riskler ve tehditler vurgulanarak, bilgi güvenliğinin özel hayatın gizliliği hakkına göre alınması gereken önlemler ve dikkat edilmesi gereken hususlar açıklanmaktadır. Bu kapsamda Bursa ilinde yaşayan rastgele seçilen 234 katılımcıya sosyal ağlarda güvenlik konusunda anket uygulaması yapılmıştır. Sosyal ağlarda güvenlikle ilgili elde edilen sonuçlar değerlendirilerek, önerilerde bulunulmuştur.

Anahtar Kelimeler — Sosyal Ağlar, Bilgi Güvenliği, Tehditler, Açıklar, Güvenlik Farkındalığı

Abstract — Evolutions in communication technology give rise to security problems in virtual media along with itself as well as providing many facilities, such as job opportunities, generalizing them, attaining the proper and sufficient information fast, work and time efficiency. The problems people have experienced lately show that using this social media may lead to some new threats and risks along with itself as well as making it easy to be socialized.

The required measures and points to take into account are explained in this study according to the right of privacy in the field of information security by emphasizing the potential risks and threats. In this context, 234 random participants, who were living in Bursa, were chosen for the survey regarding security in social networks. Based on the assessment of the outcomes of the survey on security in social networks, recommendations were provided.

Index Terms — Social Networks, Information Security, Threats, Shortfalls, Security Awareness

I. GİRİŞ

Sosyal ağlar son yıllarda ülkemizde yoğun bir şekilde kullanılmaktadır. Türkiye İstatistik Kurumu'nun (TÜİK) 2014 Ocak-Mart ayları arasında yaptığı araştırmaya göre internete erişim sağlayabilen 41.2 milyon kişinin %78'i Sosyal ağlara bağlanmıştır. 2015 yılının ilk üç ayında internet kullanan bireylerin %80,9'u sosyal medya üzerinde profil oluşturma, mesaj gönderme veya fotoğraf vb. içerik paylaşmıştır [1].
Sosyal ağlar, ister web ortamında, isterse mobil platformlarda olsun genelde aynı hizmetleri sunmalarına rağmen her bir erişim platformunda kullanıcılar açısından farklı güvenlik ve gizlilik tehlikeleri doğurmaktadır. Özellikle bu ortamları kullanan kişilerin bilgi güvenliği farkındalığının düşük olması, etik kullanım konusunda bilinç seviyesinin yetersizliği; bilinçsiz kullanmanın getireceği olumsuzluklarla karşılaşılabilecek tehlikelerin farkında olunmaması bu konuya önem verilmesi gerektiğini göstermektedir.
Bu kapsamda Bursa ilinde yaşayan rastgele seçilen 234 katılımcıya Sosyal ağlarda güvenlik konusunda uygulanan ankette çarpıcı sonuçlar elde edilmiştir. Anketten elde edilen sonuçlara göre, toplumda ciddi anlamda kullanılan sosyal ağlar konusunda bilgi güvenliği farkındalığının yetersiz olduğu ve önlemler alınması gerektiği vurgulanmıştır.

II. SOSYAL AĞLAR

Sosyal ağlar, kullanıcıların birbirleriyle tanışması, gruplar oluşturması, birbirleriyle irtibata geçmesi, tartışma ortamı oluşturularak, içerik paylaşımında bulunulması ve ortak ilgi alanlarındaki kişilerin bir araya gelebileceği internet siteleri olarak tanımlanmaktadır [2].

Dünyada kullanılan Sosyal ağlar; Facebook, Whatsapp, Twitter, LinkedIn, Instagram, Pinterest, Foursquare, Google+, Skype, Flickr. vb. olarak farklı alanlarda hizmet vermektedir. Bu tür siteler genellikle kullanıcı profilleri, kişisel bilgiler, yerleşim yeri, çalışma yeri, aile bilgileri, üyelikleri, alışkanlıkları ve hobileri gibi pek çok detay bilginin profilden takip edilebildiği veri paylaşım siteleridir. Bu kadar geniş kapsamlı paylaşımlar, son yıllarda saldırganların kaçırmayacağı, güvenliğin tehlikeye girdiği bir sosyal mühendislik alanı haline gelmiştir.

Sosyal ağları kullanıcı sayısı açısından ele aldığımızda en fazla nüfusu olan ülkeler kadar üye sayılarının olduğu görülmektedir. Bu nedenle sosyal ağlar siber mühendislerin, hackerların, kullanıcı verisi toplamak amacıyla şirketlerin hedefi durumuna gelmiştir. Spamciler, sosyal ağ sitelerinden bilgi toplamak için uğraşmakta ve fırsat beklemektedirler [3].

III. SOSYAL AĞLARDA GÜVENLİK RİSKLERİ

Sosyal ağlar, bilinmeyen veya çokta farkında olunmayan pek çok yeni tehlikeleri üzerinde barındıran paylaşım siteleridir. Bu yüzden sosyal ağlarda güvenlik önlemlerinin alınması önem arz etmektedir.

Sosyal ağlardaki güvenlik açıklıklarının temel nedenleri; bu ağların kuruluş amaçları nedeniyle, mahremiyet ilkelerine uyulmaması, ortamın yönetiminin ve kontrolünün nasıl yapıldığını kullanıcıların tam olarak bilmemesi veya kavramaması ve en önemlisi kullanıcıların kişisel bilgilerini paylaşarak kendilerini bu ortamda hedef haline getirmeleridir [4].

eBizMBA Eylül 2015 verilerine göre en sık kullanılan sosyal ağ sitelerine ve aylık ziyaret sayılarına bakıldığında Facebook'un 900 milyon aylık tekil kullanıcı sayısı ile birinci sırada olduğu, Twitter'ın 310 milyon ile ikinci, LinkedIn'in 255 milyon aylık tekil kullanıcı sayısı ile üçüncü sırada olduğu görülmektedir [5]. Bu yüzden kullanıcıların sosyal ağları günlük internet kullanımında vazgeçilmez bir alışkanlık haline getirmesi, onları çok yönlü tehlike ve tehditlere maruz bırakmıştır.

Sosyal ağlarda kullanıcıların adına sahte hesaplar açılmakta, kimlik taklidi yapılarak; hesapları ele geçirilen kişilerin tüm bilgilerine erişilebilmektedir. Kimlik hırsızlığı, bir kişinin

kimlik bilgilerine erişmek ve bu bilgileri sosyal ağda kendi menfaati için kullanmak demektir [5].

Spam, bir liste veya grup e-posta adresine gönderilen genelde reklam içerikli, istenmeyen e-posta anlamına gelir. Aynı şekilde bir saldırgan bu e-postaları bir sosyal ağ aracılığıyla kullanıcılara gönderip, gönderdiği kişinin kullanıcı bilgilerini elde etmeye çalışabilir [6].

Yeni bir sosyal ağa üye olunduğunda; bu ağdaki diğer kişileri bulmak üzere e-posta hesap ve parola bilgilerini girmeniz istenebilir. Bu sayede elde edilebilecek olan e-posta adresleri, gerçek kişileri beyan eden reklam firmalarına satılabilir. Üye olunan sosyal ağ sitesinin tüm e-posta haberleşmenizi tarayabileceği de unutulmamalıdır [7].

Sosyal ağ sitelerinde kullanıcılar evlilik durumlarını, eğitimlerini, adreslerini, kişisel bilgilerini, kişisel resimler gibi önemli bilgilerini paylaşmakta, hatta nerede çalıştıklarını, önceki tüm eğitimlerini, politik görüşlerini ve ilgi alanlarını da paylaşmaktadırlar [8]. Anne kızlık soyadı da pek çok alanda kullanılan gizlilik bilgisidir ancak, bu ortamlara kullanıcının anne ve dayısının dahi katılması bu bilgilerin biliniyor olmasına neden olacaktır [4].

IV. SOSYAL AĞLARDA ALINMASI GEREKEN GÜVENLİK ÖNLEMLERİ

Sosyal ağların kullanımının hayatımızda her geçen gün giderek artması, güvenlik konusunda ciddi önlemler alınmasını gerektirmektedir. Sosyal ağların tamamında kullanım koşullarının teyit edilerek, gizlilik ayarlarının yapılması ve gerekli güvenlik önlemlerinin alınması gerekir. Sosyal paylaşım ağları bilgi ve bilgisayar güvenliği açısından değerlendirildiğinde; kullanılırken sorumluluk isteyen, konuyla ilgili bilgi birikimi gerektiren, belirli bir kullanıcı bilincine ve disiplinine sahip kişiler tarafından kullanılması gereken, iletişim ve paylaşım ortamlarıdır. Doğru kullanılmadıkları takdirde, kişisel bilgilerin çalınması, istenilmeyen durumlarla karşılaşılması, beklenilmeyen tehdit ve tehlikelere maruz kalınması ve en önemlisi kişisel bilgilerin mahremiyetine zarar verebilecek pek çok olumsuzlukları içinde barındıran ortamlar olabileceği unutulmamalıdır [4].

Cisco'nun 2013 yılı için Yıllık Güvenlik Raporuna göre online siteler arasında en çok güvenlik tehdidi sosyal ağlarda, özellikle de yüksek sayıda kullanıcısı olan sosyal ağlarda meydana gelmiştir [9].

Kimlik taklidine karşı kullanıcıların sosyal ağ şifrelerinin ve sosyal ağlarda vermiş oldukları e-mail şifrelerinin güçlü olması gerekmektedir. Her türlü şifre işlemleri girilirken azami gizlilik sağlanmalıdır. internet kafe, otel ve halka açık erişim yerlerinden üyelik girişi ve şifre işlemi yapılmamalıdır [10].

Sosyal ağ sitelerine üye olunmadan önce gizlilik politikası, kullanım şartları ve özel şartlar okunarak, karşılaşılabilecek tehdit ve tehlikenin farkında olunarak bu ortamlar kullanılmalı, kişisel bilgilerin hangi şartlarla 3. şahıslarla paylaşılacağı bilincine sahip olunmalı ve ona göre karar verilerek üyelik işlemlerine başlanmalıdır [11].

Sosyal ağlara eklenen fotoğraf veya videolar bu hesapları ele geçiren kişiler tarafından, farklı amaçlar için izinsiz kullanılabilmektedir. Bu nedenle sosyal ağlarda fotoğraf ve video paylaşımında da dikkatli olunması gerekmektedir.

Sosyal ağ sitelerini kullanırken, kayıt olmak için şirket alan adı uzantılı e-posta adresi kullanılmamalıdır. Çalışılan kurumun üye olmak istenilen sosyal paylaşım sitesi için kuralları varsa bunlara uyulmalıdır. Profil sayfalarında kurumsal bilgiler paylaşılmamalıdır. Bazı sosyal ağ sitelerinde, bölge, çalışma alanı veya şirket adı gibi gruplaşmalar olmaktadır. Grup içerisine sızmaya çalışan bilişim korsanlarına karşı farkında olunmalıdır [12]

Sosyal ağ sayfalarında veya adres (URL) kısaltması hizmeti veren sitelerde, görünüşte zararlı olmayan, ancak tıklandıktan sonra kötü niyetli olduğu anlaşılabilen adresler yayınlanmaktadır [13]. Sahte sitelere karşı sadece bir e-posta mesajında veya bir web sitesinde yer alan bağlantılar üzerinden tıklanarak ağlara erişmeye çalışılmamalıdır. Mümkünse adres satırına erişmek istediğiniz web sitenin adresi ilgili yere yazılarak veya kopyalanarak web sitesine erişmeye çalışılmalıdır. Bu sayede, sosyal paylaşım sitesi gibi gösterilen tuzak sitelerin farkında olunmalıdır [14].

Verilerin güvenliğini sağlamak, ağ tabanlı saldırıların önüne geçmek ve saldırı anında farkındalık kazanmak için HTTP yerine mutlaka güvenli taşıma protokolü olan HTTPS (Güvenli Zengin Metin Transfer Protokolü) tercih edilmesi gerekmektedir. HTTPS kullanılarak gönderilen bilgiler üç temel koruma katmanı sağlayan taşıma katmanı güvenliği protokolü ile güven altına alınır:

Şifreleme : Alınan ve gönderilen veriler gizlice dinleme yapanlara karşı korumak için şifrelenir. Yani kullanıcı bir web sitesine göz atarken hiç kimse onun iletişimini "dinleyemez", sayfalar arasındaki etkinliklerini takip edemez veya bilgilerini çalamaz.

Veri bütünlüğü : Veriler aktarılırken, fark edilmeden kasıtlı olarak veya başka bir şekilde değiştirilemez ya da bozulamaz. Kimlik doğrulama: Kullanıcılarınızın kastedilen web sitesiyle iletişim kurduğu doğrulanır. Saldırılarına karşı korur ve kullanıcının güvenini sağlar [15].

Sosyal ağlar konusunda kullanıcıların farkındalıklarının arttırılması ve güvenlik önlemlerinin alınması için;

- Sosyal ağ sitelerine üye olunmadan önce gizlilik politikasından sitenin yayımlanan içeriği izleyip izlemediği öğrenilerek üye olunmalıdır.
- Tehlikelere karşı zaman zaman verilerin yedeklenmesi,
- Kişisel fotoğrafların sosyal ağlara yüklenmemesi ve eğer yüklenildiyse fotoğrafların etiketlenmemesi, kişisel bilgilerin ve iletişim bilgilerinin mümkünse paylaşılmaması,
- Hesabımızda sosyal ağ ayarlarından her türlü gizlilik ayarları yapılarak sınırlamalar getirilmesi ve gizlilik ayarlarının zaman zaman kontrol edilmesi,
- Tanımadığımız kişilerden gelen arkadaşlık tekliflerinin kabul edilmemesi,
- Dahil olduğunuz uygulamaların adımıza reklam ve yayın yapabileceğini düşünerek ayarlardan gerekli sınırlamaların getirilmesi,
- Sosyal ağ hesabımıza zaman zaman kullanıcı adı ile veya

farklı bir hesap ile giriş yapılması gerekmektedir.

V. ARAŞTIRMA

A. Araştırmanın Amacı

Bu araştırmanın temel amacı, günümüzde bilgi ve iletişim teknolojilerinin gelişmesiyle ortaya çıkan sosyal ağların kullanımının giderek artması fakat bu konuda toplumun bilgi güvenliği konusunda yeterli farkındalığının olmamasıdır. Bu ortamlarda meydana gelen güvenlik tehditlerinin önemli kişisel ve kurumsal zararlara neden olması dolayısıyla yapılan bu çalışma ile toplumun bilinçlendirilmesi ve farkındalığının arttırılması amaçlanmaktadır.

B. Araştırmanın Yöntemi

Sosyal ağlarda bilgi güvenliği konusunda 33 sorudan oluşan anket formu 234 kişi tarafından cevaplandırılmıştır. Bu anket formu, sosyal ağ sitelerini (Facebook, Twitter, Instagram, Linkedin vb.) kullanım durumlarını, alışkanlıklarını ve sosyal ağlarda güvenlik algısını belirlemeye yönelik genel bir araştırma eğilimi kapsamında hazırlanmıştır. Ayrıca soruların doğruluğunu teyid etmek amacıyla katılımcıların %25'i ile bire bir telefon görüşmesi de yapılmıştır. Ankete ilişkin istatistiksel sonuçlar yorumlanmaya çalışılmıştır.

C. Araştırmada Elde Edilen Bulgular

Kimlik soruları değerlendirildiğinde analiz grubunun% 59'i erkek, % 41'i kadındır. Yaş dağılımı açısından bakıldığında ise % 68'inin 18-25, % 6'sının 26-40 yaş aralığında olduğu ve % 26'sının ise 40 yaş ve üstünde olduğu görülmektedir.

Valid	Frequency	Percent
Erkek	137	58,5
Kadın	97	41,5
Total	234	100,0

Tablo 1: Cinsiyet Dağılımı

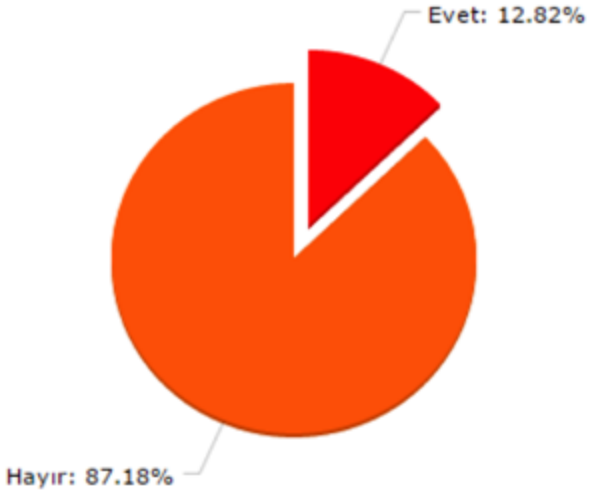
Örneklemde eğitim durumuna baktığımızda % 68'i üniversite mezunu yada öğrenci, %20'si yüksek lisans mezunu yada öğrenci, %6'sı doktora mezunu yada öğrenci, %6'sı da lise mezunudur.

- Sosyal ağ kullanıyor musunuz? Sorusuna 234 katılımcının % 100 gibi oldukça yüksek bir oranının Evet cevabını verdiği görülmektedir. Bu durumda katılımcıların tümünün sosyal ağları kullandığını söyleyebiliriz.
- Sosyal ağları kullanan katılımcıların kullandıkları araçlar sorulduğunda %65'inin akıllı telefon, %51'inin Laptop/ Netbook, %27'sinin masaüstü, %18'inin tablet ve %37'sinin hepsini kullandığı belirtilmiştir. Ayrıca katılımcıların sosyal ağ hesaplarını %53 kişisel, %2 kurumsal ve %45 her iki şekilde kullandığı belirtilmiştir.

Valid	Frequency	Percent
Evet	233	99,6
Hayır	1	0,4
Total	234	100,0

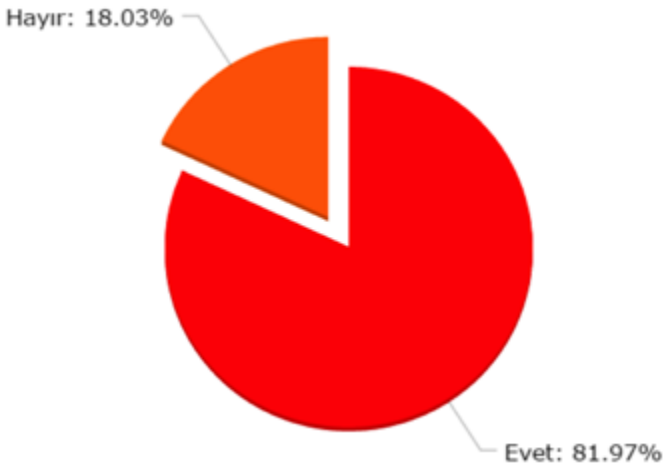
Tablo 2: Sosyal Ağ Kullanımı

- Sosyal ağ sitelerine üye olunmadan önce gizlilik politikası, kullanım şartları ve özel şartları okur musunuz? Sorusuna katılımcıların %43'ü Evet, %57'si Hayır demiştir.
- Sosyal ağ hesaplarında kendi isminizi mi kullanırsınız? Sorusuna %96'sı Evet, %4'ü de Hayır demiştir.
- Sosyal ağ hesaplarınızda yazı, resim ve video gibi paylaşımlarınız herkese açık mıdır? Sorusuna %18'i Evet, %82'si de Hayır demiştir.



Grafik 1 : Sosyal Ağların Güvenliği

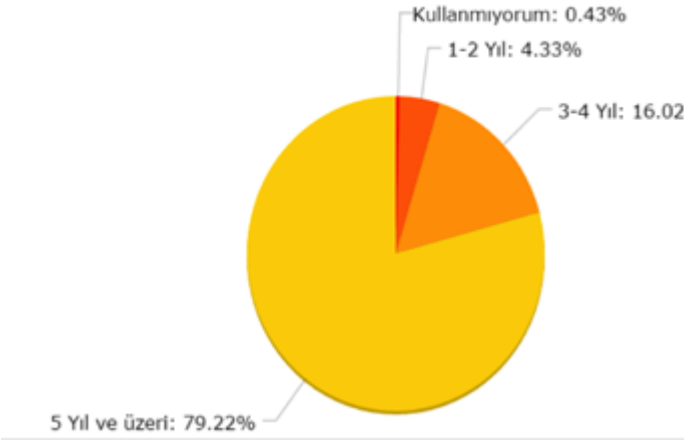
- Yukarıdaki grafikte görüldüğü gibi, Sosyal ağları güvenli buluyor musunuz? Sorusuna %13'ü Evet, %87'si de Hayır demiştir.
- Sosyal ağı kullandığınız bilgisayarınızda güvenlik programı (antivirüs, antitrojan vb.) bulunuyor mu? Sorusuna %82'si Evet, %18'i Hayır demiştir.



Grafik 2 : Sosyal Ağ Profilinin Güvenlik Ayarlarını Güncelleme

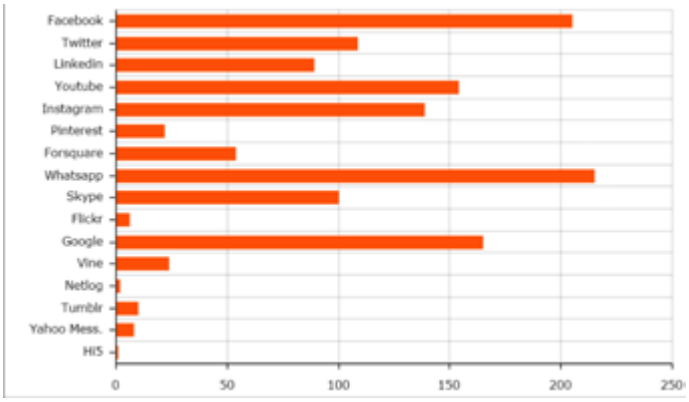
- Yukarıdaki grafikte görüldüğü gibi, Sosyal ağ profilinizin güvenlik ayarlarını güncelliyor musunuz? Sorusuna %82'si Evet, %18'i Hayır demiştir.
- Güncellemenin ise %26'sının her zaman, % 6'sının Kötü niyetli kişilerden zarar gördüğünde, %34'ü yeni güvenlik ayarları öğrenildiğinde, %34'ü sosyal ağ sitesi ayarlarının değiştirilmesi gerektiğine dair uyarı verdiğinde yapıldığı görülmüştür.
- Sosyal ağ kullanırken parolanızın çalınması, bilgisayarınıza virüs bulaşması veya güvenliğinizi tehdit eden herhangi bir olay yaşadınız mı? Sorusuna %27 Evet, %73 Hayır demiştir.

- Sosyal ağ sayfalarında zaman zaman yer alan linklere(URL) tıklar mısınız? Sorusuna % 45 Evet, %55 de Hayır demiştir.
- Sosyal ağ paylaşımları ve ekindeki dosyaları gördüğünüzde ne yaparsınız? Sorusuna %66'sı "Güvenlik nedeniyle seçerek açarım", %32'si "Hiç birisini açmam", %2'si de "Tümünü açarım" cevabını vermiştir.



Grafik 3 : Sosyal ağların Kullanımı

- Katılımcıların %79'u Sosyal ağları 5 yıl ve üzeri, %16'sının 3-4 yıldır, %4'ünün 1-2 yıldır kullandığını ve %1'inin de kullanmadığını görüyoruz.



- Yukarıdaki grafiğe baktığımızda sosyal ağlarda en çok kullanım %93 ile ilk sırada Whatsapp, %89 ile Facebook, %71 Google, %67 Youtube, %60 Instagram, %47 Twitter, %43 Skype, %39 Linkedin, %23 Forsquare ve diğerleri takip etmektedir.
- Gün içerisinde sosyal ağlarda ne kadar zaman geçiriyorsunuz? Sorusuna %40'ı 1 saatten az, %38'i 1-2 saat, %17'si 3-4 saat, %5'i 5 saat ve üzeri zaman geçirdiğini belirtmiştir.
- Sosyal ağlara aşağıdaki araçlardan en çok hangisiyle erişiyorsunuz? Sorusuna %83'ü Akıllı Telefon, %11'i Laptop/Netbook, %4'ü Masaüstü Bilgisayar ve %2'si de Tablet den eriştiğini belirtmiştir.
- Sosyal ağ parolanızı kimlerle paylaşırsınız? Sorusuna %79'u "Hiç kimseyle paylaşmam", %10'u "Diğer (Eşim ve çocuklarımla)", %7'si "Çok güvendiğim arkadaşlarımla", %3'ü "Kız arkadaşımla", %1'i "Babamla ve Annemle"dir.
- Sosyal ağ parolanız çalındığında ne yaparsınız? birden çok seçmeli sorusuna %68'i "Diğer sosyal ağ hesaplarımın şifrelerini hemen değiştiririm", %45'i "Sosyal ağ yöneticisi ile irtibata geçerim", %29'u "Bilgisayarımda casus program taraması yaparım", %8'i "Savcılığa başvururum", %7'si "Polise bildiririm", %5'i de "Bir daha sosyal ağ kullanmam" cevabını vermiştir.
- Sosyal ağ parolanızı hangi sıklıkta değiştiriyorsunuz? Sorusuna katılımcıların %34'ü "Hiç değiştirmem", %30'u "6 ayda bir", %27'si "Yılda bir",
- Hangi amaçla sosyal ağ kullanıyorsunuz? Birden çok seçmeli sorusuna Facebook için cevap veren katılımcıların %80'i "Eski arkadaşlarını bulmak için" kullandığı, %47'si "Duygu ve düşüncelerini paylaşmak", %46'sı "Gruplara üye olarak, sosyalleşmek", %45'i "Çevresindeki insanlar hakkında daha fazla bilgi sahibi olmak", %42'si "Sohbet etmek", %36'sı "Yeni arkadaşlar edinmek", "Güncel haberleri okumak", %35'i "Yazı paylaşmak", %32 "İlgili konularda bilgi edinmek", %31 "Etkinlik ve duyuruları takip etmek", %29'u "Video izlemek/Paylaşmak", %25'i "Oyun oynamak", %29'u "Mesleği ile ilgili gelişmeleri takip etmek" ve "Arkadaşlarla zaman geçirmek", %25 "Okul duyurularını" ve "Sanatsal Etkinlikleri takip etmek", %21 "Sportif etkinlikleri takip etmek" cevabını vermiştir. Sonuçlara göre, facebook kullananların en çok eski arkadaşlarını bulmak için kullandığı tespit edilmiştir.

VI. SONUÇ VE ÖNERİLER

Bu çalışmada elde ettiğimiz sonuçlara göre;

- Sosyal ağlar konusunda güvenlik farkındalığı arttıkça alacağımız güvenlik önlemleri de artmaktadır. Anket sonuçlarına göre sosyal ağları kullananların yarısının hem kişisel hem de kurumsal hesaplarını sürekli kullandıkları belirlenmiştir. Bu nedenle sosyal ağlarda kişisel ve kurumsal anlamda ilgili kurumlar tarafından eğitimler verilerek toplumda sosyal ağlarda güvenlik farkındalığı arttırılmalıdır.
- Anket katılımcılarının tamamının sosyal ağları kullandığını söyleyebiliriz. Tercih edilen sosyal ağlardan en çok Whatsapp ve Facebook tercih edildiği belirlenmiştir. Sosyal Ağ kullanımında en fazla akıllı telefonlardan bağlanıldığı görülmektedir. Katılımcıların çoğunluğunun, sosyal ağları 5 yıl ve daha fazla yıldır kullandığını da söyleyebiliriz.
- Katılımcıların %53'ünün, sosyal Ağ sitelerine üye olmadan önce gizlilik politikası, kullanım şartları ve özel şartları okumadıkları belirlenmiştir. %47'si ise şartları okuduklarını belirtmişlerdir. Bu kişilerin doğruluğundan emin olmak amacıyla katılımcıların %25 ile birebir telefon görüşmesi yapılarak, doğrulukları teyit edilmiştir. Sanal ortamlarda kullanılan sosyal ağlarda gerekli güvenlik önlemlerinin alınmaması saldırganların bilgilerimize kolayla erişebilmesine olanak sağlayarak, saldırı yapma imkanlarını arttırmaktadır. Bu yüzden de sosyal ağlarda gerekli şartların okunarak, ilgili gizlilik ayarlarının yapılması ve gizlilik ayarlarının sıkça değiştiği göz önüne alınarak takip edilmesi gerekmektedir.
- Katılımcıların çoğunluğu, sosyal ağ hesaplarında yazı, resim ve video gibi paylaşımların herkese açmadıklarını belirtmişlerdir. Katılımcıların %25 ile birebir telefon görüşmesi yapılarak da bu konuda bilgi alınmıştır. Bu sonuca göre katılımcıların bu konuda farkındalıklarının olduğu tespit edilmiştir. Katılımcıların %94'ünün üniversite mezunu veya öğrencisi olduğu dikkate alındığında farkındalığın daha fazla olduğu öngörülebilir.
- Katılımcıların büyük çoğunluğu sosyal ağları güvenli bulmadıklarını belirtmişlerdir. Bu sonuca göre, sosyal

ağların katılımcılara güven vermediği söylenebilir.

- Katılımcıların çoğu sosyal ağ kullandığı cihazlarında anti virüs programları kullandıklarını belirtmişlerdir. Kullanılan uygulama yazılımları, işletim sistemi vb. yazılımların zamanında güncellenememe-sinden dolayı güvenlik zafiyeti oluşmaktadır. Temel güvenlik önlemlerini almak için kullanılan tüm bilgisayar yazılımları güncel tutulmalı, anti-virüs ve güvenlik duvarı yazılımları mutlaka kullanılmalı ve kötücül yazılım ile spam engelleyici filtreler tercih edilmelidir [14]
- Sosyal ağ profilinin güvenlik ayarlarını güncellemenin çoğunlukla yeni güvenlik ayarları öğrenildiğinde ve sosyal ağ sitesi, ayarların değiştirilmesi gerektiğine dair uyarı verdiğinde yapıldığı görülmüştür. Sonuç olarak sosyal ağlarda güvende olmak için her zaman yapılması gereken güncellemelerin katılımcılar tarafından yapılmadığı tespit edilmiştir. Güvenlik için güncellenmelerin düzenli olarak yapılması gerekmektedir.
- Katılımcıların çoğunluğunun sosyal ağ kullanırken parola çalınması, cihazlarına virüs bulaşması veya güvenliklerini tehdit eden herhangi bir olay yaşamadıkları görülmektedir. Arkadaşların e-posta adreslerini vermekten kaçınmak için, sosyal ağ hizmetlerinin e-posta adres defterini taramasına izin vermemek bu uygulamaların dağılmasına engel olacaktır [16].
- Katılımcıların yarısı, sosyal ağ sayfalarında yer alan linklere(URL) tıkladıklarını belirtmişlerdir. Bu sonuca göre katılımcıların güvende olduğunu söyleyemeyiz. Bu linkler çoğunlukla virüs içermekte ve hesabımızı ele geçirmek için saldırganlara fırsat yaratmaktadır. Özellikle bu linkler tıklamamız için reklam içerikli ve cezbedici niteliktedir. Bu yüzden tıklanmamalıdır.
- Katılımcıların büyük çoğunluğu, sosyal ağ paylaşımları ve ekindeki dosyaları güvenlik nedeniyle seçerek açarım demiştir. Normal şartlarda güvenlik dolayısıyla dosyaların tamamının taranarak açılması gerekir.
- Katılımcıların büyük çoğunluğu, sosyal ağ parolalarını hiç kimseyle paylaşmayacağını belirtmiştir. Parolanın tek ve kişiye özel olması ilkesinin benimsenmiş olması güvenlik farkındalığının bir sonucudur. Her bir sosyal ağ hesabında kullanılan parolaların birbirinden farklı olması, art arda gelen numaralar içermemesi, karakter ve sayı içerecek şekilde kullanılması gerekmektedir.
- Katılımcıların büyük çoğunluğu, sosyal ağ parolaları çalındığında, diğer sosyal ağ hesaplarımın şifrelerini hemen değiştireceğini, katılımcıların yarısı da sosyal ağ yöneticisi ile irtibata geçeceğini belirtmiştir. Katılımcıların güvenli şifreleme konusunda azda olsa farkındalıklarının olduğunu söyleyebiliriz.
- Katılımcıların bir kısmı, sosyal ağ parolasını hiç değiştirmeyeceğini, bir kısmı 6 ayda bir veya yılda bir değiştirebileceğini belirtmiştir. Sonuç olarak, bu konuda farkındalığın olmadığı ve destek verilmesi gerektiği söylenebilir. Sosyal ağlarda güvende olmak için parolaların sıklıkla değiştirilmesi gerekmektedir. Bunun için de parolanın mümkün olduğunca zor ve karmaşık seçilmesi önemlidir. Bilgisayar korsanlarının finansal veya diğer hesaplara girerken sık kullandıkları bir yöntem de hesap giriş sayfasındaki "Parolamı unuttum" bağlantısına tıklatmaktır. Hesabınıza girebilmek için doğum gününüz, oturduğunuz şehir, lisedeki sınıfınız veya annenizin kızlık soyadı gibi güvenlik sorularına verdiğiniz cevaplarınızı ararlar. [16].
- Sonuç olarak, yapılan çalışmalar, web hackleme olaylarının yaklaşık %50'sinin sosyal ağ sitelerinde olduğunu göstermektedir. Breach Security, web hackleme veri tabanları üzerinde çalışmaktadır ve çevrimiçi atakların 2008 yılında %19, 2009 yılında ise %30 oranla sosyal ağ sitelerine olduğu görülmektedir [17].
- Bu yüzden ülkemizde son yıllarda önem arz eden sosyal ağlarda güvenlik konusunda sosyal medya, emniyet müdürlükleri (siber suçlar birimi), ilgili kurum ve kuruluşlar ile üniversitelerin, ilgili derneklerin topluma sosyal ağlarda bilgi güvenliği konusunda eğitimler vererek, toplumu bilinçlendirmesi ve farkındalık yaratması gerekmektedir.

KAYNAKLAR

[1] http://www.tuik.gov.tr/ [21 Ağustos 2015 tarihinde erişilmiştir].

[2] KurumsalHaberler, Sosyal Medya Nedir, 2010. Available: http://www.kurumsalhaberler.com/pr/sosyal-medyanedir.aspx .
[07 Ağustos 2015 tarihinde erişilmiştir].

[3] D. Hobson, Social networking – not always friendly, Computer Fraud
& Security, cilt 2008, no. 2, p. 20, 2008.

[4] U. Yavanoğlu, Ş. Sağıroğlu ve İ. Çolak, "Sosyal ağlarda Bilgi Güvenliği Tehditleri ve Alınması Gereken Önlemler", Politeknik Dergisi , cilt 15, no. 1, pp. 15-27, 2012.

[5] eBizMBA, Top 15 Most Popular Social Networking Sites. Available: http://www.ebizmba.com/articles/social-networkingwebsites [22 Eylül 2015 tarihinde erişilmiştir].

[6] Strighini, G., Kruegel, C., Vigna, G, "Detecting Spammers on Social Networks", ACSAC'10 Austin, Texas, ABD, 6-10, (2010).

[7] Sancho, D., "Security Guide to Social Networks", White-Paper Trend Micro Inc., (2009).

[8] M. Qi ve D. Edgar-Nevill, Social networking searching and privacy issues, Information Security Technical Report, cilt 2011, no. 16, pp. 74-78, 2011.

[9] "Cisco Annual Security Report", (2013). Avaible: http://www.cisco.com/web/offer/gist_ty2_asset/Cisco_2013_ASR.pdf [21 Eylül 2015 tarihinde erişilmiştir].

[10] Siber Güvenlik Tehdit Merkezleri Open Web Application Security Project: www.owasp.org / [2012] [19 Ağustos 2015 tarihinde erişilmiştir].

[11] "Facebook Security", Available: http://www.facebook.com/security?v=app_4949752878, 2010 [21 Eylül 2015 tarihinde erişilmiştir].

[12] "A Privacy Paradox: Social Networking in the United States", Available: http://firstmonday.org/ojs/index.php/f%20m/article/view/1394/1312,%202010 [20 Eylül 2015 tarihinde erişilmiştir].

[13] Sancho, D., "Security Guide to Social Networks", White-Paper Trend Micro Inc., 2009.

[14] Canbek G., Sağıroğlu Ş., "Bilgi ve Bilgisayar Güvenliği: Casus Yazılımlar ve Korunma Yöntemleri", ISBN: 975-6355-26-3, Grafiker, Ankara, 2006.

[15] Available: https://support.google.com/webmasters/answer/6073543?hl=t r
[22 Eylül 2015 tarihinde erişilmiştir].

[16] Microsoft , Sosyal ağ güvenliği için 11 ipucu, Microsoft, 2015. Available:.http://www.microsoft.com/trtr/security/online-privacy/social-networking.aspx . [21 Ağustos 2015 tarihinde erişilmiştir].

[17]Computer Fraud & Security, Hacking attacks target social networking, ELSEVIER, 2009.

GELİŞMİŞ ISRARCI TEHDİTLER VE GIT ÖRNEKLERİNİN KARŞILAŞTIRILMASI

Esra Söğüt[1], O. Ayhan Erdem[2]

1 Gazi Üniversitesi, Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü 06500, Teknikokullar, Ankara-Türkiye, Telefon: +90 312-2028561, esrasogut@gazi.edu.tr

2 Gazi Üniversitesi, Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü 06500, Teknikokullar, Ankara-Türkiye, ayerdem@gazi.edu.tr

Özet — Gelişmiş Israrcı Tehdit (GIT) bilgisayarların, bilgisayar sistemlerinin ve kullanıcıların karşılaştığı en büyük tehditlerden birisidir. Son derece gelişmiş nitelikte, özel olarak ve az sayıda hazırlanmış GIT'lerin tespit edilebilmesi de kolay olmamaktadır. Deneyimli ve sabırlı saldırganlar tarafından oluşturulan GIT'ler dünya genelinde farklı yerlerde ve farklı alanlarda karşımıza çıkabilmektedir. Kullanıcılarda farkındalık oluşturmak amacıyla bu çalışmada GIT hakkında detaylı bilgi verilmiştir. Ayrıca dünya genelinde büyük etkiler meydana getiren GIT örnekleri incelenmiş ve bu örneklerin çalışma yapılarına göre karşılaştırılmaları yapılmıştır.

Anahtar Kelimeler — Gelişmiş Israrcı Tehdit, Siber Savaş, Siber Güvenlik, Stuxnet

Abstract — The Advanced Persistent Threat (APT) is one of the greatest threats faced by computers, computer systems and users. APT has highly advanced qualities and a small number of specially crafted so APT cannot be detected easily. APTs can appear in different places and areas worlwide which are created by experienced and patient attackers. This study aims to create awareness about the APT. In addition, APT samples which constituting larger impact on the world are analyzed and comparisons are made according to the working structures.

Index Terms — Advanced Persistent Threats, Cyber Warfare, Cyber Security, Stuxnet

I. GİRİŞ

Günümüzde internetin ve bilgisayarın kullanıldığı alanlar giderek artmaktadır. Kullanılan ve saklanılan bilgiler de birikmekte ve bunların güvenliği büyük bir sorun haline gelmektedir. Kişilere olabileceği gibi kamu kurum ve kuruluşlarına, kritik altyapılara veya büyük şirketlere yönelik siber saldırılar gerçekleşebilmektedir. Sahip olunan bilginin önemine göre saldırılar da değişebilmektedir.

Özellikle kritik altyapılar, büyük şirketler, telekom operatörleri ve kamu kurumlarına yönelik saldırılar büyük zararlar verebilecek boyutlarda olabilmektedir. Bu alanda yeni bir saldırı türü sayılabilecek Gelişmiş Israrcı Tehdit(GIT)'i gösterebiliriz [1]-[10]. Bu tehdit hedefini belirleyip, başarılı oluncaya kadar çalışmasını sürdürmektedir. Gelişmiş Israrcı Tehdit'in hedefe sızdığını ve orada çalıştığını anlamak kolay olmamaktadır. Bu sebeplerle etkisi diğer saldırı türlerine göre çok daha büyük olabilmektedir.

Bu çalışmada GIT hakkında bilgi verilmiştir. Birinci bölüm giriş olarak düzenlenmiş ve bu bölümde çalışma hakkında genel bilgiler sunulmuştur. Çalışmanın ikinci bölümünde GIT'lerin ne olduğundan ve GIT'lerin çalışma yapılarının nasıl olduğundan bahsedilmiştir. Üçüncü bölümde ise ele alınan GIT örnekleri incelenmiş ve belirlenen özelliklere göre kıyaslama yapilmiştir. Seçilmiş olan Stuxnet, Duqu, Flame, Red October ve MiniDuke uygulamaları hakkında bilgi verilip karşılaştırmalı olarak incelemeler yapılmıştır. Çalışmanın son bölümünde GIT örneklerinin incelenmesi sonucunda elde edilen sonuçlar aktarılmış ve sonuç verileri tabloda gösterilmiştir.

II. GELİŞMİŞ ISRARCI TEHDİT VE ÇALIŞMA YAPISI

Gelişmiş Israrcı Tehdit (GIT), günümüzün siber dünyasında büyük bir tehdit oluşturmakta ve etkisi giderek artan bir saldırı haline gelmektedir. ABD Ulusal Standartlar ve Teknolojisi Enstitüsü (NIST) tarafından GIT için yapılan tanım şu şekildedir: Bilgi seviyesi ve tecrübesi yüksek olan saldırgan, gerekli kaynakları kullanarak çoklu saldırı vektörleri (siber, fiziksel gibi) ile amaçlarına ulaşmak için kendisine uygun fırsatlar oluşturur. Buradaki amaçlar genellikle, hedeflenen kuruluşların bilgi teknoloji altyapısını oluşturan ve onu kapsayan sistemin içinde bulunmak ve ileriki zamanlarda da faaliyete geçebilmek için doğru şekilde konumlanmaktır. Gelişmiş Israrcı Tehdit: (i) amaçlarını gerçekleştirmek için uzun süre takipte kalır, (ii) hedefin savunma sistemine uyum sağlar, (iii) belirlenen amacı gerçekleştirmek için gerekli olan etkileşim seviyesini sağlar ve onu korur [2].

Yukarıda yer alan tanımdan da anlaşılacağı üzere GIT, kaynağı iyi şekilde sağlanmış yetenekli ve kararlı saldırgan ya da saldırganlar tarafından gerçekleştirilen saldırılardır. Etkisi, yol açabileceği zararları ve oluşabilecek sonuçlar tam olarak bilinememekte veya tahmin edilememektedir. Özel olarak hazırlanmış ve birçok aşaması olan GIT'ler farklı özelliklere sahip olsa da çalışma yapıları olarak benzerlik göstermektedir. GIT saldırı evrelerini tanımlamak için kullanılan yönteme Saldırı Ölüm Zinciri (Intrusion Kill Chain) denilmektedir ve tipik bir GIT saldırısı şu altı evreden oluşmaktadır: keşif ve silahlanma, teslim ve dağıtım, ilk sızma ve saldırı, komuta ve kontrol, yanal hareket ve veri çekme. Bu evrelerin sıralı haldeki gösterimine, GIT akış şeması olarak Şekil1'de yer verilmektedir [3]-[5].

Şekil 1 : Gelişmiş Israrcı Tehdidin Akış Şeması [3]-[5]

A. Keşif ve Silahlanma (Bilgi Toplama)

GIT için en önemli adımdır ve saldırıya başlamadan önce gerekli olan hazırlık evresidir. Saldırganın saldıracağı hedefi tanıması için gerekli olan işlemler bu evrede gerçekleştirilir. Kurbanla ilgili olabildiğince araştırmanın yapılması ve elde edilen bilgilerin toplanması gerekmektedir. Bunun için açık kaynak istihbarat araçları, sosyal mühendislik yöntemleri kullanılmakta ve zaafiyet tarama işlemleri yapılmaktadır.

B. Dağıtma

Bu aşama hedef belirlendikten ve hedefle ilgili bilgiler toplandıktan sonra gerçekleşir. Bu evrede saldırganlar hazırladıkları, sistemde açık bulma kodlarını veya sistemi sömürme kod parçacıklarını (exploit) hedeflerine yönlendirmektedir. 2004-2010 yılları arasında Lockheed Martin Computer Incident Response Team (LM-CIRT)'in yaptığı gözlemlere göre yaygın olarak kullanılan dağıtma yöntemlerinden üç tanesi e-posta eklentileri, web siteleri ve USB taşınabilir medya araçları olarak belirlenmiştir [3].

C. İlk Sızma-Saldırı

Dağıtma aşamasında silah olarak kullanılan istismar ve sömürü kod parçacıkları hedef sistemde başarıya ulaşınca bu evreye geçilir. Saldırgan, hedef sisteme yetkisiz erişim hakkını ilk kez elde ettikten sonra ilk sızma işlemine sıra gelmektedir. Bu evrede hedef sistemin güvenlik açığından, işletim sistemi veya uygulama açığından faydalanılabilir ya da hedef farkında olmadan kendisi kodları çalıştırarak sızmayı başlatabilir.

D. Komuta ve Kontrol

Hedef sistemde yetkisiz izinler elde edilerek, uzaktan erişim özelliğine sahip trojan veya arkakapı sisteme yerleştirilerek kurban sistemde kalıcılık sağlanmaya çalışılır. GIT, başarılı bir arkakapı kurulması ile komuta ve kontrol sistemini kullanabilmektedir. Komuta ve kontrol sistemi kurulunca kendi hâkimiyetini yitirmiş hedef sistem ele geçirilmiş olur. Saldırgan, hedef sistemde tespit edilmemek ve dikkat çekmemek için yasal hizmetleri veya herkesin kullanımına açık olan araçları tercih etmektedir.

E. Yayılma

Yayılma aşaması diğer aşamalara göre daha fazla zaman almaktadır. Saldırgan istediği bilgileri elde edinceye kadar kendini belli etmeden uzun süre boyunca çalışabilmektedir. Ele geçirilmek istenen sistem ile kontrol-komuta sunucuları arasında iletişim kurulması, GIT unsurlarının hedef ağ içerisinde hareket etmesine zemin hazırlamaktadır. Hedef sistem üzerinde kontrolün sağlanıp ağ üzerinden kontrolün genişletilmesi, sistemin özelliklerini keşfetmek ve sistemle ilgili önemli bilgileri toplamak için olanak sağlamaktadır.

F. Veri Çekme

Hassas, önemli veya gizli bilgilerin ele geçirilmesini amaçlayan saldırganlar için veri çekme aşaması kritik öneme sahiptir. Bilgiler dışarıya aktarılırken genellikle test amaçlı kullanılan deneme sunucuları kullanılmaktadır. Bilgiler genellikle sıkıştırılmış ve şifrelenmiş halde iletilir ve saldırgan deneme sunucusundan kendi sistemine bilgileri geçirir.

III. GIT UYGULAMALARI VE ÇALIŞMA YAPILARINA GÖRE KARŞILAŞTIRILMALARI

Gelişmiş Israrcı Tehditler siber dünyadaki savaş için çok büyük tehlike oluşturmaktadır. GIT'lerin meydana getirdikleri ya da getirecekleri etkiler hemen anında anlaşılamamaktadır. Siber savaş aracı olarak kullanılan GIT uygulamaları giderek artmakta ve saldırıya uğrayan kurbanlar büyük zarar görebilmektedir. Günümüze yakın zamanlarda tespit edilmiş ve siber güvenlik alanında önemli yer edinmiş GIT uygulamaları ele alınmaktadır. Sahip olduğu özelliklerine ve çalışma yapılarına göre incelenen ve karşılaştırılan uygulamalar meydana getirdikleri etkilerine ve tespit edilme tarihlerine göre seçilmiştir. Ele alınan GIT'ler Stuxnet, Duqu, Flame, Red October ve MiniDuke uygulamalarıdır. Bu uygulamalar incelenerek elde edilen bilgiler Tablo1 üzerine karşılaştırılmalı olarak yerleştirilmiştir.

A. Stuxnet

İlk olarak ele alınan GIT uygulaması olan Stuxnet Haziran 2010'da tespit edilmiş ve güvenlik alanında büyük yankı uyandırmıştır. İran'daki SCADA (Merkezi Denetleme, Kontrol ve Veri Toplama) sistemlerini hedef alan bu GIT'in, incelemeler sonucunda sahip olduğu karmaşık yapısı ile diğer sıradan kötücül yazılımlardan farklı olduğu anlaşılmıştır [6]. Programlanabilir mantık denetleyicisi (PLC) sistemlerini, endüstriyel kontrol sistemlerini (ICS) ve Windows sisteminin kullandığı Siemens Step-7 yazılımını da kontrol altına alarak İran'ın nükleer yakıt tesisine zarar vermeyi amaçlamıştır [7]. Uranyum zenginleştirmede kullanılan ve önemli görevleri olan santrefüjlerin çalışmasını bozarak tesise fiziksel olarak zarar vermiş ve yaklaşık iki ile dört yıl olarak sistemi geriye itmiştir [8]. Stuxnet'in sisteme nasıl bulaştığı tam olarak bilinmemekte fakat taşınabilir sürücülerle sisteme bulaştığı tahmin edilmektedir [9].

Bulaştığı sistemde kendi kendini çoğaltabilme becerisine sahip olan Stuxnet, karşılaştırılan diğer GIT uygulamaları arasında bu yönüyle farklılık oluşturmaktadır. Sistem içerisinde taşınabilir medya ile veya ağ üzerinden çoğalabilmektedir. Bir diğer farklılık ise keylogger özelliğine sahip olmamasıdır. Keylogger özelliği ile kurbana ait hassas ve önemli bilgiler (şifreler, kullanıcı bilgileri gibi) toplanabilmekte veya çalınabilmektedir. Keylogger bileşenleri ile kurban sisteme ait ekran görüntüsü alma, e-posta mesajlarını yakalama, konuşmaları kaydetmek için bilgisayara ait mikrofonu kullanma gibi işlemler yapılabilmektedir. Stuxnet için bu işlemlerden söz edilmemektedir. Şifreleme yöntemi olarak diğer GIT uygulamaları gibi dizi şifreleme algoritması kullanılmıştır [9],[10]. Veri güvenliğinde kullanılan şifreleme algoritmalarından biri olan dizi şifrelemesi, bir anahtardan üretilen anahtar dizisi ile mesajda yer alan tüm harflerin özel bir algoritma (XOR işlemi gibi) kullanılarak sırayla şifrelenmesidir. Dizi şifreleme algoritması simetrik şifreleme ailesinden kabul edilmektedir [11]-[15]. Stuxnet, verileri komuta-kontrol sunucularına gönderirken kodlamak için ve kendi küçük parçalarının şifresini çözmek için de dizi şifreleme algoritmasını kullanmaktadır.

B. Duqu

Eylül 2011 tarihinde tespit edilen Duqu, hedef sistemlere sızmak için MS Word yöntemini kullanmaktadır [16]. Microsoft Word dosyalarının içerdiği True Type yazı tipi ayrıştırma sıfırıncı gün açıklığı (CVE-2011-3402) ilk saldırı vektörü olarak kullanılmıştır [9],[17]. Duqu da Stuxnet gibi SCADA sistemlerini hedef almakta fakat sistemlere zarar vermek yerine casusluk yaparak bilgi toplamayı amaçlamaktadır. Duqu uygulamasının, Stuxnet gibi sistemlere zarar verme amacı taşıyan saldırılar için istihbarat sağlamak amacıyla ve oluşacak saldırının etkinliğini arttırmak amacıyla hazırlandığı düşünülmektedir. Ayrıca SCADA sistemleri ile ilgili kritik bilgilerin tespiti yapılarak saldırı yapılacak sistemin zayıf ve güçlü yönlerinin ele geçirilmesi bu uygulama ile sağlanmıştır. Bulaştığı sistemde Stuxnet gibi kendi kendini çoğaltma özelliğine sahip olmayan Duqu otomatik olarak ağ içinde veya system içinde kopyalanamamaktadır. Keylogger özelliği ile kurbanla ilgili bilgiler çeşitli yollarla ele geçirilmektedir [9],[16]. Çalınan bilgilerin ve yapılandırma dosyalarının şifrelenmesi için diğer GIT uygulamaları gibi dizi şifreleme algoritması kullanılmıştır [18]. Bunun yanında diğer GIT uygulamalarından farklı olarak AES-CBC (Gelişmiş Şifreleme Standardı-Blok Şifreleme Zinciri) modu kullanılmıştır [10]. Bilgileri şifreleyebilen ve deşifre edebilen bir simetrik blok şifrelemesi olan AES, elektronik verileri korumak için kullanılabilecek şifreli bir algoritma belirtir [19]. CBC modu, daha önce şifrelenmiş bloklar ile düz metin bloklarını zincirleyerek birleştiren ve bu şekilde şifreleme işlemi yapan gizlilik modudur [20],[21]. AES'e ait beş güvenlik modundan biri olan CBC modu Duqu tarafından kullanılmaktadır.

C. Flame

Karşılaştırılan diğer bir GIT uygulaması olan Flame 2012 yılının Mayıs ayında tespit edilmiştir. İran, İsrail, Batı Şeria, Sudan, Suriye, Lübnan, Suudi Arabistan ve Mısır gibi Orta Doğu ülkelerini hedef alan Flame devlet kuruluşları ve eğitim kurumları gibi önemli yerlere ulaşmıştır [6],[22],[23]. Stuxnet gibi bulaştığı sisteme zarar vermek amacı gütmeyen Flame, bilgi toplamak amacıyla hareket etmiştir. Bulaştığı çok sayıda bilgisayardan veri sızdırmış ve hırsızlık yapmayı sürdürmüştür. Tespit edildiği zamandan beş ile sekiz yıl öncesinden beri aktif olduğu tahmin edilmektedir [22]. Flame'in sistemlere nasıl bulaştığı tam olarak bilinmemekte ama taşınabilir sürücülerle veya oltalama saldırılarıyla bulaştığı tahmin edilmektedir.

Flame, kurban sistemde kendi kendini çoğaltma özelliğine sahip değildir. Bulaştığı sistemde manuel olarak kopyalanabilmektedir. Sahip olduğu keylogger özelliği ile bulaştığı sistemdeki önemli bilgileri çalabilmektedir [9]. Diğer GIT uygulamalarından farklı olarak, keylogger bileşenleri ile web kamerasını (ses ve görüntü kaydetmek için) da kullanabilmektedir. Ayrıca Bluetooth ve Wifi özelliklerini, USB ve depolama aygıtlarını da veri çalmak için kullanabilmektedir. Bluetooth özelliği aktif olduğunda çevredeki cihazlara da ulaşıp sızma ve çalma işlemlerini gerçekleştirebilmektedir. Bu özellikler de Flame'in etkisini ve önemini arttırmaktadır. Yapılandırma dosyalarını ve yakaladığı verileri şifrelemek için dizi şifreleme algoritması, RC4 algoritması ve Substitution şifrelemesi kullanmıştır [6],[22]. Diğer GIT uygulamalarından farklı olarak kullanılan Substitution şifrelemesi, verilerede bayt bayt değiştirme yaparak şifreleme anlamına gelir. Burada her karakterin yerine farklı karakter konularak şifreleme yapılır ve uygun bir tablo oluşturularak her karaktere karşılık gelen karakter bu tabloda saklanır. RC4 algoritması ise şifrelenecek veriyi akan bir bit dizisi olarak algılar ve önceden belirlenen anahtar ile veriyi şifreler. RC4, rastgele olarak ürettiği anahtar akışlarını, hem şifreleme hem de şifreyi çözme sırasında XOR işlemi ile mesaja uygulamaktadır [24]-[27].

D. Red October

Ele alınan GIT uygulamalarından olan Red October Ekim 2012 tarihinde tespit edilmiş ve Doğu Avrupa, Batı Avrupa, Kuzey Amerika gibi bölgeleri etkisi altına almıştır. Bu bölgelerdeki kamu kurumlarına, devlet birimlerine, bilimsel araştırma merkezlerine, diplomatik birimlere, askeri birimlere, enerji/nükleer araştırma birimlerine gibi çok sayıda yere sızarak casusluk faaliyetleri sürdürmüştür. Bulaştığı sistemleri ele geçirmek veya çökertmek amacı gütmemektedir. Oldukça geniş bölgelerde çalışan bu GIT uygulamasının, gizli bilgileri ve jeopolitik öneme sahip istihbaratları 2007 yılından beri topladığı bilinmektedir [28],[29]. Yüksek profilli kurbanlar seçen Red October'ın çaldığı bilgileri ne için kullandığı veya kullanacağı tam olarak bilinememektedir. Elde edilen bu bilgiler karaborsada satılabilir veya doğrudan kullanılabilir niteliktedir. Hedef sistemlere sızma yöntemi olarak MS Word (CVE-2012-0158, CVE-2010-3333), MS Excel (CVE-2009-3129) ve Java (CVE-2011-3544) güvenlik açıklıkları kullanılmıştır [30]. Geleneksel saldırı hedeflerine ek olarak akıllı telefonlar da etki altında kalmıştır. İphone, Nokia veya Windows Mobile gibi mobil cihazlardan veri çalma özelliğine de sahiptir. Çalınan bilgiler telefona ait özellikler, telefon defteri, kişiler, arama geçmişleri, ajanda veya mesajlar olabilmektedir [28],[29].

Red October, kurban sistemde kendi kendini çoğaltma özelliğine sahip değildir ve bulaştığı sistemde manuel olarak çoğalabilmektedir. Keylogger özelliğine sayesinde bulaştığı sistemdeki bilgileri, klavye hareketlerini kaydederek veya ekran görüntüsü alarak çalabilmektedir [9],[28]. Yakaladığı verileri şifrelemek için dizi şifreleme algoritmasının yanında ROR işlemi de kullanılmıştır (XOR+ROR) [31],[32]. Diğer GIT uygulamalarından farklı olarak kullanılan ROR işlemi, sağ tarafta yer alan bitlerden belirtilen kadarının düşmesiyle solda açılan yere belirtilen bitlerin yerleşmesi işlemidir. Red October şifreleme işleminde bu algoritmalardan faydalanmaktadır [33]-[35].

E. MiniDuke

İncelenen son GIT uygulaması olan MiniDuke Şubat 2013 tarihinde tespit edilmiştir [36]. Almanya, Ukrayna, Portekiz, Romanya, Çek Cumhuriyeti, İrlanda, Birleşik Krallık, Macaristan ve Türkiye'nin de dâhil olduğu 23 ayrı ülkeyi hedef alan MiniDuke devlet kuruluşlarını, büyükelçilikleri, araştırma merkezlerini, sağlık kuruluşlarını, sosyal vakıfları ve özel şirketleri kurban olarak seçerek önemli yerlere ulaşmıştır. Bulaştığı sisteme zarar vermek amacı gütmeyen fakat bilgi toplama amacı taşıyan bir GIT uygulamasıdır. Bulaştığı çok sayıda bilgisayardan veri sızdırmakta ve casusluk faaliyetlerini sürdürmektedir. Bu GIT uygulamasının saldırganlardan komuta almak için birisi Panama'da diğeri ise Türkiye'de bulunan iki sunucuya bağlandığı bilinmektedir [37]. MiniDuke diğer GIT'lerden farklı olarak Twitter'ı ve

Google Arama özelliğini kullanabilmektedir. Twitter'a girerek (kullanıcıdan bağımsız olarak) önceden oluşturulan hesaplara tweet atılmasını ve bu şekilde saldırgan-kurban arasında iletişim kurulmasını sağlamaktadır [38]. Twitter aktif olmadığında ise Google Arama uygulamasını kullanmaktadır [39]. Kurban sistemlere sızma yöntemi olarak PDF dosyaları sosyal mühendislik yöntemleriyle gönderilmektedir. PDF dosyalarına ait güvenlik açıklığından (CVE-2013-6040) faydalanılarak sistemlere sızılmaktadır [36],[38].

Kurban sistemde kendi kendini çoğaltma özelliğine sahip olmayan MiniDuke, bulaştığı sistemde manuel olarak çoğalabilmektedir. Keylogger özelliğine sahiptir. Bu özellik sayesinde bulaştığı sistemdeki bilgileri veya ekran görüntülerini çalabilmektedir [10]. Yakaladığı verileri şifrelemek için dizi şifreleme algoritmasını kullanmaktadır. Sahip olduğu arka kapılar ile dizin oluşturma, dosya kopyalama, dosya taşıma, dosya kaldırma, süreçleri durdurma, yeni kötücül yazılımları indirme ve onları çalıştırma gibi işlemleri gerçekleştirebilmektedir. Tüm bunları yapabilen MiniDuke uygulaması 20KB boyutundadır [36],[38].

F. GIT Uygulamalarının Çalışma Yapılarına Göre Karşılaştırılması

Farklı zamanlarda yapılmış GIT örnekleri karşılaştırılarak, elde edilen bilgilere Tablo1'de yer verilmiştir. Karşılaştıma işlemi birçok farklı özelliğe göre yapılmaktadır. Karşılaştırılan GIT'ler için Stuxnet, Duqu, Flame, Red October ve MiniDuke uygulamaları seçilmiştir [9],[10]. Bu örnekler seçilirken her bir GIT uygulaması için tespit edilme tarihleri ve oluşturdukları etkileri göz önüne alınmıştır.

Farklı GIT uygulamaları sahip oldukları özelliklere göre karşılaştırılmakta ve elde edilen sonuçlar Tablo1'de gösterilmektedir. Farklı GIT uygulamaları kıyaslanırken, bu uygulamaları birbirinden ayırabilecek nitelikte olan ve her bir uygulama için önemli olan temel özellikler ele alınmaktadır. Bunlar: ilk sızma şekli, şifreleme yöntemleri, kendini çoğaltma yöntemleri, tespit edilme tarihi, hedefleri, amacı, etkileri ve keylogger özelliğine sahip olup olmamasıdır. İncelenen GIT uygulamaları 2010, 2011, 2012 ve 2013 yıllarında tespit edilen ve dünya genelinde önemli etkiler oluşturan uygulamalardır.

GIT Uygulamaları	Stuxnet	Duqu	Flame	Red October	MiniDuke
İlk Sızma Şekli	Bilinmiyor	MS Word	Bilinmiyor	MS Word, Excel ve Java	PDF
Şifreleme Yöntemleri	Dizi şifreleme	Dizi şifreleme, AES-CBC	Dizi şifreleme, RC4, Substitution	Dizi şifreleme, XOR+ROR	Dizi şifreleme
Kendini Çoğaltma Yöntemi	Taşınabilir medya ile veya ağ üzerinden	El ile	El ile	El ile	El ile
Tespit Edilme Tarihi	Haziran 2010	Eylül 2011	Mayıs 2012	Ekim 2012	Şubat 2013
Hedef / Kurban	İran'daki SCADA sistemleri	İran'daki SCADA sistemleri	Orta Doğu ülkeleri (Devlet kuruluşları, eğitim kurumları...)	Doğu Avrupa, Batı Avrupa bölgeleri... (Hükümetler, bilimsel araştırma merkezleri...)	Almanya, Ukrayna, Portekiz, Türkiye... (Hükümetler, özel şirketler...)
Amaç	Sistemi ele geçirme ve çökertme	Bilgi toplama	Bilgi toplama	Bilgi toplama	Bilgi toplama
Etkileri	Nükleer yakıt tesisine sızarak arızalara sebep olması	SCADA sistemleri ile ilgili kritik bilgilerin tespiti	Bulaştığı çok sayıda bilgisayardan veri sızdırması	Oldukça geniş alanlarda gizli bilgileri toplaması	Bulaştığı çok sayıda bilgisayardan veri sızdırması
Keylogger Özelliği	Hayır	Evet	Evet	Evet	Evet

Tablo 1 - Farklı GIT'lerin karşılaştırılması [9,10]

IV. SONUÇ VE TARTIŞMA

Bu çalışmada, siber dünyada görülen ve siber savaş için kullanılan GIT adını alan uygulamalar incelenmiştir. Bu uygulamaların sahip olduğu çalışma yapıları, oluşturduğu etkiler, gösterdikleri faaliyetler ve tipik özellikleri hakkında farkındalık oluşturulmaya çalışılmıştır. Ele alınan Stuxnet, Duqu, Flame, Red October ve MiniDuke uygulamaları belirlenen özelliklere göre karşılaştırıldığında ortak olan ve farklı olan durumlarla karşılaşılmıştır.

GIT uygulamalarının ilk sızma şekillerine bakıldığında MS Word güvenlik açıklığının Duqu ve Red October uygulamalarında kullanıldığı, Stuxnet ve Flame hakkında kesin bilginin bulunamadığı ve MiniDuke için PDF güvenlik açığının kullanıldığı görülmektedir. Şifreleme yöntemleri olarak XOR algoritmasının ele alınan uygulamalar için ortak olduğu ve buna ek olarak kullanılan şifrelemelerin de olduğu bilinmektedir. Kendi kendine çoğalma becerisine Stuxnet dışındaki uygulamaların sahip olmadığı ve her uygulamanın tespit edilme tarihinin farklı olduğu anlaşılmaktadır. Hedef olarak görülen alanların değişkenlik gösterdiği ve genel olarak devlet kurumlarına ve alt yapı sistemlerine saldırı yapıldığı elde edilen sonuçlar arasında yer almaktadır. Uygulamaların amaçları ve keylogger özelliğine sahip olup olmamaları karşılaştırıldığında ise diğerlerine göre Stuxnet uygulaması farklılık göstermektedir.

Siber saldırıların gelecekte evrensel olarak daha kolay yöntemlerle yapılacağı ve etkilerinin daha da artacağı tahmin edilmektedir. Siber savaş aracı olarak kullanılan GIT'lere karşı tam koruma sağlayabilmek mümkün olamayabilir fakat gerekli önlemlerin alınması ve oluşabilecek etkilerin azaltılması sağlanabilir. Örneğin kurum veya kuruluşlar için doğrudan internete bağlanılmadan önce, kullanılacak ağın denetlenmesi yapılabilir. Güvenilir ve onaylanmış kaynaklar dışındaki kaynaklardan yazılım indirilmesi engellenebilir. Geleneksel korunma yöntemlerinin yeterli kalmadığı durumlar olduğu için yeni nesil çözümler kullanan yazılımlar tercih edilebilir. Çalışanların olduğu kadar bireysel kullanıcıların da bilinçlendirilmesi sağlanabilir. Çok ayrıntılı ve kapsamlı özelliklere sahip olan GIT'ler ile ilgili yapılan çalışmalar arttırılarak ülkemiz için yeni araştırmacıların bu konularda yetiştirilmesi sağlanabilir.

KAYNAKÇA

[1] FireEye Inc., "FireEye Advanced Threat Report: 2013," The FireEye Threat Prevention Platform, Special Report, 2013.

[2] Joint Task Force Transformation Initiative, "Managing Information Security Risk: Organization, Mission, and Information System View," NIST Special Publication, (800-39), 800-39, 2011.

[3] E. M. Hutchins, M. J. Cloppert, ve R. M. Amin, "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains," Leading Issues in Information Warfare & Security Research, 1, 80, 2011.

[4] P. Chen, L. Desmet, ve C. Huygens, A study on advanced persistent threats. Berlin: Springer-Heidelberg, 2014, pp. 63-72.

[5] Mandiant Research Lab., A.P.T. Exposing One of China's Cyber Espionage Units, 2013.

[6] Bencsáth, B., Pék, G., Buttyán, L., & Felegyhazi, M. (2012). The cousins of stuxnet: Duqu, flame, and gauss. Future Internet, 4, 971-1003.

[7] Faisal, M., & Ibrahim, M. (2012). Stuxnet, Duqu and Beyond. International Journal of Science and Engineering Investigations, 2, 75-78.

[8] Kara, M. (2013). Siber Saldırılar Siber Savaşlar ve Etkileri (Doktora tezi, İstanbul Bilgi Üniversitesi).

[9] Virvilis, N., & Gritzalis, D. (2013, September). The big four-what we did wrong in advanced persistent threat detection?. In Availability, Reliability and Security (ARES), 2013 Eighth International Conference. IEEE.

[10] Adebayo, O. S., & AbdulAziz, N. (2014, November). An intelligence based model for the prevention of advanced cyber-attacks. In Information and Communication Technology for The Muslim World (ICT4M), 2014 The 5th International Conference. IEEE.

[11] CodingUnit Programming Tutorials. (2010, 10 Mart). Exclusive-OR (XOR) Encryption. Erişim Tarihi: 1 Temmuz 2015, http://www.codingunit.com/exclusive-or-xor-encryption.

[12] Wikipedia. (2014, 2 Haziran). XOR Cipher. Erişim Tarihi: 1 Temmuz 2015, https://en.wikipedia.org/wiki/XOR_cipher

[13] Wikipedia. (2015, 5 Mayıs). Drcrypt. Erişim Tarihi: 2 Temmuz 2015, https://tr.wikipedia.org/wiki/Drcrypt

[14] İTÜBİDB. (2013, 7 Eylül). Şifreleme Yöntemleri. Erişim Tarihi: 2 Ağustos 2015, http://bidb.itu.edu.tr/seyirdefteri/blog/2013/09/07/şifreleme-yöntemleri

[15] Cryptography Beta.(2014, 4 Ekim). How is XOR used for encryption? Erişim Tarihi: 2 Temmuz 2015, http://crypto.stackexchange.com/questions/19470/how-is-xor-used-for-encryption

[16] Bencsáth, B., Pék, G., Buttyán, L., & Félegyházi, M. (2012, April). Duqu: Analysis, detection, and lessons learned. In ACM European Workshop on System Security (EuroSec) (Vol. 2012).

[17] Thakur, Vikram. Symantec Official Blog. (2011, 1 Kasım). Duqu: Status Updates Including Installer with Zero-Day Exploit Found. Erişim Tarihi: 20 Temmuz 2015, http://www.symantec.com/connect/w32-duqu_status-updates_installer-zero-day-exploit

[18] Symantec Security Response. W32.Duqu: The precursor to the next Stuxnet. Technical Report Version 1.4, Symantec, 23 Kasım 2011.

[19] NIST, FIPS PUB 197. Advanced Encryption Standard (AES). Kasım 2001.

[20] Dworkin, M. Recommendation for Block Cipher Modes of Operation: Methods and Techniques. NIST Special Publication 800-38A, Aralık 2001.

[21] S. Frankel, R. Glenn, NIST, S. Kelly. Network Working Group. (2003, Eylül). The AES-CBC Cipher Algorithm and Its Use with IPsec. Erişim Tarihi: 3 Temmuz 2015, https://tools.ietf.org/html/rfc3602#section-2.1

[22] sKyWIper Analysis Team. sKyWIper (a.k.a. Flame a.k.a. Flamer): A complex malware for targeted attacks. Technical Report Version 1.05, CrySyS Lab, Budapest University of Technology and Economics Department of Telecommunications, May 31 2012.

[23] Gostev, A. Securelist Official Blog. (2012, 28 Mayıs). The Flame: Questions and Answers. Erişim Tarihi: 26 Haziran 2015, https://www.securelist.com/en/blog/208193522/The_Flame_Questions_and_Answers

[24] K. Avi. Lecture 8: AES: The Advanced Encryption Standard. Lecture Notes on "Computer and Network Security", Purdue University. 1 Mayıs 2015.

[25] Evren, Ş. (2008, 21 Şubat). Yerine Koyma Şifrelemesi (Substitution Cipher). Erişim Tarihi: 1 Ağustos 2015, http://bilgisayarkavramlari.sadievrenseker.com/2008/02/21/yerine-koyma-sifrelemesi-substitution-cipher/

[26] Karataş, A. (2013, 28 Eylül). Şifreleme Algoritmaları. Erişim Tarihi: 20 Haziran 2015, https://adnankaratas.wordpress.com/2013/09/28/sifreleme-alogoritmalari/

[27] Evren, Ş. (2008, 17 Nisan). YRC4 Şifrelemesi (RC4 Cipher, ARC4, ARCFOUR). Erişim Tariihi: 3 Ağustos 2015, http://bilgisayarkavramlari.sadievrenseker.com/2008/04/17/rc4-sifrelemesi-rc4-cipher-arc4-arcfour/

[28] Kaspersky Labs' Global Research & Analysis Team. (14 Ocak 2013). The "Red October" Campaign – An Advanced Cyber Espionage Network Targeting Diplomatic and Government Agencies. GReAT Report. Erişim Tarihi: 28 Mayıs 2015, https://securelist.com/blog/incidents/57647/the-red-october-campaign/

[29] Kaspersky Labs' Global Research & Analysis Team. (14 Ocak 2013). "Red October" Diplomatic Cyber Attacks Investigation. GReAT Report. Erişim Tarihi: 28 Mayıs 2015, https://securelist.com/analysis/publications/36740/red-october-diplomatic-cyber-attacks-investigation/

[30] McAfee Labs. Operation Red October. McAfee Labs Threat Advisory Report, 2013.

[31] Kaspersky Labs' Global Research & Analysis Team. (17 Ocak 2013). "Red October". Detailed Malware Description 1. First Stage of Attack. GReAT Report. Erişim Tarihi: 29 Mayıs 2015, https://securelist.com/analysis/publications/36830/red-october-detailed-malware-description-1-first-stage-of-attack/#1

[32] Kaspersky Labs' Global Research & Analysis Team. (17 Ocak 2013). "Red October". Detailed Malware Description 3. Second Stage of Attack. GReAT Report. Erişim Tarihi: 30 Mayıs 2015, https://securelist.com/analysis/publications/36802/redoctober-detailed-malware-description-3-second-stage-of-attack/

[33] Yliluoma, J. (2014, Ocak). Bit mathematics cookbook. Erişim Tarihi: 20 Temmuz 2015. http://bisqwit.iki.fi/story/howto/bitmath/

[34] Atmel Resmi Websitesi. ROR- Rotate Right through Carry. AVR Assembler Instructions. Erişim Tarihi: 17 Haziran 2015, http://www.atmel.com/webdoc/avrassembler/avrassembler.wb_ROR.html

[35] R. Rivest. Network Working Group. (1998, Mart). A Description of the RC2(r) Encryption Algorithm. Erişim Tarihi: 27 Temmuz 2015, https://www.ietf.org/rfc/rfc2268.txt

[36] Raiu C., Soumenkov I., Baumgartner K., Kamluk V. Global Research and Analysis Team., "The MiniDuke Mystery: PDF 0-day Government Spy Assembler 0x29A Micro Backdoor," Technical Report, Kaspersky Lab, 2013.

[37] Kaspersky Labs' Global Research & Analysis Team. (3 Haziran 2014). Miniduke is back: Nemesis Gemina and the Botgen Studio. GReAT Report. Erişim Tarihi: 11 Haziran 2015, https://securelist.com/blog/incidents/64107/miniduke-is-back-nemesis-gemina-and-the-botgen-studio/

[38] Kaspersky Lab. (27 Şubat 2013). Kaspersky Lab Identifies 'MiniDuke', a New Malicious Program Designed for Spying on Multiple Government Entities and Institutions Across the World. Report. Erişim Tarihi: 11 Haziran 2015, http://www.kaspersky.com/about/news/virus/2013/Kaspersky_Lab_Identifies_MiniDuke_a_New_Malicious_Program_Designed_for_Spying_on_Multiple_Government_Entities_and_Institutions_Across_the_World

[39] CrySyS Malware Intelligence Team, Kaspersky Labs GREAT Team. Miniduke: Indicators. Technical Report Version 1.00, CrySyS Lab, Budapest University of Technology and Economics Department of Telecommunications, 2013.

Esra SÖĞÜT, Eskişehir'de doğdu. İlk, orta ve yüksek öğrenimini Eskişehir'de tamamladı. 2012 yılında Eskişehir Osmangazi Üniversitesi Bilgisayar Mühendisliği Bölümü'nden mezun oldu. Bir yıl sonra, Gazi Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı'nda yüksek lisans eğitimine başladı. 2013 yılında Gazi Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü'ne araştırma görevlisi olarak atandı. Halen, Gazi Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü'nde araştırma görevlisidir. İlgi alanları: bilgisayar ağları, kötücül yazılımlar.

Prof. Dr. O. Ayhan ERDEM, Ankara'da doğdu. İlk, orta ve yüksek öğrenimini Ankara'da tamamladı. 1989 yılında Gazi Üniversitesi Fen Bilimleri Enstitüsü'nden Yüksek Lisans, 2001 Yılında doktora derecesi aldı. Amerika Birleşik Devletleri, 1990 yılında İndiana Üniversitesinde yoğun İngilizce eğitimini, Purdue Üniversitesinde Bilgisayar Teknolojisi Eğitimini tamamladı. Bilgisayar programlama dilleri, bilgisayar ağları, temel bilgi teknolojileri, bilgisayar sistemleri konularında çok sayıda kitapları, uluslararası ve ulusal dergilerde makaleleri bulunmaktadır. Halen Gazi Üniversitesi Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümünde Profesör ünvanlı öğretim üyesi olarak çalışmaktadır. Evli ve üç çocuk babasıdır.

DERIVING PRIVATE DATA IN VERTICALLY PARTITIONED DATA-BASED PPCF SCHEMES

Murat Okkalioglu, Mehmet Koc, and Huseyin Polat

Abstract — New companies might lack enough data for collaborative filtering. This problem can be overcome if different companies collaborate by sharing their data. However, these companies should protect their user data. Therefore, companies must take privacy measures to prevent data disclosure. There are some studies in the privacy-preserving collaborative filtering literature focusing vertically partitioned binary rated data for two-party cases. In this study, we focus on such two-party vertically distributed privacy-preserving collaborative filtering schemes and experimentally analyze privacy offered by a specific scheme against three different attack types in the literature. We show that data can be derived through such attacks.

Index Terms — Privacy, collaborative filtering, binary, vertically partitioned data, attack scenarios

I. INTRODUCTION

Amount of information that customers can process has some limits and this limit makes the decision making process difficult. E-companies might collect implicit (browsing, purchase history, time spent, etc.) and explicit data (ratings, reviews, etc.) about their customers for collaborative filtering (CF), which is a technique to offer predictions on user data [1]. A typical CF system is composed of an n×m matrix with n users and m items. The users can express their ratings as binary (like or dislike).

Since CF systems usually have a sparse matrix, data sparsity is a crucial problem [1]. User participation is important to obtain accurate referrals. However, users may be unwilling to participate in providing their true preference due to privacy risks like unsolicited marketing, price discrimination, unauthorized access, and government surveillance [1]. Thus, privacy-preserving collaborative filtering (PPCF) aims to protect privacy and provide accurate referrals [1]. Online vendors might collaborate by sharing their ratings to fill their missing ratings. Two companies could have ratings for the different set of items by the same customers. This type of partitioning is known as vertically partitioned data (VPD). There are different VPD-based PPCF schemes to offer predictions [2, 3].

Although PPCF schemes promise individual privacy, there are some studies arguing that they do not protect the privacy as much as believed [4, 5, 6, 7]. Data perturbed by randomization has predictable nature and data perturbed by this method can be extracted using spectral filtering (SF) [4]. Zhang et al. [5] propose two different techniques to disclose user ratings masked by randomization. The authors in [6] analyze how to derive rated items in PPCF, which are disguised by randomized response technique (RRT). In [7], the authors study three attack scenarios for a two-party horizontally partitioned data (HPD)-based PPCF scheme proposed in [3]. In this study, however, we examine how to derive private data in VPD-based binary PPCF scheme proposed in [3]. Our aim is to show how much privacy can be achieved by a specific VPD-based PPCF binary scheme, where data is partitioned between two parties.

II. RELATED WORK

Polat and Du [8] apply RRTs on binary data to disguise ratings. Kaleli and Polat [9] utilize RRTs to produce predictions on naïve Bayesian classifier (NBC). While the previous studies handle the central server-based scenarios, there are some other studies focusing on partitioned data [2, 3, 10, 11, 12]. The researchers in [10, 11, 12] also propose multi-party schemes. The authors in [10] utilize NBC for HPD- and VPD-based data. Self-organizing maps-based predictions are proposed for HPD- [11] and VPD-based [12] schemes.

Kargupta et al. [4] propose an SF technique to extract the original data perturbed by random perturbation. Their method is based on obtaining theoretical boundaries of maximum and minimum values of eigenvalues of the noise matrix. Some researchers study the bounds of the reconstruction error by SF methods [13]. Principal component analysis is also utilized to reconstruct the original data by exploiting data correlations [14]. When the correlation is high, more accurate reconstructions can be performed for random perturbation. Zhang et al. [5] examine how to derive true data in PPCF schemes. They utilize singular value decomposition and k-means clustering to reconstruct the original data. They apply k-means clustering to get the data in groups for discrete and continuous valued data. They discretize the continuous data into k segment and an item is assigned to the median value of the segment it belongs to after clustering. Binary PPCF scheme in [8] is investigated in terms of disclosing which items are rated [6]. The authors utilize publicly collected information about the target data set and manage to retrieve this private information.

Our approach in this paper is to deal when data is partitioned vertically between two parties with binary data. Okkalioglu et al. [7] study how much privacy is offered when data is partitioned horizontally between two-parties. They utilize possible attack techniques (acting as an active user and knn-based) and propose an attack technique, perfect match attack. We extend our prior study [7] for a VPD-based binary PPCF scheme.

III. PRELIMINARIES

Polat and Du [3] offer a top-N (TN) prediction scheme for the active user, a, among the item list (Na) she wants a prediction. Privacy measures are taken by the parties to overcome possible privacy breaches. From now on, A and B will denote each party. This scheme selects the users who have high positive and negative correlations with a claiming that accuracy might be increased if the best similar and dissimilar users are selected as follows:

$$W_{au} = \frac{t(R_s) - t(R_d)}{t(R)} \quad (1)$$

In Eq. (1), Wau is the similarity weight between the user u and a. t(Rs), t(Rd), and t(R) are the number of similarly, dissimilarly, and commonly rated items by both u and a, respectively.

After determining Wau, neighbors are picked on best-N or threshold methods. In the best-N neighbors' selection, N number of users with the highest correlations is picked as neighbors. Threshold neighbors' selection method picks its neighbors among the users whose correlations surpassing a threshold (τn) value. Next, number of likes (lj) and dislikes (dj) among the selected neighbors are estimated, where j is item number. Then, ldj = lj - dj is calculated. If ldj > 0, the item will be liked by a. Otherwise, it will be disliked.

Private similarity computation protocol (PSCP) is utilized to privately compute the similarities. PSCP considers the cases where a's incoming query is sparse or dense. Either party who receives the incoming query can apply PSCP. Assume that A applies the protocol as a master party. First, A finds the number of rated items, M, by a and the protocol follows two different path on its density. In the sparse case, if M is less than [m/2], then A finds the unrated items of a. A random number YAa is created by A from the range (1, m-M). YAa number of cells is selected to fill using private default vote computation protocol calculating a default vote for the items on ldj values. The dense case handles the situation, where M is greater than m/2. A uniform random number YAr is selected by A from the range (1, M). Then, A randomly selects YAr items among the M rated items and removes them from a's query. In VPD, target items (Na) that a is looking for a prediction might belong to different parties. This scheme offers two different cases. The first one deals with the case, where all Na items belong to the one of the parties. The second case is designed when Na items are shared between parties. The first case is called Case-All and the second one refers to Case-Split.

Items belong to both parties in VPD. N_a items might belong to the one of parties. Case-All scheme deals with this special case assuming that B has all items [3]: (1) a sends her corresponding ratings to both parties and N_a to only B. A computes the required values to calculate similarities using PSCP. (2) A sends the partial similarity values to B through a. B finds its own partial similarity values between users it holds and a. Then, B calculates the final similarities (Wau) adding the partial similarities received from A to its own calculated ones. (3) B selects the best neighbors. (4) ld_j values are calculated and sorted by B. TN is returned to a. In Case-Split, N_a items are split between parties: (1) a sends a query and her ratings to both parties. B finds the partial similarities between its users and a using PSCP; and sends them to A through a. (2) A computes its own partial similarities. Then, A finds the similarities by adding values from B. (3) A selects the best Nn neighbors. It lets B know which neighbors are selected and the similarity signs. (4) A forms a neighborhood by employing random N_n and τ_n. A computes ld_{Aj} with this new neighborhood and lets B know them. B then calculates ldj values by adding ldAj values from A to ld_{Bj} values. B finally sends TN list to a.

IV. ATTACK SCENARIOS

We describe the possible three attack scenarios that can be applied to VPD-based PPCF schemes; and evaluate their performance in terms of privacy. The first aspect of privacy preserves actual ratings and the second aspect protects rated items.

A. Acting as an active user in multiple scenarios

If there is a malicious party, it can try sending multiple queries to learn the other party's matrix. Consider a case, where the malicious party sends multiple queries and alters only one cell each time. The malicious party can track the changes in the output (similarity weights) and decide the rating value of the items whose value has been manipulated. Assume that the malicious party invokes an initial query and stores the similarity weights. Then, it manipulates a single rating and sends the altered query to the other party to learn the new similarity weights. After receiving similarity weights for the manipulated query, it compares them with the ones from the earlier query. If there is an increase in the similarity weight between a and u, then the manipulated value is kept by the user. The malicious party can reach such a decision because the increase in the similarity weights means a higher correlation between a and u. If there is a decrease, then it concludes that u keeps the value in the first query. If there is no change in the similarity value, this means that the manipulated item is not rated by u. This notion can be applied for each user so that the whole matrix can be disclosed in such a scenario. This attack is a threat for both privacy aspects. Case-All and Case-Split VPD-based PPCF schemes are subjected to this attack. Collaborating parties have to exchange the partial similarity values of individuals to calculate the final similarities for both schemes. This interaction makes it possible to perform this attack because the malicious party will have an access to the similarities for each user. Due to the PSCP, the success of this attack is affected. We examine how effective PSCP is against this attack using trials.

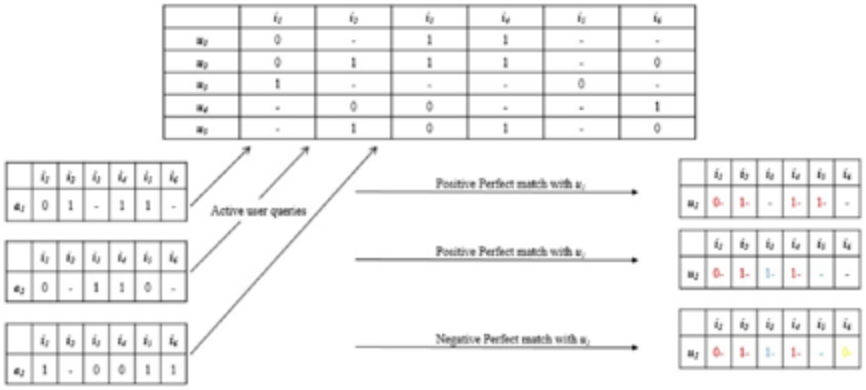


Fig. 1. Perfect match attack

B. knn-based scenarios

This attack is proposed by Calandrino et al. [15]. It assumes that the attacker has a history of ratings of a target user and appends k fake users into the CF system with an exact copy of known history. When a prediction is requested for one of the fake users, it is highly probable that k neighbors will be selected among k-1 fake users and the target user. Since k-1 users are identical, the predictions are expected to come from the target user. This attack discloses the data of whether an item is rated or not, so the second privacy aspect is under threat. Since both VPD-based schemes utilize the best neighbors selected among all neighbors, knn-based attack can be performed. Additionally, the party intending for the attack does not need to have a history of the target user. Parties have already had a history of each user. Thus, the malicious party can use its own part of the ratings as the history of the target user and manipulate neighbors by inserting k fake users. We hypothesize that PSCP is not effective to prevent from this attack because randomly removing or appending ratings into a's vector does not change the similarity weights between a and k-1 users plus the target user.

C. Perfect match attack

This attack is proposed in [7] for an HPD-based binary PPCF scheme. We apply this attack for VPD-based scheme. The attack exploits the similarity value exchanged between parties. Each party calculates the partial similarity values. Although PSCP is applied by each party as a privacy measure, we believe that this scheme is subjected to this attack. Assume that B acts as the master party and no privacy measure is taken. A calculates the similarities between its user and a, sends them to B. If the similarity between any user of A and a is either 1 or -1, such similarities are called as perfect matches [7]. This means that the commonly rated items between these two users are either identical or opposite, respectively. Hence, B can conclude that the corresponding user who has a perfect match with a either identically voted or not voted for any of its items if the similarity is 1. If the similarity is -1, they either vote opposite or not voted for any of the items. The attack is depicted in Fig. 1.

Both VPD-based schemes calculate the similarities by two-party collaboration. One party calculates its own partial similarities and sends them to the master party for similarity calculation. Then, the similarities are calculated by adding up the incoming partial similarity values with the ones calculated off-line. The master party can use these interim similarity values received from the other part to identify perfect matches. Thus, this attack is applicable for both VPD-based schemes.

V. EXPERIMENTS

Experiments are performed using MovieLens Million (MLM) and Jester data sets. MLM contains one million ratings from 6,040 users for 3,952 items. It was collected by GroupLens (www.cs.umn.edu/research/GroupLens). Ratings are on a 5-star scale. Jester is a dense data set and ratings are continuous scale between -10 and 10 (http://eigentaste.berkeley.edu/dataset/). Precision and recall are used as evaluation metrics. Precision is the ratio of how many retrieved items are related. Recall is the ratio of how many relevant items are selected. MLM ratings greater than or equal to 3 are converted to 1 (like) and the rest converted to 0 (dislike). Jester ratings greater than 2 are rated 1 and the rest is rated 0. Experiments are repeated 10 times. 500 and 2,000 users are picked randomly from MLM and Jester, respectively. Jester users are picked among the users with at least 60% rated cells for a denser data set. We display the outcomes in Table 1, where WP means privacy measures are taken while NP means no privacy measures are taken.

A. Experiment I

In this experiment, multiple active queries are sent by the malicious party to derive information. An active query is created with the random density between 10% and 50% and it is sent to the master party. PSCP can mitigate this attack. Thus, we devise the attack against VPD-based schemes with and without PSCP to measure the success of PSCP. Each query sent to the server is altered by PSCP. This means that the partial similarities returned from the other party does indeed belong to the altered active queries. As a result, it is expected that PSCP would be an effective privacy measure because each subsequent queries will be almost independent from each although they are distinct with only one item.

As seen in Table 1, precision results are perfect in NP case. Every item that this attack can recover obviously belongs to the original data. Recall rates denote the percentage of these recovered items to the total relevant items including the ones that are not recovered. We do not attempt to recover every item. Multiple queries whose density is randomly picked between 10% and 50% are sent to the master party. In WP case, it is clear that precision and recall results deteriorate due to PSCP. Precision rates with the sparse data set, MLM, is very low; this is mainly because observing unrated items is more difficult with PSCP. We send an active query by manipulating only one item from the original active query to recover the manipulated item. Two subsequent active queries have to produce the same similarity weight to mark the manipulated item to be recovered unrated. However, marking the manipulated item unrated is a rare possibility since PSCP alters each incoming query based on the density. MLM has many unrated items so it is highly possible that many unrated items are marked rated. Precision calculation is dramatically affected for MLM because the number of actually rated items is very low. Thus, making mistakes by marking unrated items as rated creates great number of false positives for a sparse data set. This possibility is still valid for the dense data set, Jester; yet unrated items are not as many as the sparse case, MLM, and false positive rate remains low. This means that the penalty is greater for sparse data sets in terms of false positive due to sparsity. Recall rates basically indicate that there is a decline of the percentage of recovered items in WP compared to NP due to PSCP's role on altering active queries.

		NP				WP			
		Case-All		Case-Split		Case-All		Case-Split	
		MLM	Jester	MLM	Jester	MLM	Jester	MLM	Jester
Acting as an active user attack	Precision	1.000	1.000	1.000	1.000	0.028	0.592	0.030	0.760
	Recall	0.293	0.288	0.603	0.760	0.205	0.191	0.380	0.690
knn-based attack	Precision	0.178	0.828	0.136	0.808	0.204	0.807	0.085	0.801
	Recall	0.374	0.716	0.368	0.706	0.245	0.537	0.259	0.703
Perfect match attack	$Precision_1$	0.037	0.913	0.034	0.894	0.010	0.761	0.010	0.648
	$Recall_1$	0.173	0.443	0.189	0.441	0.017	0.042	0.017	0.032
	$Precision_2$	0.976	0.157	0.978	0.193	0.992	0.251	0.992	0.373
	$Recall_2$	0.331	0.514	0.344	0.537	0.098	0.077	0.086	0.111

Table I - Empirical outcomes for the three attacks

B. Experiment II

This experiment performs knn-based attack, which requires history of a user and the parties in VPD-based schemes have such information. We set k as 100 and 100 fake users are appended into the data set for each test. This attack asks for a TN prediction for one fake user to pick the best k neighbors among the k-1 fake users and the target user whose history is known. There might be more than k best neighbors, then we pick the first k's. Thus, this attack could be less efficient if there are some other highly correlated users other than appended k-1 users. Calculation of the similarity metric is an important factor here; it is computed by only considering the commonly rated items. Assume that a user with only one item rated could have the perfect similarity with the target user if the target user rated that item. Therefore, such a user could join among the best neighbors by only one cell. knn-based attack has such a limitation for VPD-based schemes due to the similarity metric. However, PSCP might not introduce significant privacy for the same reason. Appending or removing random entries into/from the query will have no effect on the similarity calculations with k-1 fake users. Since there is no change in the fake users vector, altering the active query has no effect on the commonly rated entries. We do not expect a clear trend in WP case in both precision and recall rates. Consequently, PSCP is not expected to introduce a level of privacy.

There is no clear trend in both precision and recall rates as seen in Table 1 for this attack because the similarity calculation is determined by only commonly rated cells. All k-1 fake users and the target users in NP case are still among the best neighbors after PSCP (WP). There might be some users joining into or staying out of the best neighbors if PSCP removes some items from the active query. A user in the best neighbors can be kept out of the best neighbors if there remain no commonly rated items after PSCP. A new user might join into the best neighbors if all commonly rated items match after some items are removed from the active query by PSCP. Note that k-1 fake users and the target user similarity are not affected because their commonly rated items remain same. However, the other users among the best users might have some opposite rating for the appended items in the active query by PSCP. Thus, those users are kept out. Since k-1 fake users and the target user are not kept out of the best neighbors due to PSCP, there is no clear trend in terms of precision and recall. The success of the attack can vary based on how the nominated users for the best k neighbors change. It can be concluded that PSCP does not offer a level of privacy as hypothesized.

C. Experiment III

Perfect match attack is fulfilled here. Its success relies on finding perfect matches between the users and the query. Thus, we empirically test how much density would be better to capture more perfect matches. We perform tests with active queries of 5%, 10%, 20%, and 50% density rates without employing privacy measures. 5% is chosen as the best density rate for this attack. This attack reveals two privacy breaches. We evaluate different precision and recall metrics for these two breaches: (1) either the actual value of the relevant item or it is unrated and (2) unrated entries. The first precision1 and recall1 measure the first breach. Since there are two possibilities in this case, we perform a coin toss to decide. Precision2 and recall2 measure the success of the second breach. It gives statistics about the items which are successfully marked unrated.

For the first breach, a clear effect of PSCP can be seen in Table 1. When PSCP is applied, we see certain decrease in all cases. Although PSCP helps privacy compared to the NP case, precision1 for Jester in WP is not negligible. Recall1 shows that the amount of data disclosed can be considered very low in WP. Jester reveals more privacy breaches than MLM in NP and WP cases. Sparse data possibly captures perfect matches on small number of commonly rated items and the other rated items in the active query are either marked 1's, 0's or unrated in the user vector. This process deteriorates results. In terms of the second breach, results are better for MLM due to sparsity. However, WP case reveals significant amount of data for MLM. Recall2 are too low for both sets while precision2 are significant for MLM.

VI. CONCLUSION

We study three attack techniques that can be applied to VPD-based binary PPCF schemes proposed by Polat and Du [3]. Acting as an active user attack in multiple scenarios is very effective without privacy measures; however, PSCP provides privacy in terms of precision for sparse data set. knn-based attack performs similar results when privacy measures are taken and not taken. Results for perfect match attack displays that it is effective to exploit the first and

second privacy breaches for denser and sparse data sets, respectively. PSCP has been originally designed for acting as an active user in multiple scenarios attack and it is successful to mitigate its effect. However, PSCP is not very useful for knn-based attack due to similarity weighting calculation.

As a future goal, we would like to consider if we can improve privacy provided by these schemes against three attacks. Multi-party PPCF schemes will be examined in terms of privacy breaches.

REFERENCES

[1] A. Bilge, C. Kaleli, I. Yakut, I. Gunes, and H. Polat, "A Survey of Privacy-Preserving Collaborative Filtering Schemes," Int. J. Softw. Eng. Knowl. Eng., Vol. 23, No. 08, pp. 1085-1108, 2013.

[2] Polat and W. Du, "Privacy-Preserving Collaborative Filtering on Vertically Partitioned Data" Lect. Notes Comput. Sci., Vol. 3721, pp. 651-658, 2005.

[3] H. Polat and W. Du, "Privacy-Preserving Top-N Recommendation on Distributed Data," J. Am. Soc. Inf. Sci. Technol., Vol. 59, No. 7, pp. 1093-1108, 2008.

[4] H. Kargupta, S. Datta, Q. Wang, and K. Sivakumar, "On the Privacy Preserving Properties of Random Data Perturbation Techniques," in Proceedings of the 3rd IEEE International Conference on Data Mining, Melbourne, FL, USA, 99-106, 2003.

[5] S. Zhang, J. Ford, and F. Makedon, "Deriving Private Information from Randomly Perturbed Ratings," in Proceedings of the 6th SIAM International Conference on Data Mining, Bethesda, MD, USA, 59-69, 2006.

[6] M. Okkalioglu, M. Koc, and H. Polat, "On the Discovery of Fake Binary Ratings," in Proceedings of the 30th Annual ACM Symposium on Applied Computing, Salamanca, Spain, 901-907, 2015.

[7] M. Okkalioglu, M. Koc, and H. Polat, "On the Privacy of Horizontally Partitioned Binary Data-based Privacy-Preserving Collaborative Filtering" in Proceedings of the 10th DPM International Workshop on Data Privacy Management, Vienna, Austria, 2015.

[8] H. Polat and W. Du, "Achieving Private Recommendations Using Randomized Response Techniques," Lect. Notes Comput. Sci., Vol. 3918, pp. 637-646, 2006.

[9] C. Kaleli and H. Polat, "Providing Private Recommendations Using Naïve Bayesian Classifier," Advances in Soft Computing, Vol. 43, pp. 168-173, 2007.

[10] C. Kaleli and H. Polat, "Privacy-Preserving Naïve Bayesian Classifier-based Recommendations on Distributed Data," Comput. Intell., Vol. 31, No. 1, pp. 47-68, 2015.

[11] C. Kaleli and H. Polat, "Privacy-Preserving SOM-based Recommendations on Horizontally Distributed Data," Knowledge-Based Syst., Vol. 33, pp. 124–135, 2012.

[12] C. Kaleli and H. Polat, "SOM-based Recommendations with Privacy on Multi-party Vertically Distributed Data," Journal of the Operational Research Society, Vol. 63, No. 6, pp. 826-838, 2012.

[13] S. Guo, X. Wu, and Y. Li, "Determining Error Bounds for Spectral Filtering based Reconstruction Methods in Privacy Preserving Data Mining," Knowl. Inf. Syst., Vol. 17, No. 2, pp. 217-240, 2008.

[14] Z. Huang, W. Du, and B. Chen, "Deriving Private Information from Randomized Data," in Proceedings of the 24th ACM SIGMOD International Conference on Management of Data, 37-48, 2005.

[15] J. A. Calandrino, A. Kilzer, A. Narayanan, E. W. Felten, and V. Shmatikov, "``You Might Also Like:" Privacy Risks of Collaborative Filtering," in Proceedings of the IEEE Symposium on Security and Privacy, Oakland, CA, USA, 231-246, 2011.

SOSYAL MEDYA VERİLERİ ÜZERİNDEN SİBER İSTİHBARAT FAALİYETLERİ

S. Savaş, N. Topaloğlu

Özet — Toplumsal olayların önceden tespit edilebilmesi veya gerçekleşen olaylarda geriye dönük araştırma yapılabilmesi için, sosyal medya verileri çok önemli bir istihbarat ortamı haline gelmiştir. Siber istihbarat günümüzde gittikçe daha da önem kazanmaktadır. Ülkeler istihbarat kurumlarında siber istihbarat birimleri oluşturmaya başlamıştır. Gereksinimler doğrultusunda farklı program ve algoritmalar kullanılarak, sosyal medya veri yığınlarından istendik sonuçlara ulaşmak mümkündür. İstihbarat açısından çok önemli olan sosyal medya siteleri ticari, akademik veya güvenlik amaçlarıyla analiz edilebilir. Bu çalışmada Türkiye'de sosyal medya üzerinde çok tartışılan bir olay analizi gerçekleştirilmiştir. Elde edilen veriler içerisinden bilgi çıkarımı ve görselleştirme işlemleri yapılmıştır. Bazı anlamlı bilgilere ulaşılmıştır. Türkiye'de TT listesinde pek çok paylaşım reklam amaçlı yapılmaktadır. Gerçek verilere ulaşmak için öncelikle veriler içerisinde temizleme işlemi gerçekleştirilmiştir. Veriler URL'lerden ve etiketlerden temizlendiğinde daha anlamlı bilgiler ortaya çıkmıştır. Verileri analizi sırasında görülmüştür ki toplum hafızası, benzer olaylar arasında köprü kurabilmektedir. Ayrıca etiketli verilerde de yüksek merkezilik oranları göstermiştir ki bazı kullanıcılar ve twitler, pek çok farklı kullanıcı arasında köprü görevi görmektedir.

Anahtar Kelimeler — Siber güvenlik, siber istihbarat, istihbarat, nodexl, sosyal medya analizi

Abstract — Social Media data became very important to identify community events before or to make retrospective research on actual events. Cyber intelligence is gaining importance more today. Countries began to create cyber intelligence units in their intelligence departments. Reaching needed information from social media data stacks is possible with using different programs and algorithms in accordance with requirements. Social media sites which are very important for intelligence can be analyzed with commercial, academical or security purposes. In this study, an event which is discussed very much in Turkey on social media was analyzed. Accesing information and visualization was made on data. Some meaningful results were reached. In Turkey TT list, plenty of sharing is done for advertising purposes. First, data cleaning process was performed in the data to reach the actual data. When data was cleaned from URL's and hashtags, more meaningful information were accessed. During analysis, it has been seen that community memory is able to establish a bridge among similar events. Also high centrality of data which includes mentions showed that some users and twits are serving like bridges between a lot of different users.

Index Terms — Cyber security, cyber intelligence, intelligence, nodexl, social media analysis

I. GİRİŞ

Her geçen gün artan sosyal medya kullanımıyla birlikte dünya üzerinde daha fazla insan birbiriyle etkileşim halinde bulunmaktadır. Aslında insanların birbiri arasındaki iletişim ve etkileşimleri daha 1929 yılında Macar yazar Frigyes Karinthy tarafından yayınlanan "Láncszemek" (Zincirler) adlı kısa hikâyesinde "6 Degrees of Seperation" teorisinde dile getirilmiştir. Karinthy'nin teorisine göre dünya üzerindeki herhangi iki kişi arasında en fazla 6 kişi aracılığıyla bir bağlantı bulunmaktadır[1,2]. Daha sonra Amerikalı sosyolog Stanley Milgram 1967 yılında "Small World Experiment" adlı çalışması ile bu teoriyi desteklemiştir. Çalışmasında Milgram, rastgele seçilmiş kişilere Boston'a ulaştırmaları için kartlar vermiştir. Onlardan da ulaşacağı adrese yakın olduğunu düşündükleri tanıdıklarına vermelerini istemiştir. Çalışma sonunda görülmüştür ki kartlar ortalama 6 adımda istenilen yere ulaşmıştır[1,2]. Sonralarda 6 adım teoremi gittikçe önem arz etmiştir. 1996 yılında Virjinya Üniversitesi'nde okuyan Brett C. Tjaden tarafından merkeze Kavin Bacon'u yerleştiren ve yaklaşık 3 milyon aktör ve aktrisin birbirleriyle aralarındaki bağları gösteren http://oracleofbacon.org/ sitesi oluşturulmuştur. Bu site de teoremin yaygınlaştırılmasına büyük katkı sağlamıştır. Son olarak MSN, 180 milyon kullanıcısının 30 milyar üzerindeki elektronik posta trafiği incelediğinde herhangi iki insanın birbirinden ortalama 6,6 derece uzakta olduğunu bulmuştur[3,4]. Dünya üzerinde insanlar birbiriyle bu kadar yakın etkileşimde olunca sosyal medya sitelerinin de kitleleri harekete geçirmek etkisi artmaktadır. Dünyaca ünlü Time dergisi, 2010 yılında Facebook'un kurucusu ve sahibi Mark Zuckerberg'i yılın adamı seçerken, 500 milyon kişilik sanal devletin başkanı, insanların hayatlarını yaratıcı ve iyimser olarak değiştiren kişi olarak adlandırmıştır[5]. Bu başlıkta dahi derin anlamlar yatmakta ve sosyal medyanın insanlar üstündeki etkisine dikkat çekmektedir.

Gün geçtikçe hem sosyal medya siteleri hem de bu siteleri kullanan insan sayısı artmaktadır. Son yıllarda gerek Dünyada gerekse Türkiye'de sosyal medyanın kitleler üzerindeki etkisini ve kitleleri harekete geçirmekteki etkinliğini gösteren bazı olaylar olmuştur. 18 Aralık 2010 tarihinde Tunus'da başlayan ve daha sonra Mısır, Libya, Suriye, Bahreyn, Cezayir, Ürdün ve Yemen'de büyük çaplı, Moritanya, Suudi Arabistan, Umman, Irak, Lübnan ve Fas'ta ise küçük çaplı[6,7,8] olayların yaşandığı Arap Baharı sosyal medyanın etkinliğine gösterilebilecek en önemli olaylardandır. Bu süreçte bahsedilen ülkelerde başta Facebook ve Twitter gibi sosyal medya siteleri olmak üzere, sosyal siteleri kullanan kullanıcı sayıları katlanarak artmıştır. Örneğin Mısır'da Facebook kullanıcı sayısı 5,5 milyondan 8,5 milyona ulaşmış, Libya'da Twitter kullanıcıları ilk aylarda 600Bin artmıştır[9]. Arap Baharının tüm Dünyaya Facebook ve Twitter'ın kitleleri harekete geçirmekteki etkisini ispatlamasının ardından pek çok araştırmacı, sosyal medya sitelerinin kitleler üzerindeki etkilerini araştırmaya yönelik çalışmalar yapmıştır.

Benzer şekilde Türkiye'de 27 Mayıs 2013 tarihinde başlayan ve Gezi Parkı olayları olarak İstanbul'dan tüm Türkiye'ye yayılan olaylar sırasında başta Twitter olmak üzere sosyal medyanın toplum üzerindeki etkisi belirgin şekilde görülmüş, pek çok TV kanalında sosyal medya konulu programlar yapılmış, gazetelerde sosyal medyanın etkilerine yönelik haberler yapılmıştır. Olaylar sürecindeki

bir rapora göre twitterda üç ana etiket(#) ile 8.49 milyon mesajın yayınlandığı görülmüştür. Olaylarla ilgili toplamda 100 milyonun üzerinde mesaj yayınlanmıştır. Olayların başlangıcında Türkiye'de günlük aktif twitter kullanıcısı 1.8 milyon civarında iken, 10 günlük süre içinde bu rakam yaklaşık 9,5 milyona ulaşmıştır[10].

Günümüzdeki bir başka sosyal medya olayı ise Doğu Türkistan olaylarıdır. 2015 yılı Haziran-Temmuz aylarında Facebook ve Twitter üzerinde paylaşılan, Çin'in Doğu Türkistan'da katliam yaptığına yönelik haberler üzerine Türkiye'de tepkiler giderek çoğalmıştır. Bu tepkiler sokakta Çinli olduğu düşünülen insanları darp etmeye kadar gitmiştir[11,12].

Türkiye ve Dünyada bunlara benzer örnekler çoğaltılabilir. Bu olaylar ülkelerin gerek devlet istihbaratı gerekse ticari amaçlı istihbarat anlamında sosyal medya üzerine yoğunlaşması gereğini ortaya çıkarmıştır. Çünkü sosyal medya günümüzde, sanalın gerçeğe dönüşmesini sağlayan, bir anlamda "hayatın kendisi" haline dönüşmeye başlamıştır.

Bu çalışmada istihbarat kavramı açıklanmış ve sosyal medyanın devlet istihbaratı amaçlı nasıl kullanılabileceği belirtilmiştir. Daha sonra en önemli sosyal medya sitelerinden biri olan Twitter'ın kitleler üzerindeki etkisi araştırılmış ve yapılan çalışmalara örnekler verilmiştir. Twitter üzerinde bir örnek durum incelemesi ve analizi yapılarak sonuçları açıklanmıştır.

II. İSTİHBARAT

İstihbarat, Türk Dil Kurumunun Türkçe Sözlüğüne göre; "Yeni öğrenilen bilgiler, haberler, duyumlar ve bilgi toplama, haber alma" olarak tanımlanmaktadır[13]. İstihbarat ile ilgili yapılan bazı çalışmalardaki tanımlara göre:

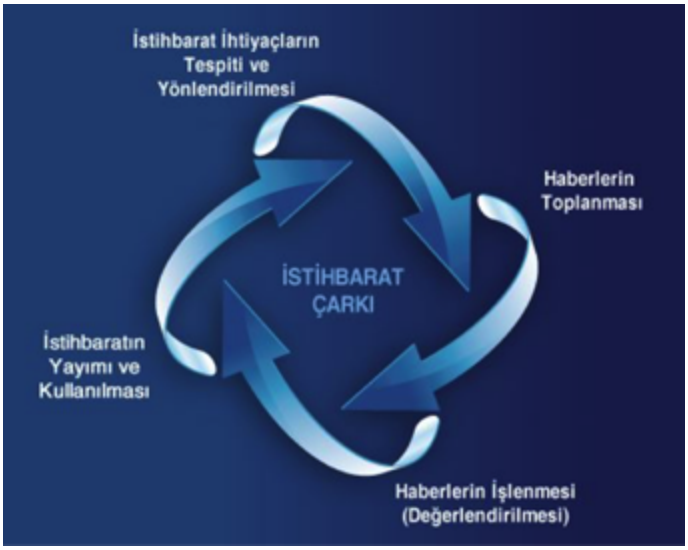
Gültekin Avcı, istihbarata şu şekilde bir tanım getirmektedir: "İstihbarat, muhtelif imkân ve vasıtaları kullanarak, herhangi bir konuda enformatik materyal temini ve temin edilen bilgilerin ham halden kurtarılarak işlenmesi, kıymetlendirilmesi ve yorumlanarak bunlardan bir netice çıkarılmasıyla ilgili faaliyettir."[14].
Ümit Özdağ'ın tanımına göre; istihbarat ulaşılabilen bütün açık, yarı açık ve/veya gizli kaynaklardan her türlü aracın kullanılması sonucunda elde edilen her türlü veri, malumat ve bilginin ulusal genel veya ulusal özel plandaki politikaların gerçekleştirilmesi ve ulusal politikalara zarar verilmesinin engellenmesi amacı ile toplandıktan sonra önemine ve doğruluğuna göre sınıflandırılması, karşılaştırılması, analiz edilerek değerlendirilmesiyle ulaşılan bilgidir[15].

Warner'a göre ise istihbarat; faaliyetleri yönlendirmek üzere önceden bilinmesi gereken her türlü konu ile ilgilenmektedir[16].

Kısaca istihbarat, devletlerin başındaki karar vericilerin ülkelerinin güvenliklerini sağlama, bekalarını koruma, belirsizlikleri azaltma ve çıkarlarının arttırılmasını sağlamada önemli bir role sahiptir[17].

Milli İstihbarat Teşkilatı Müsteşarlığı'na göre, "Devlet İstihbaratı, devletin bütünlüğünü, rejimin emniyetini sağlamak için, millî politika ile tespit edilen millî hedefleri elde etmek üzere devlet organlarının yaptığı istihbaratın tümüdür. Başka bir ifadeyle, Millî Güvenlik Politikaları'nın oluşturulması için gerekli bilgileri sağlayan ve ilgili bütün devlet istihbarat kuruluşlarının işbirliği ve koordinasyonu ile üretilen istihbarattır."[18].

Bu tanımlamalardan, istihbarat faaliyetleri ile ilgili sadece devletlerin değil, şirketlerin, reklam ajanslarının, şahısların da rakipleri ile ilgili bilgi edinip bu bilgileri değerlenderip analiz ederek gelecekle ilgili hazırlıklar yaptığı anlaşılmaktadır.



Şekil 1 - MİT – İstihbarat Çarkı[19]

21. yüzyılda bunların dışında istihbarata eklenen bir bilgi boyutu bulunmaktadır. Siber uzay olarak adlandırılan bu boyut, bilgi çağı teknolojisinin ulusal güvenliğe hem ödülü hem cezası sayılabilir. Bu alanda internet, hem bilgi toplama hem de bu bilgiye dayalı operasyonlar yapma olanağı sunmaktadır. Siber mücadele organize suçlarla ve terörle mücadelede önemli bir araç olarak kullanıldığı gibi, bu örgütler tarafından ulusal güvenlik kurumlarının veri tabanlarına ulaşmak ve stratejik bilgileri elde etmek için de kullanılmaktadır. Siber mücadelenin bilgi çağının ilerlemesi ile diğer mücadele alanlarını daha fazla etkilemeye başlayacağı öngörülmekte ve ulusal güvenliğin sağlanması için stratejik olarak ele alınması gerekmektedir[20].

III. SOSYAL MEDYA VE İSTİHBARAT

İnternet kullanımın yaygınlaşmasıyla birlikte sanal ortamda pek çok website alternatifleri oluşmaya başlamıştır. Bu siteler içerisinde ise günümüzde en popüler olanları sosyal medya siteleri olmuştur. İnsanlar birbirleriyle iletişim ağları oluşturmakta, yeni kişiler tanımakta, düşüncelerini sanal ortamlarda paylaşmakta, resim, müzik, video paylaşımları yapmakta, nerede olduğunu, ne yediğini, ne içtiğini, ne yaptığını paylaşmaktadır. Bu kadar yoğun bilginin aktığı ortamlarda, bu bilgiyi yararlı veya zararlı, ticari veya bireysel kullanmak isteyen kişilerin olması kaçınılmazdır. Burada devreye sibergüvenlik kavramı girmektedir. Siber güvenlik, "siber ortamda, kurum, kuruluş ve kullanıcıların varlıklarını korumak amacıyla kullanılan araçlar, politikalar, güvenlik kavramları, güvenlik teminatları, kılavuzlar, risk yönetimi yaklaşımları, faaliyetler, eğitimler, en iyi uygulamalar ve teknolojiler bütünüdür"[21] şeklinde tarif edilebilir.

Siber İstihbarat: İstihbaratın faaliyet alanları; devletin kontrol fonksiyonundan ötürü tehdidin seviyesine göre yakın ve

uzak tehlikelerin engellenmesi amacıyla karar vericilere bilgi desteği sağlamak, propaganda, psikolojik harekât gibi örtülü operasyon yöntemleri ile olayları yönetmek ve düşman veya muhtemel düşmanın istihbarat faaliyetlerini engellemek olduğu dikkate alındığında, siber uzayda bu amaçlı yapılan faaliyetler bütünü "siber istihbarat" olarak kavramsallaştırılabilir[22]. Devlet istihbaratında durum böyle iken, siber istihbarat kavramı pek çok farklı amaçla da kullanılabilir. Ticari, akademik ve güvenlik amaçları, bu amaçların birkaçını oluşturmaktadır.

Ticari siber istihbarat, günlük hayatımızda her an karşımıza çıkabilmektedir. Arama motorlarında arama yaptıktan sonra, sosyal medya sayfamızda benzer ürünlerin önerilerinin görülmesi büyük veri(big data) disiplininin konusu olduğu kadar, ticari siber istihbaratın da konusudur. "Serkan google'da antivirüs arattı" bilgisi bir istihbarattır. Çünkü istihbarat en genel tabiriyle –kaydadeğer bir bilgi-dir. Bu istihbarat sonrasında facebook sayfasında Serkan'a çeşitli online alışveriş sitelerinden antivirüs önermek ise bu siber istihbaratın ticari olarak kullanılmasıdır. Bu şekilde mikro siber istihbaratlardan başka, firmalar veya ekonomi dünyası için yapılmış makro siber istihbarat çalışmaları bulunmaktadır. 2011 yılında Twitter üzerinde J. Bollen, H. Mao ve X. Zeng tarafından Dow Jones Industrial Average (DJIA) şirketlerinin hisse senetlerinin oranlarını tahmin için bir çalışma gerçekleştirilmiştir. Çalışmada günlük kapanış değerleri ile insanların Twitter'daki tutumları kullanılmış ve %86,7 oranla günlük artış/düşüş değerleri tahmin edilebilmiştir[23]. Benzer bir çalışma 2011 yılında X. Zhang, H. Fuehres ve P. A. Gloor tarafından Dow Jones, NASDAQ ve S&P 500 üzerinde gerçekleştirilmiştir. İnsanların Twitter'da paylaştıkları endişeleri ve düşünceleri üzerinden 6 aylık bir veri toplama aşaması sonrasında yapılan çalışmada, hisseler ve kişilerin duyguları üzerindeki korelasyonlar hesaplanarak açıklanmıştır[24]. 2012 yılında R. Agnihotri ve arkadaşları, sosyal medyanın satış ve pazarlamada nasıl kullanılabileceğini göstermek üzere bir çalışma gerçekleştirmişlerdir[25]. M. M. Mostafa tarafından 2013 yılındaki başka bir ticari sosyal istihbarat çalışmasında Twitter kullanıcılarının, Nokia, T-Mobile, IBM, KLM ve DHL gibi büyük firmalar hakkındaki görüşleri analiz edilmiştir. Bu çalışmada firma karar alıcılarına, kullanıcı görüşlerini kararlarda göz önünde bulundurma imkanı sunulmuştur[26].

Akademik siber istihbarat çalışmaları, siber ortamlardaki verilerin analiz edilerek elde edilen bilgilerle yeni çalışmaların önünü açmak ve siber dünyada akan verinin potansiyellerini göstermek için yapılan çalışmalardır. 2014 yılında X. Tang ve C. C. Yang, sosyal medya verileri içindeki gizli bilgileri ortaya çıkarmak için iki aşamalı bir sistem önermişlerdir. Bu çalışmada Dinamik Stochastic Blockmodel ve Geçici Dirichlet Süreci aşamaları 3 farklı test grubuna uygulanmış ve klasik algoritmalardan daha verimli sonuç üreten bir sistem ortaya çıkmıştır[27]. 2014 yılında A. Weichselbraun, S. Gindl ve A. Scharl tarafından, sosyal medya verilerinde anlamsal bilgilerin keşfine yönelik yeni bir yöntem önerilmiştir[28]. 2014 yılında M. C. Yang ve H. C. Rim tarafından yazılan Twitlerin konulara göre popülerliğini tespit amaçlı bir çalışma gerçekleştirilmiş sonuçları açıklanmıştır[29]. Sosyal medya verileri üzerinde benzer şekilde gerçekleştirilen pek çok akademik çalışma bulunmaktadır. Kullanıcı sayıları ve sosyal medya sitelerine ilgi arttıkça, bu ortamda yapılan akademik çalışma sayısı da artmaktadır.

Güvenlik amaçlı siber istihbarat çalışmaları günümüzün en önemli konularından biri olmuştur. Sosyal ağların günlük hayata büyük oranda yansımasının ardından bireysel, kurumsal ve devlet güvenliği konularına siber düzeyde çözüm gereksinimi ortaya çıkmıştır. Bireylerin veya kurumların güvenliklerinin sağlanması, olası dış tehditlere karşı savunma anlamına gelmektedir. Gerçek hayatın kendisi durumuna dönüşmekte olan sanal dünyada bireylerin ve kurumların gerçek hayatta karşılaşabileceği tehlikelerin benzerleri bulunmaktadır. Devlet güvenliği açısından ise siber güvenlik bazı farklılıklar göstermektedir. Bu farklılık, yapılan saldırı ve sızmaları önleme olduğu gibi siber istihbarat ile olası ihtimalleri önceden tahmin, teşhis ve önleme de olabilir. Sosyal medya sitelerinin kitleler üzerindeki etkileri dünya üzerinde özellikle Arap Baharı olaylarından sonra daha da dikkatle incelenmeye başlamıştır. 2012 yılında Sir D. Omand, J. Bartlett ve C. Miller, Sosyal Medya İstihbaratına giriş adında bir yayın ile istihbarat yaklaşımlarına yeni bir istihbarat türünü eklediklerini duyurmuş ve bu istihbaratın nasıl yapılacağını çalışmalarında açıklamışlardır[30].

Sosyal medya siteleri ülkelerdeki faaliyetleri için o ülkelerde çeşitli anlaşmalar yapmaktadır. Bu anlaşmalarla devletler sosyal medyada kendilerine bazı faydalar sağlamaktadır. Buna örnek olarak Türkiye'de son çıkan internet yasası da gösterilebilir. Ülkeler sosyal medya sahipleri ile yaptıkları anlaşmalar sayesinde istedikleri kişilere yönelik istihbarat elde edebilmektedirler. Her ne kadar da Facebook ve Twitter gibi sosyal medya devlerinin katı kuralları olsa da, bazı durumlarda devletler doğrudan veya dolaylı yollarla bu sitelerden istediklerini alabilmektedirler. "Sosyal medya takip şirketlerinden Visible Technologies, CIA'in girişim sermayesi firması olan In-Q-Tel'den finansman sağlamaya başlamıştır. Batı'daki birçok istihbarat şirketi de daha detaylı internet kullanıcı bilgilerini incelemeyi sağlayacak internet teknolojileri geliştirmek için bütçe ayırmaktadır[31]. Arap Baharı sürecinde, ülkelerde pek çok kişi sosyal medya aracılığıyla ayaklanma önderliği yapmaktan dolayı yargılanmıştır. Ülkemizde de gezi olayları sonrasında sosyal medya aracılığıyla kitleleri harekete geçirmekten dolayı çeşitli gözaltılar ve yargılamalar olmuştur.

Günümüzde sosyal medyanın ulaştığı güç seviyesi göz önüne alındığında, devletler tarafından sosyal medya verileri üzerinde istihbarat çalışmaları yapmanın artık kaçınılmaz olduğu anlaşılacaktır. Bu çalışmalar gerek sitelerle belirli anlaşmalar çerçevesinde gerekse devletlerin kendi yöntem ve teknikleriyle yürütülebilmektedir. Burada esas olan kitlelerin hareketlerini önceden tespit ederek olası tehlikeleri eylem haline dönüşmeden önce engelleyebilmek ve/veya olası şüpheli kişilerin takibi ve engellenmesidir. Toplum ve devlet güvenliği için ülkelerin sosyal medya verilerine kayıtsız kalması mümkün değildir.

Bu çalışmada, Türkiye'de kadına şiddet ve bunun karşısında toplumsal tepkilerin ölçümüne, toplum görüşlerin anlaşılmasına ve karar vericilerin benzer durumlarda toplumun nabzını nasıl tutabileceğine örnek olması açısından bir çalışma gerçekleştirilmiştir.

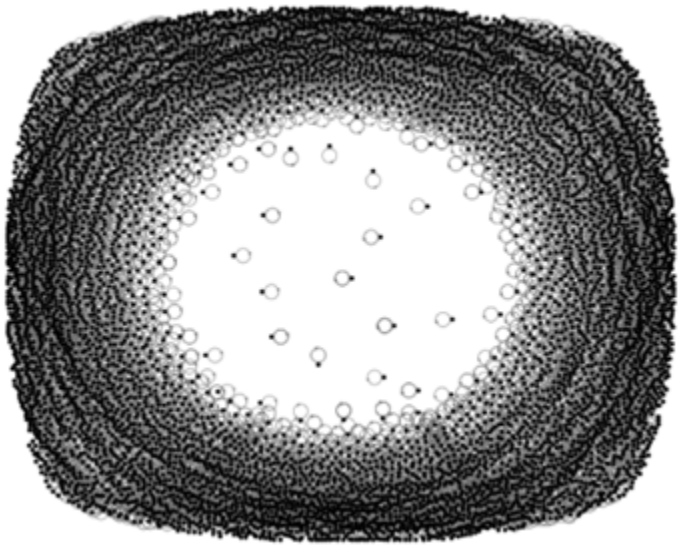
IV. TWİTTER VERİLERİ ÜZERİNDE UYGULAMA

Türkiye'de, üniversite öğrencisi bir genç kız 11 Şubat 2015 tarihinde kaybolmuş ve araştırmalar sonucu kendisinin cinayete kurban gittiği ortaya çıkmıştır. Bu cinayete tüm toplumun tepkisi çok büyük ve etkili olmuştur. Kadına şiddete pek çok ortamda tepki gösterilmiştir. Bu ortamlardan en önemlisi ise sosyal medya olmuştur. Toplum, sanatçılar, politikacılar, sporcular, yerli ve yabancı basın olmak üzere her kesimden bu konuyla ilgili twitler yayınlanmıştır. Toplumun nabzını tutan en önemli sosyal medya sitelerinden olan Twitter üzerinde böyle durumlarda, bu verilerden istihbarat çalışması yapmak, ilgili birimlerin amaçlarından biri olmalıdır. Kitleler sosyal medya mesajları üzerinden harekete geçebilmekte, bireysel veya kitlesel şiddete başvurabilmekte, gösteri veya eylem organize edebilmektedir. Bu süreç içerisinde yazdığı twitler yüzünden ifade vermeye çağrılan veya basında açıklama yaparak düzeltmeler yapmaya çalışanlar dahi olmuştur.

Olayın ortaya çıkmasını takiben twitter dinlenerek 15 Şubat – 15 Mart 2015 tarihleri arasındaki bir aylık süreçte NodeXL[32] temasıyla, twitter üzerinden konu ile ilgili 62273 veri çekilmiştir. Bu veriler üzerinde yapılan analiz sonuçları açıklanmıştır.

Twitter'dan elde edilen veriler öncelikle "Tweet", "Mentions" ve "ReplyTo" olmak üzere üç ayrı gruba ayrılmıştır. Her grup kendi içinde değerlendirilmiştir. Sonuç olarak ortaya, 22626 Twit, 38766 Etiket(bahsetme) ve 796 Cevap verisi çıkmıştır. 85 veri ise boş veya sadece simgesel veri olarak temizlenmiştir.

Yapılan işlemlerden sonra, Fruchterman-Reingo algoritmasıyla her üç veri grubuna ait ağ grafiği Şekil 2,3 ve 4'te görüldüğü gibi olmuştur.



Şekil 2 - Twit verileri ağ grafiği.

Şekil 3 - Etiket verileri ağ grafiği.

Şekil 4 - Cevap verileri ağ grafiği.

Şekil 2'de görülen düğümlerin tamamı kendine dönmektedir çünkü sadece twit verileri işlenmiştir. Şekil 3 ve Şekil 4'te ise etiket ve cevap verileri, kişiler arasındaki bağlantıyı göstermektedir.

Yazılan twitler bütün halinde analiz edildikten sonra içerisinde çok fazla Uniform Resource Locator (URL) bulunması, ülkemizde Trending Topic (TT) listesinin çoğunlukla reklam amaçlı kullanıldığını göstermiştir (Tablo 1).

Top Domains in Tweet in Entire Graph	Entire Graph Count
com.tr	1899
7gundematematik.com	1145
haberler.com	962
google.com	837
cnn.com	748
time.com	574
televizyongazetesi.com	427
feedburner.com	373
sondakikaturk.com	303
nytimes.com	229

Tablo I - En çok kullanılan domain listesi

Bu sonucu destekleyen diğer bir bulgu, en çok kullanılan etiket (# hashtag) tablosunda ortaya çıkmıştır. Tablo 2'de görüldüğü gibi ilk üç sırayı konu dışında bulunan "lys", "aof" ve "ygs" etiketleri almıştır. Bu twitler ve içerisindeki URL'ler incelendiğinde, Tablo 1'de görülen bazı web sitelerin sınav reklamına yönelik twitler olduğu görülmüştür.

Top Hashtags in Tweet in Entire Graph	Entire Graph Count
lys	1026
aof	826
ygs	761
haber	359
özgecanaslan	355
aöf	319
sondakika	304
ozgecan	278
gerçekşuki	146
özgecan	143

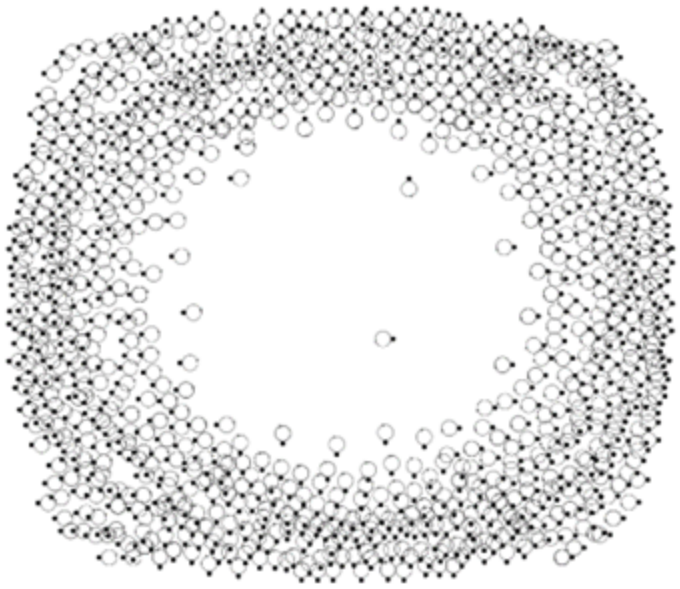
Tablo II - En çok kullanılan etiketler listesi

Anahtar kelime grupları incelendiğinde ise Tablo 3'te görüldüğü gibi gündem konusu ile ilgili kelime grupları belirgin şekilde ortaya çıkmıştır. Buradan anlaşılmaktadır ki pek çok kullanıcı, gündem konuları kelimelerini kullanarak pek çok farklı konuda da paylaşım yapmaktadır.

Top Word Pairs in Tweet in Entire Graph	Entire Graph Count
özgecan,aslan	3419
özgecan,için	2084
turkish,men	823
men,wearing	778
wearing,miniskirts	774
aslan,için	767
lys,ygs	761
aof,lys	734
özgecan,aslan'ın	717
woke,up'	697

Tablo III - En çok kullanılan anahtar kelime grupları (ikili)

Twit verileri daha sonra URL'lerden arındırılmış ve 1919 twit üzerinde analizler gerçekleştirilmiş ve ortaya çıkan ağ grafiği Şekil 5'te görüldüğü gibi olmuştur.



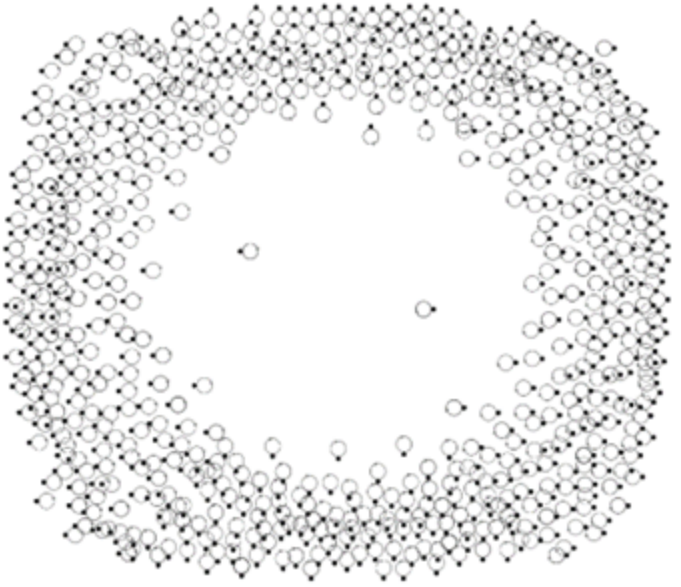
Şekil 5 - URL'lerden arındırılmış twit grafiği

Cinayet ile ilgili açılan etiketler, reklam linklerinden ayrıldığında ortaya ilginç bir durum çıkmıştır. Başka bir şehirde, başka bir nedenle, başka bir cinayete kurban giden genç bir üniversite öğrencisinin ismi de veriler içerisinde görülmüştür. Böyle bir durumda toplum hafızasının olaylar arasında köprüler kurarak tepki gösterebilme olasılığı da görülmüştür (Tablo 4).

Top Hashtags in Tweet in Entire Graph	Entire Graph Count
ozgecan	138
mehmetm3tinkazandırıyor	38
ozgecanaslan	21
celilyamann	21
fıratcakıroğlu	17
ozgecanicinminietekgiy	13
sendeanlat	11
fıratçakıroğlu	11
özgecanaslan	9
news	8

Tablo IV - En çok kullanılan etiketler

Son olarak twit verileri içerisinde # bulunan veriler de temizlenerek 1212 veri elde edilmiş ve ortaya çıkan ağ grafiği Şekil 6'da görüldüğü gibi olmuştur.



Şekil 6 - URL ve #' lerden arındırılmış twit grafiği

Bu verilerden NodeXL ile twitler içerisinde kullanılan kelime grupları sayıları belirlenmiş ve ortaya çıkan anahtar kelime grupları Tablo 5'te görüldüğü gibi olmuştur. haber twitleri ile etiketli twitler temizlendiğinde elde edilen veriler, konu ile en alakalı twit veriler olmuştur.

Top Word Pairs in Tweet in Entire Graph	Entire Graph Count
ozgecan,aslan	136
ozgecan,icin	69
ozgecan,için	34
azizacar,celilyaman	33
ozgecan,gibi	22
kutlu,olsun	15
bu,kadar	12
bir,ozgecan	12
ozgecan,ve	10
aslan,icin	9

Tablo V - En çok kullanılan kelime grupları (Twit)

Etiket içeren twit verileri daha karmaşık bir yapı içermektedir. Burada kelimeler üzerinde işlem yapılabildiği gibi, bağlantılar üzerinde de analizler yapılabilmektedir. Etiketli veriler üzerinde yapılan analiz işlemi sonrasında en çok kullanılan kelime grupları Tablo 6'da görüldüğü gibidir.

Top Word Pairs in Tweet in Entire Graph	Entire Graph Count
aylinnazlıaka,özgecanaslan	161
turkey,sendeanlat	135
arkasıyarın,bizeneoldu	113
ozgecanaslan,ozgecanicinminietekgiy	90
turkish,ozgecanaslan	88
oscars2015,vatanyahutsüleymanşah	78
vatanyahutsüleymanşah,söylemesemolmaz	78
söylemesemolmaz,beyazfutbolsizlerle	78
turkey,özgecanaslan	74
döndümvededimki,dayatma	74

Tablo VI - En çok kullanılan kelime grupları (Etiketli Twit)

Sosyal ağ analizinde en çok kullanılan kavramlardan birisi merkeziliktir. En sık kullanılan merkezilik çeşitleri ise yakınlık merkeziliği (closeness centrality) ve arasındalık merkeziliği (betweennes centrality)'dir. Yakınlık merkeziliği (Eşitlik 1) bir birimin diğer birimlere grafikteki uzaklığının toplamıdır (Otte & Rousseau, 2002).

$$c(i) = \sum_{i} d_{ij} \qquad (1)$$

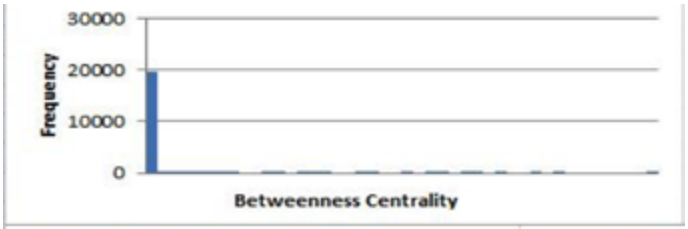
[33]

Arasındalık merkeziliği (Eşitlik 2) ise bir birimin ağda diğer birimler arasında bulunma derecesidir. Bir birimin yüksek derece arasındalığı varsa, köprü görevindedir[33,34].

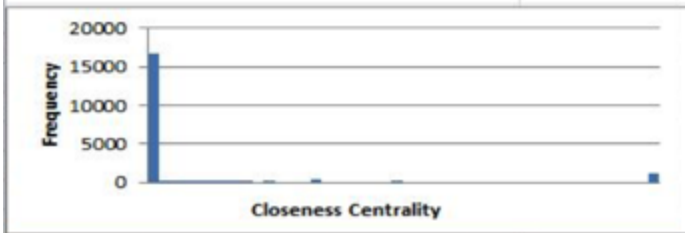
$$b(i) = \sum_{j,k} \frac{g_{jik}}{g_{jk}} \qquad (2)$$

[33]

Analiz edilen veriler üzerinde ortaya çıkan merkezilik sonuçları Tablo 7'de görüldüğü gibi olmuştur.



Minimum Betweenness Centrality	0,000
Maximum Betweenness Centrality	44817261,308
Average Betweenness Centrality	65787,287
Median Betweenness Centrality	0,000



Minimum Closeness Centrality	0,000
Maximum Closeness Centrality	1,000
Average Closeness Centrality	0,080
Median Closeness Centrality	0,000

TABLO VII - Merkezilik sonuçları

Minimum arasındalık merkeziliğinin 0(sıfır) olması, bağımsız twitlerin olduğu anlamına gelmektedir. Ancak ortalama arasındalık merkeziliğine bakıldığında, 65787,287 gibi yüksek bir oran olması, birbiriyle bağlantılı pek çok twit verisinin olduğu anlamına gelmektedir. Ayrıca maximum yakınlık merkeziliğinin ve maximum arasındalık merkeziliği oranlarına bakıldığında, elde edilen twit verileri içerisinde birbiriyle bağlantılı pek çok twitin olduğu anlaşılmaktadır.

V. SONUÇLAR VE ÖNERİLER

Ortaya çıkan tablo ve grafiklerden de anlaşıldığı gibi, sosyal medya verileri pek çok bilginin ortaya çıkarılabileceği bir büyük veri kaynağıdır. Bu çalışma içerisinde Türkiye'de sosyal medya üzerinde çok tartışılan bir olay analizi gerçekleştirilmiştir. Elde edilen veriler içerisinden bilgi çıkarımı ve görselleştirme işlemleri yapılmıştır. Bazı anlamlı bilgilere ulaşılmıştır. Türkiye'de TT listesinde pek çok paylaşım reklam amaçlı yapılmaktadır. Twitler içerisindeki web site adresleri bunu göstermiştir. Gerçek verilere ulaşmak için öncelikle veriler içerisinde temizleme işlemi gerçekleştirilmektedir. Veriler URL'lerden ve etiketlerden temizlendiğinde daha anlamlı bilgiler ortaya çıkmaktadır. Verilerin analizi sırasında aynı twitler içerisinde farklı cinayetlerden birlikte bahsedilmesiyle görülmüştür ki toplum hafızası, benzer olaylar arasında köprü kurabilmektedir. Ayrıca etiketli verilerde de yüksek merkezilik oranları göstermiştir ki bazı kullanıcılar ve twitler, pek çok farklı kullanıcı arasında köprü görevi görmektedir.

Toplumsal olayların önceden tespit edilebilmesi veya gerçekleşen olaylarda geriye dönük araştırma yapılabilmesi için, sosyal medya verileri önemli bir istihbarat ortamı haline gelmiştir. Siber istihbarat günümüzde daha da önem kazanmaya başlamıştır. Ülkeler istihbarat kurumlarında siber istihbarat birimleri oluşturmaya başlamıştır. Gereksinimler doğrultusunda farklı program ve algoritmalar kullanılarak, sosyal medya veri yığınlarından istendik sonuçlara ulaşmak mümkündür. Bu kadar yoğun verinin aktığı ortamlara kayıtsız kalınmamalıdır. İstihbarat açısından çok önemli olan sosyal medya sitelerinden ticari, akademik veya güvenlik amaçlarıyla analiz edilebilir.

Sosyal medya verileri üzerinde analiz işlemleri yaparken karşılaşılan sorunlardan bazıları haber linkleri, TT listesinde görünmek için sadece belirli kelimeleri kullanarak twit yazmak, sadece simgeler içeren twit verileri, kısaltılmış kelimeler içeren twitler, farklı diller kullanılarak yazılan etiketli twitler olmuştur. Bu sorunlar araştırmacıların çözmesi gereken sorunlardan başlıcaları olmuştur. Ayrıca sosyal medya siteleri üzerinde gerçekleştirilen duygu analizi işlemleri, toplum nabzını tutmaya yönelik önemli bir konudur. Türkçe diline ait araştırmaların azlığı, Türkçe kelime kütüphanesinin tam olarak oluşturulmamış olması da bu çalışma sırasında karşılaşılan eksiklikler olmuştur. Araştırmacılara sosyal medya verileri üzerinde siber istihbarat, veri temizleme, Türkçe duygu analizi işlemleri konuları, gelecekte araştırma konuları olarak önerilmektedir.

KAYNAKLAR

[1] Koçak, Z. (2006-5). Six Degrees Of Seperation. Hukuk Gündemi Dergisi, 108.

[2] Tufan, S. (2013, 09 10). Aramızdaki Mesafe: En fazla 6 kişi. 08 15, 2015 tarihinde INNOVA: http://www.innova.com.tr/blog/yazi.asp?ID=117&baslik=Aramizdaki-mesafe_-En-fazla-6-kisi adresinden alındı

[3] Whoriskey, P. (2008, 08 02). Instant-Messagers Really Are About Six Degrees from Kevin Bacon. 08 15, 2015 tarihinde The Washington Post: http://www.washingtonpost.com/wp-dyn/content/article/2008/08/01/AR2008080103718.html

[4] Smith, D. (2008, 08 03). Proof! Just six degrees of separation between us. 08 15, 2015 tarihinde The Guardian: http://www.theguardian.com/technology/2008/aug/03/internet.email

[5] Grossman, L. (2015, 12 15). PERSON OF THE YEAR 2010. 08 15, 2015 tarihinde TIME: http://content.time.com/time/specials/packages/article/0,28804,2036683_2037183,00.html

[6] nedir.com. (tarih yok). Arap Baharı Nedir? 08 5, 2015 tarihinde: http://arapbahari.nedir.com/

[7] Vikipedi. (2015, 06 23). Arap Baharı. 08 10, 2015 tarihinde Vikipedi: https://tr.wikipedia.org/wiki/Arap_Bahar%C4%B1

[8] Kınık, A. M. (2012). Arap Baharı Bağlamı'nda Sosyal Medya-Birey Etkileşimi ve Toplumsal Dönüşüm. 21. Yüzyılda Eğitim ve Toplum Eğitim Bilimleri ve Sosyal Araştırmalar Dergisi, 87-98.

[9] Babacan, M. E., Haşlak, İ., & Hira, İ. (2011). Sosyal Medya ve Arap Baharı. Akademik İncelemeler Dergisi, 63-92.

[10] Banko, M., & Babaoğlan, A. R. (2013). Gezi Parkı Sürecine Dijital Vatandaş'ın Etkisi.

[11] Haber, Mynet (2015, 07 04). Sultanahmet'te Çinli sandıkları Koreli turist grubuna saldırdılar. 08 10, 2015 tarihinde Mynet Haber: http://www.mynet.com/haber/guncel/sultanahmette-cinli-sandiklari-koreli-turist-grubuna-saldirdilar-1909429-1

[12] Cumhuriyet. (2015, 07 09). Bu sefer de Çinli sanılan kadını dövdüler. 08 10, 2015 tarihinde Cumhuriyet: http://www.cumhuriyet.com.tr/video/video/318071/Bu_sefer_de_Cinli_sanilan_kadini_dovduler.html

[13] TDK. Türk Dil Kurumu. 03 29, 2015 tarihinde Türk Dil Kurumu: http://www.tdk.gov.tr/

[14] Avcı, G. (2004). İstihbarat Teknikleri: Aktörleri - Örgütleri ve Açmazları. İstanbul: Timaş Yayınları.

[15] Özdağ, Ü. (2010). İstihbarat Teorisi. Ankara: Kripto Yayınları.

[16] Warner, M. 03 29, 2015 tarihinde Central Intelligence Agency: https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/vol46no3/article02.html

[17] Özçoban, C. (2014). Uluslararası İlişkiler Ana Bilim Dalı Yüksek Lisans Tezi. 21. Yüzyılda Ulusal Güvenliğin Sağlanmasında Siber İstihbaratın Rolü. İstanbul, Türkiye: T.C. Harp Akademileri Stratejik Araştırmalar Enstitüsü.

[18] MİT. 03 29, 2015 tarihinde Milli İstihbarat Teşkilatı: https://www.mit.gov.tr/tarihce/giris.html

[19] MİT2. 03 29, 2015 tarihinde Milli İstihbarat Teşkilatı: http://www.mit.gov.tr/t-cark.html

[20] Sechser, T. S. (2003). Intelligence and Prediction In An Unpredictable World. Eisenhower National Security Series, 4.

[21] Alkan, M. (2012). Siber Güvenlik ve Siber Savaşlar. Ankara: Siber Güvenlik Siber Savaşlar TBMM İnternet Komisyonu.

[22] Bayraktar, G. (2014). Harbin Beşinci Boyutunun Yeni Gereksinimi: Siber İstihbarat. Güvenlik Stratejileri Dergisi, 119-147.

[23] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 1-8.

[24] Zhang, X., Fuehres, H., & Gloor, P. A. (2011). COINs2010: Collaborative Innovation Networks Conference. Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear". Procedia - Social and Behavioral Sciences.

[25] Agnihotri, R., Kothandaraman, P., Kashyap, R., & Singh, R. (2012). Bringing Social into Sales- The Impact of Salespeoples Social Media Use on Service Behaviors and Value Creation. Journal of Personal Selling & Sales Management, 333-345.

[26] Mostafa, M. M. (2013). More than words: Social networks' text mining for consumer brand sentiments. Expert Systems with Applications, 4241–4251.

[27] Tang, X., & Yang, C. C. (2014). Detecting Social Media Hidden Communities Using Dynamic Stochastic Blockmodel with Temporal Dirichlet Process. ACM Transactions on Intelligent Systems and Technology, 36.

[28] Weichselbraun, A., Gindl, S., & Scharl, A. (2014). Enriching semantic knowledge bases for opinion mining in big data applications. Knowledge-Based Systems, 75-85.

[29] Yang, M. C., & Rim, H. C. (2014). Identifying interesting Twitter contents using topical analysis. Expert Systems with Applications, 4330–4336.

[30] Omand, S., Bartlett, J., & Miller, C. (2012). Introducing Social Media Intelligence(SOCMINT). Intelligence and National Security, 801-823.

[31] Papic, M., Noonan, S., & Çeviri(Ece Dündar). (2011).

Sosyal Medya: Bir Protesto Aracı. Türk Kütüphaneciliği, 165-172.

[32] Smith, M. (2015, 02 23). NodeXL: Network Overview, Discovery and Exploration for Excel. 08 10, 2015, CodePlexProject Hosting for OSS: http://nodexl.codeplex.com/

[33] Otte, E., & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. Journal of Information Science, 441-453.

[34] Blogger, G. (2012, 05 24). Tools for Transparency: A How-to Guide for Social Network Analysis with NodeXL. 08 15, 2015 tarihinde Sunlight F: https://sunlightfoundation.com/blog/2012/05/24/tools-for-transparency-a-how-to-guide-for-social-network-analysis-with-nodexl/

Serkan Savaş : 1982 yılında Ankara'da doğdu. Lisans eğitimini Gazi Üniversitesi Teknik Eğitim Fakültesi Bilgisayar Sistemleri Öğretmenliği Bölümü'nde tamamladı. Yüksek Lisans eğitimini Gazi Üniversitesi Bilişim Enstitüsü Elektronik-Bilgisayar Eğitimi Anabilim Dalı'nda tamamladı. Halen Gazi Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği (Tek. Fak.) Anabilim Dalı Doktora öğrencisi olarak eğitimine devam etmekte ve Kızılcahamam Mesleki ve Teknik Anadolu Lisesi'nde Müdür Yardımcısı olarak çalışmaktadır. Veri Madenciliği, Büyük Veri, Siber Güvenlik ve Siber İstihbarat alanlarında çalışmaları vardır.

Nurettin Topaloğlu : Elektronik Bölümü mezunu olup, Gazi Üniversitesi Bilgisayar Mühendisliği Bölümü'nde Öğretim Üyesi olarak çalışmaktadır. Bilişim teknolojileri ve sistemleri ilgi alanına girmektedir.

E-POSTALARDA ADLİ BİLİŞİM VE KARŞI ADLI BİLİŞİM TEKNİKLERİ

R. MARAŞ, E. B. CEYHAN, Ş. SAĞIROĞLU

Ş.S., Gazi Üniversitesi Mühendislik Fakültesi, Eti Mh. Yükseliş Sk. No: 5, Maltepe / Ankara. (e-posta: ss@gazi.edu.tr)
E.B.C., Gazi Üniversitesi Mühendislik Fakültesi, Eti Mh. Yükseliş Sk. No: 5, Maltepe / Ankara. (e-posta: eyupburak@gazi.edu.tr)
R.M., Gazi Üniversitesi Mühendislik Fakültesi, Eti Mh. Yükseliş Sk. No: 5, Maltepe / Ankara. (e-posta: rifatmaras@gazi.edu.tr)

Özet — Gelişen teknoloji ile beraber iletişim araçlarının çeşitleri ve sayısı hızla artmaktadır. Günümüzde e-postalar en sık başvurulan iletişim araçlarından birisidir. Bununla birlikte e-postalar işlenen suçlarda araç olarak kullanılmaktadır. Bu nedenle e-postaların adli olarak incelenmesi ve e-postalara ilişkin her kaydın delil olabilecek şekilde elde edilmesi büyük önem arzetmektedir. İyi bir bilirkişi, analizci veya inceleyici personel e-postalardaki işleyişi iyi bilmelidir. Ayrıca suç işleyenlerin kendilerini gizleme ve delilleri saklama/yok etme yöntemlerine de hâkim olmalıdır.
Bu çalışmada adli bilişim ve karşı adli bilişimin ne olduğu açıklandıktan sonra, e-postaların adli olarak incelenmesi ve karşı adli bilişim teknikleri açıklanmıştır. E-postaların araç olarak kullanıldığı suçlarda, suçluların muhtemel hareket tarzları ortaya konmuş ve adli bilişim uzmanının karşılaştığı olaylarda hangi hususlara dikkat edeceği açıklanmıştır.

Anahtar Kelimeler — Adli bilişim, e-posta, karşı adli bilişim, başlık bilgisi, anonim e-posta.

Abstract — The number and means of communication are rapidly increasing with the deveoping technology. E-mail is one of the means of communication that is mostly used nowadays. However, e-mail is being used as a tool of committing crime. Thus, forensic examination of e mails and getting records of each e-mail to be used as an evidence assume great importance. A qualified expert, analyst or examiner must be aware of how the e-mail system works. Also, they must know the ways of spoliation of evidence that criminals use and their ways of hiding themselves.
In this study, after computer forensics and antiforensics are described, forensic examination of e-mails and antiforensic techniques are clarified. In addition, criminals' probable course of action in the crimes, in which e-mail is used as a way of, and the issues forensic experts should pay attention to in the events they come across are explained.

Keywords — Forensics, e-mail, antiforensics, headers, remailer.

I. GİRİŞ

E-posta kelimesi, İngilizce'deki "electronic mail" ifadesinin Türkçe anlamı olan "elektronik posta" kelimelerinin kısaltmasıdır. E-posta hizmeti de, gerçek hayattaki posta hizmeti model alınarak gerçekleştirilen hizmettir. Gerçek dünyada bir mektup yazılmasına müteakip, zarf içerisine konularak ve zarfın üzerine gönderici, alıcı isimleri ve adresleri gibi bilgiler yazılarak postaneye verilir. Postane görevlileri de zarf üzerindeki bilgilere (başlık bilgileri) bakarak mektubu ilgili adrese gönderirler. E-postalar da bu yapıya benzer şekilde çalışır. Gerçek hayattan farkı ise, aracı olarak insanlar değil bilişim sistemlerinin kullanılmasıdır [1].

İnternetin gelişmesiyle beraber, günümüz iletişim sistemlerinde telefonlardan sonra e-postaların kullanıldığı görülmektedir. Bunun en önemli sebepleri arasında, e-posta hizmetlerinin çoğunlukla ücretsiz ve kullanımının basit olması söylenebilir. GSM servis sağlayıcılarının internet kullanımını artırma yönündeki çalışmalarıyla birlikte e-postaların, kısa mesajın (SMS) yerini alacağı öngörülmektedir [2].

Bununla birlikte e-postalar yalnızca haberleşme için değil, çoğu zaman internet ortamında sanal kimlik olarak da kullanılmaktadır. Ticari internet sitelerinden hizmet almak maksadıyla üye olma ve sanal ödeme yapma gibi işlemler e-postalar kullanılarak yapılmaktadır. Bununla birlikte günümüzde e-posta hizmetini aktif olarak kullanan kişilerin hesap bilgileri, bilgisayar korsanları (hackerlar) ya da kötü niyetli kişilerce ele geçirildiğinde çok üzücü durumlarla karşılaşılmaktadır. Bu da e-postalarda güvenliğin ne kadar önemli olduğunu göstermektedir [2].

E-postaların işlenen suçlarda araç olarak kullanılması gerçeği de önemini artırmaktadır. Bu nedenle işlenen suçların aydınlatılmasında kullanılması ve dijital delil olarak kabul edilmesi göz önüne alındığında, e-postanın gönderilişinden alıcısına ulaşana kadarki yaşam döngüsünün çok iyi incelenmesi gerekmektedir.

Bu çalışmada, adli bilişim ve karşı adli bilişim konuları detaylıca açıklandıktan sonra e-postaların adli bilişim incelemelerindeki yeri ve bu incelemelerde dikkat edilmesi gereken konular tartışılmış ve adli incelemelere karşı koyma teknikleri ele alınmıştır.

II. ADLİ BİLİŞİM VE KARŞI ADLİ BİLİŞİM

A. Adli Bilişim

Adli Bilişim, işlenen bir suçun aydınlatılması ve suçluların tespiti için suç esnasında kullanılan bilişim nesnelerinden ihtiyaç duyulan sayısal delillerin elde edilmesini sağlamaktadır [3].

Ayrıca adli bilişim; elektromanyetik ya da elektro optik ortamda saklanan veya bu ortamlarca iletilen; ses, görüntü, metin veya bunların birleşiminden oluşan her türlü bilişim nesnesinin, dijital delil niteliği taşıyacak şekilde elde edilmesi, muhafaza edilmesi, incelenmesi ve sonucun adli makamlara sunulması çalışmalarıdır [4].

B. Karşı Adli Bilişim

Dijital deliller üzerinde adli bilişim yöntemlerinin başarılı olamaması için geliştirilen yöntemlerdir. Purdue Üniversitesinden Dr. Marc Rogers'a göre karşı adli bilişim; "olay mahallinde bulunan delillerin varlığını, miktarını ve/veya kalitesini olumsuz yönde etkilemek, incelenmesini ve analizini zorlaştırmak hatta imkânsız hale getirmek için kullanılan her türlü yöntem/aksiyondur" [5].

Diğer bir deyişle karşı adli bilişim; adli delilin ortadan kaldırılması, silinmesi, karartılması ve anlamsızlaştırılması, delil ekleme, adli bilişim alanında personelin veya uygulamalarının etkisiz hale getirilmesi ve bütün bu faaliyetler neticesinde adli makamların etkilenmesinin sağlanmasıdır [3].
Bilişim dünyasındaki gelişmeler ile birlikte gelecekte,

ülkelerarası savaşların ve terör olaylarının çoğunun sanal dünyada gerçekleşebileceği değerlendirilmektedir. Buna paralel olarak iletişim, ulaşım, enerji ve e-vatandaşlık gibi birçok hizmetin tamamen bilgisayara bağımlı olarak çalışacak olması nedeniyle, sanal ortamlarda güvenlik kavramının önemi artmıştır. Bununla birlikte adli bilişim alanında standartların henüz belirlenmemesi ve yapılan çalışmaların yeterli seviyeye ulaşmamış olması adli bilişimin akademik bir disiplin haline gelme ihtiyacını ortaya koymaktadır.

Adli bilişim inceleme teknikleri ve karşı teknikler, birbirinin çalışma mantığını anlama temeline dayanmaktadır. Karşı adli bilişim teknikleri; verilerin silinmesi, veri gizleme, karıştırma, yanlış yönlendirme, aldatma, engelleme ve şifreleme olarak sıralanabilir. Metasploit Antiforensics, Linux Kernel Module, ADMutate, Evidence Eliminator, SecureClean, Steganos, StealthDisk, Rootkit FU gibi araçlar, adli bilişim incelemelerini tamamen etkisiz hale getirebilmekte veya delil bulma sürecini oldukça uzatabilmektedir [6].

III. E-POSTALARIN YAPISI VE GÖNDERİM AŞAMALARI

Bu bölümde e-postaların yapısı, e-postayı meydana getiren bölümler, başlık bilgileri ve gönderim aşamaları incelenmiştir.

A. E-postaların Yapısı

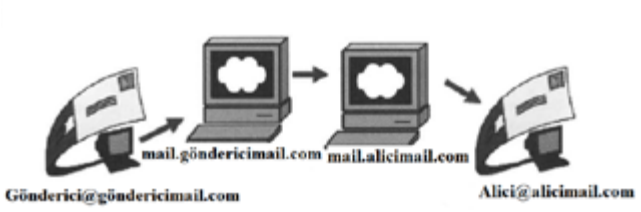
E-postalar uygulama katmanında kime, konu ve mesaj alanı (body) bölümlerinden oluşur. Ayrıca hizmet alınan e-posta servis sağlayıcı tarafından opsiyonel olarak eklenen CC (Carbon Copy) ve BCC (Blind Carbon Copy) bölümleri de bulunabilir [7]. Genel olarak e-postalar şu bölümlerden oluşur;

Kime: Mesajın gönderileceği e-posta adresinin yer aldığı bölümdür.
Konu: Mesajın konusu veya başlığının ifade edildiği bölümdür.
Mesaj alanı: Mesajın tam içeriğinin yer aldığı bölümdür.
CC: Mesajın gönderileceği bir başka e-posta adresinin yazılacağı bölümdür.
BCC: Mesajın bir kopyasının gönderileceği ancak diğer alıcılardan gizlenen e-posta adreslerinin yazıldığı bölümdür.

B. E-posta Gönderim Aşamaları

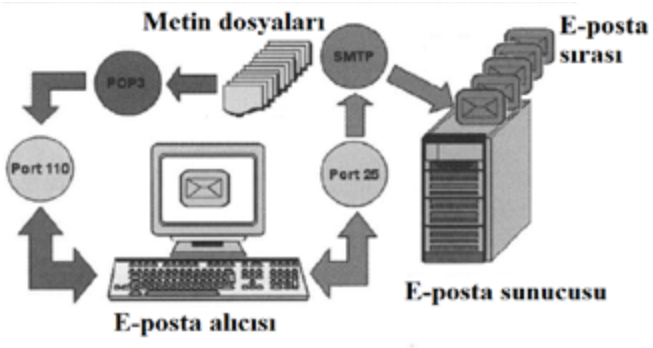
E-posta gönderme ve alma işlemi, posta alıcısı ve posta sunucusu olmak üzere iki türlü sistemden oluşmaktadır. E-posta başlık bilgilerini anlamak için öncelikle bir e-postanın A noktasından B noktasına doğrudan gitmeyeceği bilinmelidir. Her e-posta, yaşam döngüsü içerisinde en az dört bilgisayardan geçer. Şekil 1'de görüldüğü gibi e-posta transferinde alıcı ve gönderici arasında, göndericinin ve alıcının mail sunucuları olmak üzere en az iki bilgisayar daha vardır [1].

E-posta alışverişine bir örnek şu şekildedir:
Gönderici : gönderici@göndericimail.com
Alıcı : alıcı@alicimail.com
Mail sunucumuz : mail.göndericimail.com
Alıcı sunucusu : mail.alicimail.com



Şekil 1. Örnek e-posta alışverişi [1].

Şekil 2'de örnek bir e-posta alışverişi gösterilmektedir. Gönderici ve alıcı arasında e-posta gönderilmesinde kullanılan SMTP (Simple Mail Transfer Protocol) sunucusu ile e-posta alımında kullanılan POP3 (Post Office Protocol) sunucusu vardır. POP3, e-postanın posta sunucusundan alıcının bilgisayarına alınmasını sağlar. SMTP ise, e-postanın göndericinin bilgisayarından posta sunucusuna gönderilmesini sağlar [7]. Aşağıdaki şekilde standart bir e-posta altyapısı gösterilmiştir. E-postanın alım işlemi, POP3 ile 110 numaralı porttan, gönderim işlemi ise SMTP ile 25 numaralı porttan gerçekleşir. Kullanılan bu port numaraları değiştirilebilir.



Şekil 2. Standart bir e-posta altyapısı [1].

IV. E-POSTALARDA ADLİ BİLİŞİM

Genel olarak e-postaların adli olarak incelemesi şu konuları kapsamaktadır;

- E-posların başlık bilgilerinin incelenmesi,
- E-posta servis sunucularının incelenmesi,
- Bilgisayarda kurulu e-posta uygulamalarının (MS Outlook, Thunderbird, vb.) incelenmesi,
- Bilgisayarın bağlı olduğu ağların (network forensics) incelenmesi,
- İşletim sistemlerinde e-postalara yönelik tutulan kayıtların (registry) ile internet tarayıcıların (browser) incelenmesi,
- E-postalarda metin madenciliği [5, 8].

A. E-posta Başlık Bilgilerinin İncelenmesi

İletişimi gerçekleşen her e-posta, başlık bilgisine sahiptir ve bu başlık bilgilerinin incelenmesi adli ve idari soruşturmalarda büyük önem taşımaktadır. E-postalar tıpkı kargo şirketlerinin, kargonun taşınması ve takip bilgilerine benzer şekilde bir gönderim hikâyesini taşır. Bu hikâye başlık bilgilerinde barınır. Başlık bilgileri;

- Göndericinin e-posta adresi,
- Göndericinin IP (Internet Protocol) adresi,
- Göndericinin e-posta sunucusu,

- Alıcı veya alıcılara ait e-posta adresleri ile IP adresi,
- E-postanın gönderildiği tarih ve saat ile zaman dilimi,
- Gönderici ve alıcının hizmet aldığı e-posta servis sağlayıcının sunucularına ait bilgileri barındırır [2, 4].

E-postalara ait başlık bilgileri şu şekilde açıklanabilir [1,2,5,6,7,19];

From: E-postanın kimden geldiğini gösteren başlık alanıdır. İçeriği çok kolay değiştirilebileceği için en az güvenilir başlık alanıdır.

Reply to: E-postaya bir cevap yazılırsa, cevabın hangi adrese gönderileceğini bildirir.

Return-path: Reply-to başlığına benzemektedir. Eğer e-posta gönderiminde bir hata meydana gelirse, hata mesajının hangi adrese gideceğini belirtir.

Received: Bu alanda yer alan bilgi e-posta iletişimi ile ilgili verdiği detaylı ve gerçekçi bilgiden dolayı oldukça önemlidir. Postanın göndericiden alıcıya ulaşana kadarki tüm bilgisayar/sunuculara ait bilgileri barındırır. Received alanı da diğer başlık alanları gibi değiştirilebilir fakat son Received bilgisi mutlak surette gönderici sunucu tarafından ekleneceğinden gerçek bilgi verecektir.

Timestamp: E-postanın alıcıya ait e-posta sunucusuna ulaştığı zamandır. İlk ve son timestamp bilgilerine bakılarak e-posta sunucularının performanslarına dair bilgiler edinilebilir.

For recipient: Alıcı e-posta adresidir. E-postanın kime gönderildiği bilgisini verir.

Date: E-postanın ilk kaynakta oluşturulma zamanını gösterir. User-agent: E-posta göndericisinin hangi yazılımı kullandığını gösterir.

X-Başlıkları: İstemci ve sunucu dışında özel e-posta yazılımların eklediği başlıklardır. Gerçek başlık değerleri ile karışmaması için X ile başlar. Örneğin bazı webmail uygulamaları gönderdikleri e-postalara X-Originating-IP:[1.2.3.4] şeklinde başlık alanı ekler ve bu başlık bilgisi istemcinin IP adresini gösterir.

E-posta başlık bilgileri çevrim içi (online) siteler ve özel olarak yazılmış bazı yazılım araçları ile de analiz edilebilmektedir. Bununla birlikte içeriğinde kişisel veya gizli olabilecek veriler olduğunda, e-postaları çevrim içi sitelerde analiz etmek uygun değildir. Adli veya idari soruşturmanın da gizliliğine halel getirebilecek işlemlerden kaçınılmalıdır. Google Apps-MessageHeader, MxToolBox, IptackerOnline ve Gaijin e-posta analiz sitelerine verilebilecek örneklerdendir [9-12].

Bu araçlardan en sık kullanılanlardan birisi Google firmasına ait olan Google Apps-MessageHeader'dır. Bu sayfada, bir e-postaya ait başlık bilgisinin analizi ve Gmail, Hotmail, Yahoo, AOL, Excite gibi e-posta servis sağlayıcılar ile Outlook, Mozilla, Apple Mail ve Opera gibi uygulamalardan e-posta başlık bilgilerinin nasıl elde edileceğine dair detaylı bilgiler verilmiştir [13].

Nirsoft firmasına ait IPNetInfo isimli yazılım da e-posta başlık bilgilerinin analizinde oldukça başarılıdır. Bu yazılımı, rekabet ettiği diğer yazılımlardan ayıran en belirgin özelliği ise, e-posta başlığında yer alan IP adresleri hakkında bilgi taraması yaparak, bu IP adreslerinin hangi kurum ya da kişiye ait olduğu, IP adresinin konumu, ait olduğu firma veya kişinin e-postası, telefon numarası, faks numarası ve adresi gibi çok önemli bilgileri toplaması ve sunmasıdır [14].

B. E-Posta Servis Sunucuları

İnternet dünyasında e-posta servis sağlayıcıları, sağladıkları hizmetleri ücretli veya ücretsiz olarak sunabilirler. Devlet kurumları ve özel şirketler genelde ücretsiz e-posta hizmeti sunmaktadırlar. Kurumlar ve özel şirketler, e-posta sunucularını doğrudan sahip oldukları sunuculardan veya Microsoft Outlook ve Exchange, IBM Lotus Notes, Novell GroupWise gibi e-posta sunucularından da bu hizmeti kiralamak suretiyle sağlayabilirler. Bu sunucular da düzenli olarak yedekleme yaparlar. Bu sunucuların adli incelemesi yapılarak, sunucular üzerinden transfer edilen tüm e-posta trafikleri içerikleriyle beraber elde edilebilir [15, 16].

C. Bilgisayarda Kurulu E-posta Uygulamaları

Bilgisayarlarda e-postaların otomatik olarak takibinin yapılabildiği uygulamalar vardır. Bunların en sık kullanılanları MS Outlook, Mozilla Thunderbird ve Opera'dır. Bu uygulamalar da bir bilgisayarda yapılan e-posta haberleşmeleri hakkında bilgiler saklar ve e-postaların adli olarak incelenmesinde büyük önem taşırlar.

MS Outlook uygulamasının veri dosyaları olan OST (Offline Storage Table) ve PST (Personal Storage Table) uzantılı dosyalar, bilgisayarda kurulu MS Outlook uygulaması ile yapılan e-posta trafiğinin kaydedildiği dosyalardır. Benzer şekilde Outlook Express DBX, MBX ve IDX uzantılı dosyalarda, Grupwise uygulaması MLM ve DB uzantılı dosyalarda, Lotus NSF ve ID uzantılı dosyalarda, Apple Mac İşletim sistemi ve Thunderbird uygulaması MBOX uzantılı dosyalarda e-posta ile ilgili kayıtlar tutar [7, 15, 16].

MS Outlook Express uygulaması, kullanıcıların isim, telefon numarası, işyeri adı ve adresi gibi bilgilerin yazıldığı "vCards" oluşturmasını istemektedir. vCards incelenerek, e-posta kullanıcısı hakkında ekstra bilgilere de ulaşılabilmektedir [17].

E-postaların adli incelemesini yaparken, bilgisayarda tutulan bu dosyaları anlamlandırarak inceleme yapmaya imkân sunan bazı yazılımlar vardır. Bu yazılımlardan biri de Paraben E-mail Examiner'dir [16].

D. Bilgisayarın Bağlı Olduğu Ağdaki Kayıtlar

MS Exchange uygulamasının EDB (Exchange DataBase) uzantılı dosyası ve Lotus Notes uygulamasının ise NSF (NoteS File) uzantılı dosyası incelenerek e-posta trafiğine ulaşılır. Ayrıca bilgisayarın ağdaki canlı trafiği incelenerek e-posta servisine bağlanma zamanı, e-posta gönderim ve alım aktiviteleri gibi e-posta kayıtları incelenebilir [18].

Bu dosyalar genel adli bilişim yazılımları (X-Ways, Encase, FTK) ile incelenebileceği gibi sadece e-posta sunucuları ve

ağdaki e-posta trafiğini incelemede kullanılan Paraben's Network E-mail Examiner yazılımı ile de incelenebilir [19].

E. Bilgisayarda Kurulu İnternet Tarayıcıları

Bilgisayarda kurulu internet tarayıcılar da e-postaların adli olarak incelenmesinde büyük önem taşırlar. Internet Explorer, Google Chrome ve Mozilla Firefox gibi internet tarayıcılarının internet geçmişi ile ilgili tuttuğu kayıt bilgileri, e-postaların gönderilip alındığı zamanlarda e-posta servis sağlayıcılara (Gmail, Outlook, Yahoo Mail, vb.) bir bağlantı kuruyorsa buradan da adli olayların aydınlatılmasında ipuçları elde edilebilir [16, 20].

Ayrıca işletim sistemlerinin tuttuğu kayıtlar (registry) içerisinde de e-postaların gönderim ve alım aşamalarına ait bilgiler bulunur. Örneğin Windows işletim sistemlerinde index.dat isimli dosya internet geçmişi ile ilgili bilgiler tutar. Linux işletim sistemlerinde ise mail isimli e-posta loglarının tutulduğu dosya bulunmaktadır. Bu dosyalar incelendiğinde suçla ilişkili e-postalara ait bilgiler elde edilebilir [17, 20].

F. E-postalarda Metin Madenciliği

E-postaların gönderici tespiti ile ilgili çalışmalar yapılırken, postanın içeriği de mutlaka incelenmelidir. Her e-posta göndericisi, göndericinin seçtiği kelimeler, kısa ya da uzun yazması, kelime sayısı, üslup, şive ve gramer hataları, e-posta metninin şekli, paragraf ve cümle yapısı, başlangıç ve bitişte yer alan ifadeler, kullanılan karakter ve işaretlerle iz bırakır. E-posta metinlerinde metin madenciliği yapılarak göndericiye dair işaretler bulunabilir. Yazarı bilinmeyen ve tahmin edilemeyen e-postaların içeriği ile bilinen kullanıcıların e-posta içerikleri metin madenciliği teknikleri ile kıyaslanarak e-posta yazarı (göndericisi) tespit edilebilir [21, 22].

V. E-POSTALARDA KARŞI ADLİ BİLİŞİM TEKNİKLERİ

E-postaların adli olarak incelenmesinde, incelemeyi sonuçsuz bırakmayı, uzatmayı ve delil olabilecek verileri değiştirerek bozulmasına neden olabilecek karşı adli bilişim teknikleri;

- E-posta başlık bilgilerinin silinmesi ve değiştirilmesi,
- Sahte e-posta gönderimi,
- Geçici e-posta hesabı ile gönderim,
- E-postalara ilişkin tutulan kayıtların silinmesi,
- Vekil sunucu (proxy) ya da VPN (sanal özel ağlar) kullanılarak e-posta gönderimi şeklinde olabilmektedir [5, 8].

A. E-posta Başlık Bilgilerinin Silinmesi veya Değiştirilmesi

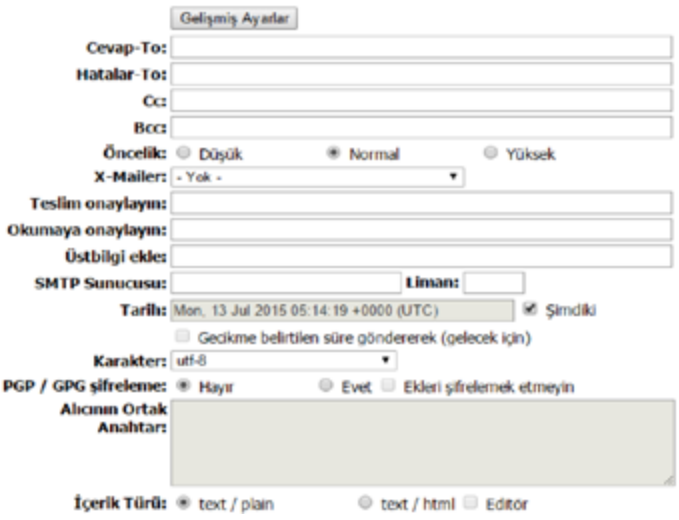
E-postalarda başlık bilgilerinde son "Received" başlığı ile son IP adresi ve zaman damgalarını içeren e-posta sunucusu bilgileri değiştirilemez. Onların dışında olan konu, tarih, mesaj-ID, gönderici-alıcı (from, to, CC, BCC), mesaj içeriği, x-mailer, x-message info ve başlangıçtaki "received" başlıkları değiştirilebilir [1].

Günümüzde birçok e-posta servis sağlayıcı başlık bilgilerinden Göndericiye ait bilgileri silmektedir. Gmail, Outlook, Yandex gibi en çok kullanılan posta sağlayıcılarından gelen e-postaların başlık bilgilerinde göndericiden alıcıya kadar tüm başlık bilgileri yer almaz. Sadece bu servis sağlayıcılardan alıcıya ulaşana kadarki başlık bilgileri yer alır. Benzer şekilde Dark Mail olarak adlandırılan serviste de e-posta başlık bilgilerinde yer alan göndericiye ait ve diğer posta sunucularına ait bilgiler, başlık bilgilerinden çıkarılmaktadır [23].

E-posta adreslerinin başlık bilgileri bazı uygulamalar kullanılarak değiştirilebilmektedir. Bu değişiklikler de inceleme yapan kişiyi yanıltabilir. Buna örnek olabilecek çevrim içi araçlardan birisi "Emkei's Mailer" sitesidir [24]. Şekil 3'te ekran çıktısı görülen bu araç sayesinde; sahte isim ve sahte e-posta adresi ile posta gönderimi yapılmaktadır. From satırına yazılan e-posta adresi hiç kullanılmayan adres olabileceği gibi var olan bir e-posta adresi de olabilir.

Şekil 3. E-posta başlık bilgilerinin manipüle edilmesinde kullanılan bir araç [24].

Bu aracın gelişmiş seçenekler bölümünde, mesaja cevap yazılırsa ya da hata mesajı alınırsa hangi adrese gönderileceği, birden fazla kişiye gönderimde kullanılan CC ve BCC alanları, e-postanın öncelik durumu, e-posta gönderiminden sonra ulaşma bilgisi, okundu ise okunma bilgisi, e-posta başlığına herhangi bir ekleme yapılacaksa bu eklemenin girileceği alan, sunucu ismi, tarih değişikliği yapılabilecek alanlar gibi e-postanın başlık bilgilerinin değiştirilmesinde kullanılabilecek birçok alan mevcuttur.



Şekil 4. Gelişmiş seçenekler [24].

Geçmiş dönemde e-posta başlık bilgilerinde sık kullanılan yöntemlerinden biri de tarih ve saat gibi zaman damgalarının değiştirilmesi üzerine olmuştur. 2009 yılının sonuna kadar bu konuda Yahoo Mail, Gmail ve Hotmail gibi servis sağlayıcılar da önlem alamamış ve kişiler ile kurumlar üzerinde birçok mağduriyetler yaşanmıştır [25].

Başlık bilgilerinin değiştirildiği bilgisini tespit işleminde göndericinin e-posta sunucusuna bakılarak uyumlu olup olmadığı incelenmelidir. Örneğin abc@gmail.com e-posta adresinden posta geldiğini varsayalım. Bu e-posta adresinin sunucusu GMail'dir. Gelen mesajın başlık bilgisi de GMail posta sunucusuna uygun olmalıdır. "X-originating IP", "X-originating E-mail" gibi başlık bilgilerinin bölümleri Gmail'in posta formatına uygun olmalıdır. Uygun değilse sahte bir e-posta olduğu çıkarımına varılabilir [2].
E-posta sunucusunun DNS ismi ve makine ismi farklı olması durumunda da başlık bilgilerinin değiştirilmiş olabileceğinden şüphelenilmelidir. From satırında e-postayı gönderen sunucunun smtp2.abc.com.tr olduğu görülür ve aynı adrese DNS sorgulaması yapıldığında farklı bir isim görülürse bu başlığın değiştirilmiş olabileceği sonucuna varılır. Bazen de e-postayı gönderen makinenin DNS ismi ile kendi üzerinde tanımlanmış ismi farklı olur ve Received kısmında iki farklı isim görülür. Örneğin makinenin ismi smtp2.xyz.com.tr olarak tanımlanmasına rağmen DNS kaydı smtp2.abc.com.tr şeklindedir [2].

B. Sahte E-posta Gönderimi

Sahte e-postalarda kullanılan çevrim içi sayfalara ait bilgiler açık kaynaklarda sıkça yer almaktadır. Örneğin www.spamhaus.org isimli sitede SBL (Spam Block List) bölümü yer almakta olup sahte, reklam veya kötü niyetli gönderildiği düşünülen e-postalardaki IP adresi ve site isimleri bu güncel veri tabanına sahip sayfadan sorgulanabilir [26].

Sahte e-posta gönderim araçlarından biri de www.wetransfer.com isimli sitedir. Bu sitede de e-posta doğrulaması yapılmadan başkası adına e-posta gönderilebilir [27].

Sahte e-posta gönderimi bir yazılım vasıtasıyla da yapılabilir. MS Windows platformunda çalışan "Private Idaho" isimli yazılım ile anonim şekilde e-posta gönderimi yapılabilir [28].

C. Geçici E-posta Hesabı

E-postayı suç işlemek amacıyla kullanan kişilerin sık başvurduğu yollar; geçici e-posta hesabı, tek kullanımlık e-posta hesabı veya günübirlik e-posta hesabı almak şeklinde olabilir. Bu yöntemlerin tercih edilme nedenleri arasında; ücretsiz ve reklamsız olmalarının yanı sıra birçoğunun kullanıcı kimliklerini gizli tutması, IP adresi kaydetmemeleri ve güvenli (şifreli) haberleşme yolunu kullanmalarıdır [23].
Bu şekilde hizmet veren çevrim içi araçlara, Hushmail, Mytrashmail, 10minutemail ve 5ymail gibi sayfalar örnek verilebilir [29-32].

www.hushmail.com sitesini incelediğimizde bu site, OpenPGP standartlarını kullanarak PGP (Pretty Good Privacy) şifreli e-posta imkânı sunmaktadır [28, 29].

Bir diğer örnekte "Remailer" olarak adlandırılan, tek kullanımlık, bir günlük ya da belirli bir süre kıstası ile kendini imha eden e-postalardır. Buna verilebilecek bir örnek de www.mytrashmail.com sitesidir [30].

D. E-postalara İlişkin Tutulan Kayıtların Silinmesi

E-postalarda karşı adli bilişim teknikleri kapsamında en önemli ve en basit delil imha yöntemlerinden birisi, e-postalara ilişkin tutulan kayıtların silinmesidir. Bu kayıtların silinmesi; e-posta uygulamalarındaki (MS Outlook, Thunderbird vb.) kayıtların silinmesi, işletim sistemlerinde e-postalara yönelik tutulan kayıtların (internet geçmişi, index.dat ve Linux-mail dosyaları) silinmesi, bilgisayarın bağlı olduğu ağlardaki kayıtların (internet kullanıcı logları, MS Exchange EDB, Lotus NSF dosyaları) silinmesi şeklinde olabilir.

E-posta servis sağlayıcılarda tutulan sunucu kayıtlarında, hizmet veren firmalarca değişiklik veya silme yapılabilir. Bu sunucularda e-posta başlık bilgilerine girecek sunucu isimleri, tarih ve saat gibi zaman damgalarında manipülasyon yapılabilir [23].

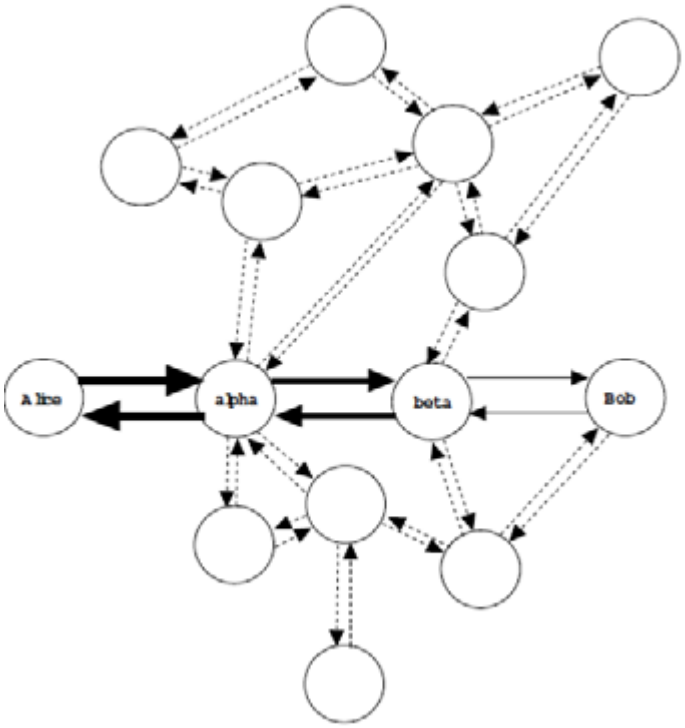
Yukarıda bahsedilen bu dosyalar güvenli silme işlemi (wipe) uygulanarak geri getirilemeyecek şekilde silinirse e-posta inceleyicisinin delil tespiti büyük oranda engellenecektir.

E. Vekil Sunucu veya VPN Kullanımı

E-posta gönderim işleminin bir vekil sunucu üzerinden yapılması durumunda gönderen kişinin tespiti zor olmaktadır. Şekil 5'te gösterildiği gibi Alice'ten Bob'a gönderilen bir e-posta, normal e-posta trafiğine ek olarak Alpha ve Beta isimli iki vekil sunucudan geçmiştir. Vekil sunucu veya VPN hizmeti veren firmaların birçoğu IP adresi değiştirme ve internet trafiğini şifrelemenin yanı sıra kullanıcı loglarını da tutmamaktadır. Bu durumda e-posta göndericisinin kimliğinin tespitini imkânsız kılmakta ya da çok uzatmaktadır [33].

Şekil 5. Alice ve Bob arasında iki vekil sunucu kullanılarak e-posta haberleşmesi [33].

En çok kullanılan ve açık kaynak kodlu VPN'e benzer hizmetleri barındıran TOR Project isimli hizmette, kullanıcılar en az dört farklı sunucudan geçerek internet ortamında dolaşır. Şekil 6'da TOR Project hizmetinde yer alan ve standart e-posta alışverişinin dışında Alpha ve Beta isimli vekil sunucular kullanılmış olup, Alpha ve Beta sunucuları arasında da çok sayıda başka vekil sunucular kullanılabileceği gösterilmiştir. Bu hizmet kullanılarak gönderilen bir e-postanın göndericisine ulaşmak neredeyse imkânsızdır. Teknik olarak mümkün olmasına rağmen uluslararası düzeyde ülkelerin anlaşma yapması gerektiğinden e-posta göndericisine ulaşmak çok zor ve uzun bir süreç gerektirmektedir [23, 33].



Şekil 6. Tor network üzerinde Bob ve Alice'in e-posta haberleşmesi [33].

VI. SONUÇ

Literatürde e-postaların adli incelenmesine yönelik akademik çalışmalar yer alsa da adli incelemeye karşı koyma konusunda yeterli akademik çalışma yer almamaktadır. E-postaların bilişim suçlarında doğrudan kullanılması veya diğer suçlarda da araç olarak kullanılması gerçeğinden yola çıkarak öncelikle e-postaların adli incelenmesi konusuna değinilmiş ve bu alanda kullanılan dünyada kabul görmüş yazılımlardan bahsedilmiştir. Bu yazılımların hangi amaçla kullanıldığı izah edilmiş, okuyucuya rehber olabilecek bir sunum yapılmıştır. Adli incelemenin tüm detaylarına inilmemiş, yol gösterilmiştir.

Çalışmada ayrıca e-postaların adli incelemesinde çalışan kişilerin nelere dikkat etmesi gerektiği vurgulanmış ve bu konuda bireysel ve kurumsal farkındalığı artırma amaçlanmıştır. E-postaların yapısı açıklanmış ve gönderim aşamaları izah edilmiştir. Bununla birlikte e-posta göndericisinin tespitinde en önemli aşamalardan biri olan başlık bilgileri detaylandırılmıştır. E-postaların başlık bilgilerinin ne olduğu, ne gibi bilgiler içerdiği, hangi alanların nasıl değiştirilebileceği ve başkası adına e-posta göndermenin kolaylığı, inceleme yapan kişileri yanıltıcı durumlar örneklerle açıklanmıştır.

E-postaların otomatik takibini yapan uygulamalar e-postalara ilişkin tutulan kayıtlar ve bu kayıtların hangi araçlarla incelenebileceği detaylandırılmıştır. Kurumsal veya bireysel olarak e-posta ile ilişkili konuları inceleme konusunda hangi detaylara dikkat edilmesi gerektiği vurgulanmıştır. Başlık bilgilerinin değiştirilmesi, e-postalara ilişkin tutulan kayıtların silinmesi, sahte e-posta gönderimi, şifreli ve geçici hesapla yapılan e-posta gönderimleri ve vekil sunucu veya VPN kullanımı gibi karşı adli bilişim teknikleri ortaya konmuştur. Profesyonel olarak bir e-posta incelenmek isteniyorsa, tek bir yöntem değil, birden fazla yöntem uygulanmalıdır. İnceleme yapılmasını kolaylaştıran uygulamaların mutlaka alternatifleri de olmalıdır. Tüm dünyada ve ülkemizde adli inceleme konusu yeni ve hızla gelişen bir alan olduğu için araştırma iyi yapılmalıdır.

Adli incelemeyi zorlaştıran, engellemeye çalışan ya da imkânsız hale getirmeyi amaçlayan karşı adli bilişim tekniklerinin bilinmesi de çok önemlidir. Bir araştırmacı veya bilirkişi, e-posta kullanılarak işlenen suçlarda, suçu işleyen hedef kişinin kendini gizleme yönünde attığı adımların neler olabileceğini iyi bilmelidir. Bu nedenle karşı adli bilişim teknikleri de çok önemlidir. İnceleyicinin herhangi bir hususu unutmasını önlemek, sistemli çalışmasını ve adım adım inceleme yapmasını sağlamak amacıyla hazırlanan kontrol listesi Tablo 1'de verilmiştir.

TABLO 1. E-POSTALARIN ADLİ OLARAK İNCELENMESİNDE KONTROL LİSTESİ

1	Gelen e-postanın hangi e-posta adresinden geldiği tespit edildi mi?
2	Gönderici e-posta adresinin doğrulaması ve halen kullanımda olup olmadığı tespit edildi mi?
3	Gönderici e-posta adresinin servis sağlayıcısından kime ait olduğu ve hangi bilgilerle kayıt olunduğu bilgileri talep edildi mi?
4	Gelen e-postanın başlık bilgileri incelendi mi?
5	Başlık bilgilerinden gönderici IP adresi tespit edildi mi?
6	Gönderici IP adresi tespit edildi ise, tarih ve saat bilgisi ile kime ait olduğuna ilişkin internet servis sağlayıcısı ile irtibata geçildi mi?
7	Gönderici IP adresinin vekil sunucu veya VPN kullanıldığına dair bir tespit mevcut mu? Mevcut ise vekil sunucu veya VPN sisteminin yöneticileri ile şüpheli IP adresinden kimlik tespitine yönelik irtibata geçildi mi?
8	Başlık bilgilerinde yer alan Received alanlarındaki IP adresi, sunucu ismi ve tarih-saat bilgilerinde herhangi bir uyumsuzluk var mı?
9	Başlık bilgilerinde yer alan DNS ismi ile DNS sorgulamasındaki isim aynı mı?
10	Başlık bilgileri gönderici e-posta sunucusunun formatına uygun mu?
11	Başlık bilgilerinin manipüle edildiğine dair bir şüphe var mı?
12	Tespit edilen IP adresini kullanan bilgisayarda suça ilişkin e-postayla ilgili kayıtlar (registry, browser, internet kayıtları) incelendi mi?
13	Suçla ilişkili bilgisayarda herhangi bir e-posta uygulaması mevcut mu? Mevcut ise incelendi mi?
14	Suçla ilişkili bilgisayarın bağlı olduğu ağdaki kayıtları incelendi mi?
15	Gönderici e-posta sunucusundaki kayıtlar incelendi mi?
16	Alıcı e-posta sunucusundaki kayıtlar incelendi mi?
17	E-posta göndericisinin kimliği belli değilse, kimliği bilinen e-postalar ile kıyaslanarak metin madenciliği uygulandı mı?

Tablo 1'deki kontrol listesinde yer alan soruların tamamı e-postaların adli incelenmesinde kullanılabileceği gibi, birkaç tanesi de kullanılabilir. Her olay, kendine özgü çalışma gerektirir ve farklı soruları içerebilir.

Gelişen teknoloji ile beraber karşı adli bilişim teknikleri de hızla artmakta ve suç işleyen kişilerin tespiti zor olmaktadır. Ancak her zorluğu aşmak için yeni yöntemler araştırılmakta ve ortaya çıkarılmaktadır. Unutulmamalıdır ki, bilişim ortamında yapılan her işlemin kaydedildiği kaçınılmaz bir gerçektir.

KAYNAKÇA

[1] L. James, "E-Mail: The Weapon of Mass Delivery", Syngress Force Emerging Threat Analysis, Canada: Springer, Bölüm 10, 289-334, 2006.

[2] H. Önal, "E-posta Başlıklarından Bilgi Toplama", Bilgi Güvenliği Akademisi, 2009.

[3] S. Sağıroğlu, "Karşı Adli Bilişim ve Siber Güvenlik", 1. Uluslararası Adli Bilişim Sempozyumu, Ankara, 2014.

[4] H. Önal, "Bilişim Sistemlerinde Adli Bilişim Analizi ve Bilgisayar Olayları İnceleme", Bilgi Güvenliği Akademisi, 2013.

[5] H. Öztürkçi, "Bilişim Suçları ve Takibi", Adeo Bilişim Teknolojileri Eğitim Merkezi, 2010.

[6] M. Dalyanda, "Adli Bilişim Sistem Analizi", Beyaz Şapka Dergisi, 4, 16-17, 2006.

[7] A. Schroader, Alternate Data Storage Forensics, USA: Syngress, Bölüm 5, 147-169, 2007.

[8] J.Gn.K.lığı, "Bilişim Suçları ve Bilişim Güvenliği Kursu". Kurs Notları, 2015.

[9] İnternet: https://toolbox.googleapps.com/apps/messageheader, (Erişim Tarihi: 7 Nisan 2015).

[10] İnternet: http://mxtoolbox.com/EmailHeaders.aspx , (Erişim Tarihi: 7 Nisan 2015).

[11] İnternet: http://www.iptrackeronline.com/email-header-analysis.php,
(Erişim Tarihi: 7 Nisan 2015).

[12] İnternet: http://www.gaijin.at/en/olsmailheader.php, (Erişim Tarihi: 7 Nisan 2015).

[13] İnternet: https://support.google.com/mail/answer/22454?hl=en, (Erişim Tarihi: 7 Nisan 2015).

[14] İnternet: http://www.nirsoft.net/utils/ipnetinfo.html, (Erişim Tarihi: 7 Nisan 2015).

[15] L Daniel ve L. Daniel, "E-mail Evidence" Digital Forensics for Legal Professionals, USA: Syngress, Bölüm 34, 239-244, 2012.

[16] C. Albrecht, Email Analysis, İnternet: http://www.gsaig.gov/ assets/File/other-documents/Forensics-EmailAnalysis.pptx.pdf, (Erişim Tarihi: 10 Nisan 2015).

[17] H. Çakır ve M.S. Kılıç, "Bilişim Suçlarına İlişkin Delil Elde Etme Yöntemlerine Genel Bir Bakış", Polis Bilimleri Dergisi, 15 (3), 23-44, 2013.

[18] T. Fair, M. Nordfelt, S. Ring ve E. Coler, "Spying on E-mail", Cyber Spying, USA:Syngress, Bölüm 8, 291-316, 2005.

[19] V.K. Devendran, H. Shahriar ve V. Clincy, "A Comparative Study of Email Forensic Tools", Journal of Information Security, 6 (2), 111-117, 2015.

[20] J. Sammons, "Internet and E-Mail" The Basics of Digital Forensics, USA: Syngress, Bölüm 8, 119-131, 2015.

[21] F. Iqbal, H. Binsalleeh, B.C.M. Fung ve M. Debbabi, "Mining writeprints from anonymous e-mails for forensic investigation", Digital Investigation, 7 (1-2), 56-64, 2010.

[22] F. Iqbal, R. Hadjidj, B.C.M. Fung ve M. Debbabi, "A novel approach of mining write-prints for authorship attribution in e-mail forensics", Digital Investigation, 5, 42-51, 2008.

[23] D. Bradbury, "Can we make e-mail secure?", Network Security, 3, 13-16, 2014.

[24] İnternet: https://emkei.cz/. (Erişim Tarihi: 10 Nisan 2015).

[25] M.T. Banday, F.A. Mir, J.A. Qadri ve N.A. Shah, "Analyzing Internet E-mail Date-Spoofing", Digital Investigation, 7 (3-4), 145-153, 2011.

[26] İnternet: http://www.spamhaus.org/sbl/ (Erişim Tarihi: 25 Nisan 2015).

[27] İnternet: https://www.wetransfer.com/ (Erişim Tarihi: 25 Nisan 2015).

[28] P. Loshin, "E-mail Security and Anonymity Practices", Practical Anonymity, USA:Syngress, Bölüm 7, 103-112, 2013.

[29] İnternet: https://www.hushmail.com/ (Erişim Tarihi: 1 Mayıs 2015).

[30] İnternet: http://www.mytrashmail.com/ (Erişim Tarihi: 1 Mayıs 2015).

[31] İnternet: https://www.5ymail.com/ (Erişim Tarihi: 1 Mayıs 2015).

[32] İnternet: http://www.10minutemail.com/ (ErişimTarihi: 1 Mayıs 2015).

[33] P. Wayner, "Anonymous Remailers", Dissappearing Cryptography, USA: Morgan Kaufmann, Bölüm 10, Sayfa 193-230, 2009.

Şeref SAĞIROĞLU, Gazi Üniversitesi Fen Bilimleri Enstitüsü Müdürü ve Bilgisayar Mühendisliği Bölüm Başkanıdır.

Eyüp Burak CEYHAN, Gazi Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümünde araştırma görevlisidir.

Rıfat MARAŞ, Gazi Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği ABD'da Yüksek Lisans öğrenimine devam etmektedir.

MOBİL CİHAZLARDA ZARARLI YAZILIM TESPİTİNDE KULLANILAN STATİK ANALİZ ARAÇLARI

Gülsüm KAYABAŞI, İbrahim Alper DOĞRU

Gazi Üniversitesi Teknoloji Fakültesi
Bilgisayar Mühendisliği Bölümü
glsmkayabasi@gmail.com, iadoğru@gazi.edu.tr

Özet — Günümüz teknolojisinde kullanılan akıllı telefonlar mobil işletim sistemlerine sahiptir. Bu mobil işletim sistemleri her kullanıcı ihtiyacına göre farklılık göstermektedir. Açık kaynak kodlu ve Linux tabanlı bir mobil işletim sistemi olan Android'in yapılan çeşitli araştırmalar sonucunda en popüler işletim sistemi olduğu görülmüştür [1]. Android'in popüler olması ve açık kaynak kod yapısı, kötücül saldırı tiplerinin ve kötü amaçlı saldırganların hedefi olmasına sebep olmuştur. Bu çalışmamızda, mobil cihazlarda zararlı yazılım tespiti yapan statik analiz yöntemlerinden bahsedilmiş ve bu yöntemlerin özellikleri karşılaştırılmıştır.

Anahtar Kelimeler — Mobil işletim sistemi, Android, kötücül yazılım, mobil uygulama güvenliği, statik analiz metotları

I. GİRİŞ

İstenmeyen kötü amaçlı yazılım kapsamına giren virüs, solucan, Truva atları, casus ve reklam içerikli yazılımlar yıllardır kişisel veya kurumsal bilgisayar sistemlerine tehdit oluşturmuşlardır [2]. Akıllı telefon kullanıcılarının sayısı arttıkça, akıllı telefon platformlarının çeşitli tehditlere duyarlı hale gelmesi de bir gerçektir [3]. Bu tehditlerden masaüstü bilgisayarlardan bilindiği gibi, akıllı telefonlar için cep telefonları ve sofistike kompakt minibilgisayarlar da etkilenmektedir [4][5].

Geleneksel bilgisayar sistemlerine benzer şekilde, akıllı platformlarda zararlı yazılım tespiti için birçok çözüm imza tabanlı yaklaşımlara dayanmaktadır [2]. İmza tabanlı ikili dosya eşleştirme yaklaşımının kolay ve verimli olduğu bilinmektedir. Ancak bunun gibi çözümler her zaman zararlı yazılım tespit etmede verimli olmamaktadır çünkü imzaların kapsamı ve kalitesi değişiklik göstermektedir. Diğer bir y.anlış negatif koşulları icat etmiştir. Öte yandan, imzaların oluşumu statik ve dinamik analiz metotlarının oluşmasını gerektirmiştir. Bu metotlar zaman alıcı ve uzmanlık gerektiren interaktif süreçlerdir.

Piyasada bulunun akıllı telefonlardaki kötü amaçlı yazılımlara karşı alınan önlemler imzalama yaklaşımına sahip olduğundan beri zayıflığa maruz kalmıştır. Bu yaklaşım imza kullanılabilir oluncaya kadar kullanıcıları yeni bir zararlı yazılım ile karşı karşıya bırakmıştır. Yapılan çalışmalardan örnek verilecek olursa; Bulygin [6] en kötü durumda 700.000'den fazla cihaz kullanarak bir MMS solucanının rastgele hedeflediği bir telefon rehberindeki numaralara yaklaşık 3 saat içerisinde bulaşabileceğini göstermiştir. Buna ek olarak, Oberhide ve arkadaşları[7], bir imza tabanlı anti virüs motoru için yeni tehditleri tespit etmesi için gerekli ortalama sürenin 48 gün olduğunu değerlendirmiştir. Akıllı telefonlardaki kötücül yazılım için bu sayılar güvenlik önlemleri genişletildiğinde virüslü bir cihazda saniyeler içersinde ciddi bir şekilde zarar verebilir. Bu durumda yeni akıllı telefon işletim sistemlerinden biri olan Android, geliştiriciler arasında özel bir ilgi kazanmıştır. Açık kaynak kurmak için, güvenlik araçları çekirdek seviyesinde geliştirilebilir. Bu da Android telefonlar üzerine kurulan sadece mobil cihazlar için kaynak kısıtları ile sınırlı kapsamlı bir güvenlik mekanizması oluşumuna izin vermiştir [7].

Mobil cihazlardaki kaynak kısıtlamaları nedeniyle, Android cihazlarda zararlı yazılımın varlığının tespit edilmesi için statik ve hafif mekanizmalara odaklanılmıştır. Zararlı yazılım tespiti için kullanılan statik analiz yaklaşımı, çok kaynak tüketmeyen ve mobil ihtiyaçlara çok iyi uyum sağlayan basit sınıflandırıcıların kullanımına izin vermiştir. Önceki yaklaşımlar, mobil cihazın hesaplama yükünü azaltmak için çoğunlukla dış sunuculara bağlıdır. Bu durumda ise, tespit işlemi için sunucudan yararlanılmıştır fakat buna bağlı olmak zorunda değildir. Böylece, ağır öğrenme mekanizması işlemi için, uzak sunucunun entegrasyonundan yararlanılmıştır [8] [9].

Bu çalışmada statik analiz yöntemi kullanan mimarilerden bahsedilmiştir. Çalışmanın geri kalan kısmı şu şekilde yapılandırılmıştır: 2.bölümde mobil cihazlarda kullanılan zararlı yazılım tespit yöntemleri özetlenmiştir, 3.bölümde statik analiz metodunu kullanan mimariler detaylı bir şekilde anlatılmıştır ve 4.bölümde ise sonuçlar ve bu mimarilerin karşılaştırmaları yapılmıştır.

II. MOBİL CİHAZLARDA ZARARLI YAZILIM TESPİTİNDE KULLANILAN METOTLAR

Modern bilgisayar ve iletişim altyapıları saldırılara karşı son derece duyarlıdır. Bu saldırıların başlamasının en önemli sebepleri; solucanlar, virüsler ve Truva atları vb. zararlı yazılımların yayılmasıdır. Bu yayılma ticari şirketlere, özel kullanıcılara ve devletlere ciddi hasarlar verebilir. Bu hasarların yayılmasının en önemli sebebi internet kullanımının yoğunlaşmış olmasıdır. Özellikle yüksek hızda internet bağlantılarındaki son büyüme yeni zararlı yazılımların oluşumunda artışlara sebep olmuştur.

Zararlı yazılım tespiti için çeşitli analiz teknikleri önerilmiştir [10]. Genel olarak statik analiz ve dinamik analiz olmak üzere ikiye ayrılır. Dinamik Analiz (davranışsal tabanlı analiz olarak da bilinir.) yönteminde dosya ve hafiza değişiklikleri, ağ erişimi ve sistem çağrıları gibi işletim sisteminin çalışma zamanında (yani programın yürütülmesi sırasında) toplanan bilgileri içeren br tespit gerçekleştirilir. Örneğin, Panorama adındaki sistem; farklı tiplerdeki zararlı yazılımın ortak işlem davranışlarını (örneğin bilgi akış analizi) ve şüpheli bilgi erişim desenleriyle dinamik yöntemin fizibilitesini çıkarır ve gösterir. Bu tespitlerin dönüşümü mobil ortama yaklaşır ancak akıllı telefonların dayattığı kaynak kısıtlamaları(CPU,güç.hafıza) yüzünden düzgün değildir [10].

Statik Analiz yönteminde; program hakkında bilgi veya ikili/kaynak kodundaki açık veya üstü kapalı gözlemlerin içerdiği davranışların olduğu tahmin edilmektedir. Statik analiz teknikleri sınırlı olmasına ragmen hızlı ve etkilidir. Çeşitli bulanıklaştırma teknikleri statik analiz yöntemini kullanmaktan kaçınmaya sebebiyet verebilir ve böylece sınırlı polimorfik zararlı yazılımlarla başa çıkma yetenekleri kılar.

Dinamik analiz yaklaşımında, çeşitli bulanıklaştırma metotları problemler meydana gelir çünkü programın gerçek davranışı monitörlenir. Bununla birlikte bu yaklaşım diğer dezvantajlara da sahiptir. İlk olarak, programın zararlı fonksiyonları aktif olacağında(örneğin zararlı yazılımın sömürdüğü hassas uygulamalar) uygun durumların simülasyonu zordur. İkincisi, her kötü amaçlı yazılım için kötü niyetli faaliyet görünümünü gözlemleme ihtiyacının ne kadar süre gerektirdiği belirsizdir [10].

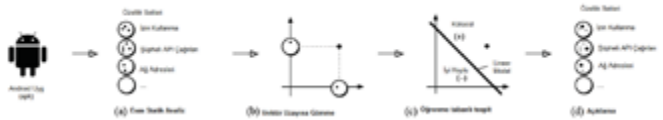
Statik analiz ve dinamik analiz çözümleri imza tabanlı ve sezgisel tabanlı iki metot kullanılarak uygulanmaktadır. İmza tabanlı metot anti virüs üreticileri tarafından sık kullanılan bir metottur ve zararlı yazılımı benzersiz bir imzayla tanımlamaya dayanmaktadır. Çok hassas olduğu durumlarda, imza tabanlı metotlar bilinmeyen bir kötücül yazılım koduna karşı kullanışlı değildir. Sezgisel tabanlı metot, uzmanlar tarafından veya makine öğrenmesi teknikleri ile tanımlanan kötücül veya iyi huylu yazılımlar için meydana gelmiş kurallara dayanmaktadır.

Son zamanlarda, sınıflandırma algoritmaları sezgisel tabanlı metot fikirlerini genişletmek ve otomatikleştirmek için kullanılmaya başlanmıştır. Bu metotlar, dosyanın ikili kodu veya uygulamanın çalışma esnasındaki davranışını temsil etmektedir. Sınıflandırıcılar uygulamaların iyi veya kötü huylu olduğunu sınıflandırarak desenlerin öğrenilmesi için kullanılmaktadır [10].

III. ZARARLI YAZILIM TESPİTİNDE KULLANILAN STATİK ANALİZ METOTLARI

Android işletim sistemi geliştikçe kötücül yazılım çeşitleri ve sayısı da artmaktadır. Bunlara yönelik koruma yöntemleri de birbirlerine paralel şekilde çoğalmaktadır. Android mobil işletim sistemine sahip cihazlardaki uygulamaların kurulumlarından önce uygulamanın bizlere verdiği bilgiler (veriler) sayesinde yapılması statik analiz metodunu açıklar. Statik analiz ile zararlı yazılım tespiti ve korunmasının en büyük avantajı kötücül yazılımın kullandığımız cihaza kurulmadan önce tespit edilmesidir. Böylece cihaz hiçbir şekilde kötücül yazılımın sebep olduğu etkilere maruz kalmaz.

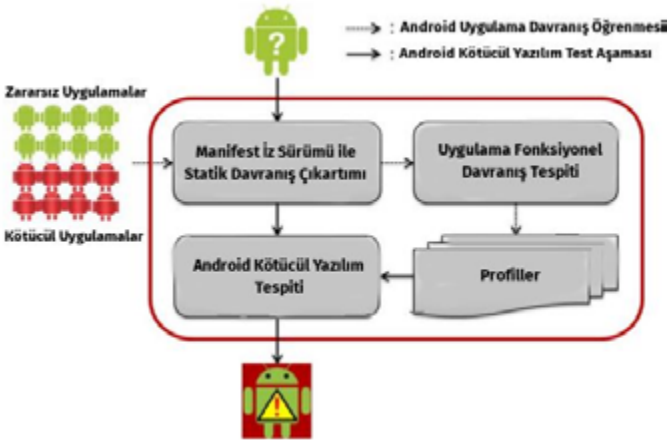
Statik analiz yaklaşımıyla kötücül yazılım tespiti yapan araçlar bulunmaktadır. Bu araçların birkaçından bahsedecek olursak Drebin[11], statik analiz yaklaşımına makine öğrenmesi yaklaşımı da ekleyerek Android'te kötücül yazılım tespitinde kullanılan bir araç olmuştur. Drebin, android uygulamasının kaynak kodlarını ve AndroidManifest dosyasını kullanarak uygulamaya ait izinleri, uygulama programlama arayüzü çağrılarını ve ağ adresleri gibi çeşitli özellikleri toplamaktadır. Bu özelliklerin tamamı bir araya gelerek bir vektör uzayına gömülmüştür ve çıkan desenler yardımıyla kötücül yazılım tespiti yapılmıştır. Drebin kötücül yazılım tespit aracı tarafından yapılan statik analiz adımlarının şematik yapısı Şekil-1'de gösterilmiştir.



Şekil-1 Drebin aracının statik analiz adımları

Şekil-1'de statik analiz yapılırken kullanılan özellik setleri oluşturulmuş ve bunlar vektör uzayının içerisine gömülmüştür. Daha sonra lineer bir model kullanılarak iyi ve kötü huylu uygulamalar meydana çıkarılmıştır.

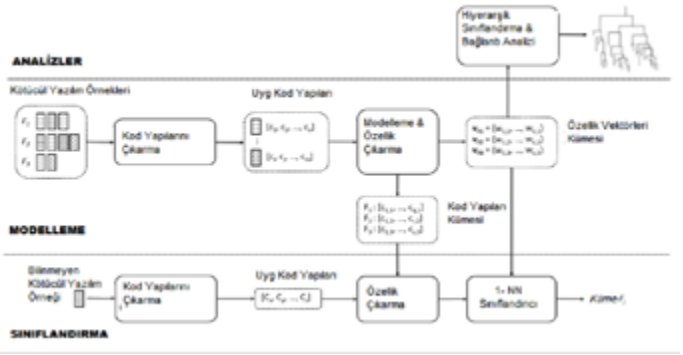
DroidMat[12] aracı ise, AndroidManifest dosyasını ve onunla alakalı izinlere yönelik uygulama programlama arayüzü (API) çağrıları üzerinden kötücül yazılım tespiti yapar. DroidMat mimari yapısı Şekil-2'de gösterilmiştir.



Şekil-2 Droid Mimari Yapısı [13]

Şekil-2'de görüldüğü üzere, manifest dosyası sayesinde statik analiz çıkarımı yapılmıştır. Analizler elde edildikten sonra fonksiyonel davranış tespiti safhasına geçilmiştir. Analizler sonucunda K-means adı verilen bir algoritma ile kötücül yazılım modelleri oluşturulmuştur. Burada oluşacak küme sayısı tekil değer ayrışımı (SVD-Singular Value Decomposition) algoritması ile belirlenmiştir. En sonunda ise k=1 olmak üzere kNN (k Nearest Neighbours) algoritması ile android uygulamasının kötücül olup olmadığı tespit edilmiştir [12].

DENDROID, iki farklı yolla yeni çıkmış bir analiz aracıdır. Bir taraftan, Android mobil işletim sistemindeki daha önce ortaya çıkmamış kötücül yazılım ailelerinin karakteristiklerini belirlerken kod yapılarını kullanma hususunda en iyisidir. Kod metotlarının iç yapısına odaklanmanın komutların özel dizilerine göre önemli bir avantajı bulanıklaştırmaya karşı direnç geliştirmiş olmasıdır. Ayrıca bu tür yapılar, kodların yeniden kullanımına dayalı hızlı geliştirme metodolojilerinin yaygın olduğu yerlerde, akıllı telefonlardaki kötücül yazılımın durumu için kullanışlı olduğu bilhassa kanıtlanmıştır [14].

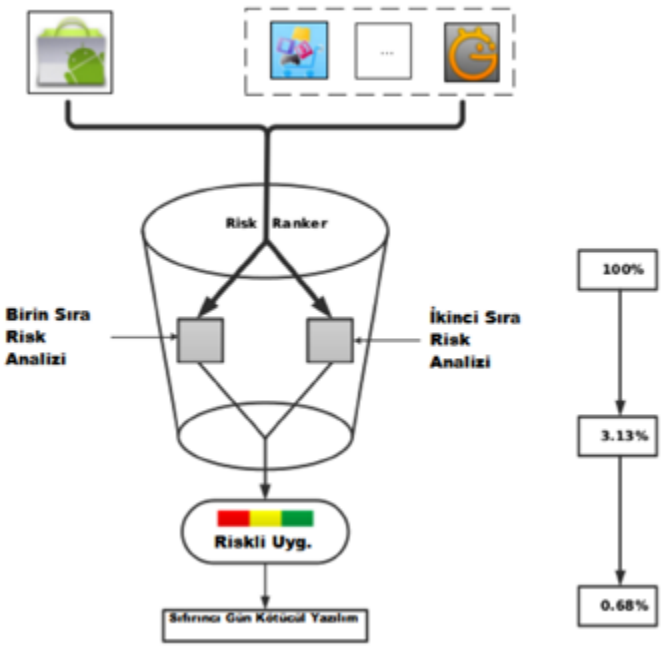


Şekil-3 DENDROID Mimari Yapısı

Öte yandan, örnekleri sınıflandırmak, kod bileşenlerini araştırmak veya kötücül yazılım ailelerinin evrimsel

bağlantıları üzerine çalışma durumlarını otomatikleştirmek için metin madenciliği tekniklerini kullanma fikrini ortaya atmıştır. Ayrıca, metin madenciliği teknikleri büyük miktarda verilerle etkin bir şekilde başa çıkarak geliştirilmiş ve özellikler adreslenen problemler için çok elverişli bir şekilde oluşturulmuştur. DENDROİD mimarisi Şekil-3'te gösterilmiştir [14].

RiskRanker, ölçeklenebilir şekilde ve sıfırıncı gün kötücül yazılımı ortaya çıkarmak için Android marketteki bütün uygulamaları elemek için dizayn edilmiştir. Güvenilmeyen uygulamalardan potansiyel risk taşıyanları ön planda tutarak ve değerlendirerek, ölçeklenebilirlik, verimlilik ve doğruluk gereksinimlerini doğal olarak artıran yönetilebilir bir alan oluşturmak amaçlanmaktadır.Daha ayrıntılı olarak kaynak verimli olma durumuna bakarak, istenilen zamanda yüzlerce veya binlerce uygulamayı ele almak için ölçeklenebilir olmalıdır. Ayrıca sistemin yeterince kısa olan listeyi elle doğrulamak için girdileri verimli bir şekilde ayıklamaya ihtiyacı vardır ve kötücül uygulamaları tespit ettiğini yeterince doğrulamalıdır. RiskRanker mimarisi Şekil-4'te gösterildiği gibidir. RiskRanker'da tehlike içeren riskler düşük, orta ve yüksek seviyede risk olmak üzere 3'e ayrılmıştır. Risk için düşükten yükseğe doğru gidecek olursak uygulamalara erişim güçleştikçe risk seviyesinin azaldığı görülmüştür. Yani risk ile uygulamaya erişim arasında ters bir orantı bulunmaktadır. RiskRanker'da iki kademeli risk analizi yapılmıştır. İlk kademesinde uygulamaların ne derecede riskli olup olmadıklarına bakılmıştır. Alçak seviyede riske sahip olanlar haricindekiler için tehlike içeriği kontrol edilmiştir. 118,318 uygulama kullanılarak test edilmiş bunların 3281 tanesi risk içermiştir. 3281 tane uygulamanın da 322 tanesinin sıfırıncı gün zararlı yazılımı olduğu tespit edilmiştir [15].



Şekil-4 RiskRanker Mimarisi

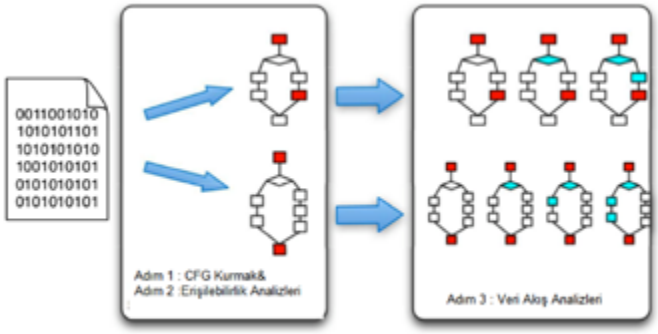
AppProfiler yönteminde vurgulanmak istenen, kötücül yazılım tespit etmek değil, yasal uygulamalar hakkında izin verilen kullanıcılara bilgi sağlamaktır. Doğal kodlar veya uygulamalarla çalışmak analizleri önlemek için olağanüstü adımlar atılmasını gerektirir. Odaklanılan konu Android API'sinin nasıl kullanıldığının veya bu API'nin harici olup olmadığının gösterilmesidir. Korumaları yıkmaya çalışmak AppProfiler altyapısının kapsamı dışındadır. Kabul edilmeyen bir gizlilik ihlali olduğunu belirleme girişimi bulunmamaktadır. Bu kapsamda çoğu davranışın zararsız olduğu keşfedilmiştir. Örneğin, çalındığı zaman izlenen bir mobil uygulama casus yazılımdan ayırt edilebilir. Diğerleri kullanıcıdan kullanıcıya farklılık gösterebilir. Örneğin, konum alan uygulamalarla ilgili bütün kullanıcılar bilgi sahibi değildir. AppProfiler'ın amacı yüklenen uygulamalar hakkında kullanıcılara fikir vermektir [16].

Kapasite sızıntılarının açığa çıkarılmasını kolaylaştırmak için Woodpecker[17] adında bir sistem geliştirilmiştir. Önceden yüklenmiş uygulamalar üzerinde veri akış analizini kullanarak, Woodpecker sistematik olarak açık veya savunmasız arayüzden tehlikeli bir izne erişimi telefonda her uygulama için analiz eder.

Şekil-5 Woodpecker Mimarisi

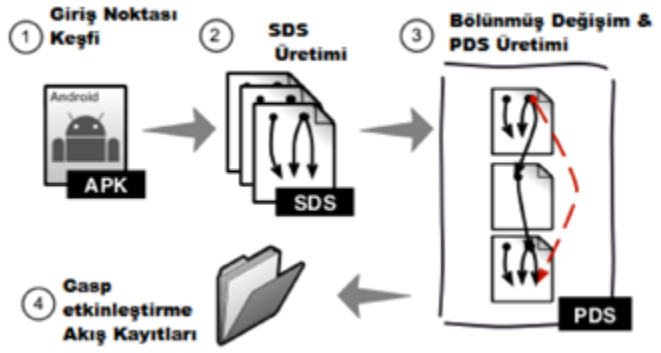
Olası kapasite sızıntılarını en iyi şekilde inceleyen sistemde iki farklı kategori bulunmaktadır. Açık kapasite sızıntıları, kendi kendine izin talep etmeyen servisler veya erişilebilir bazı arayüzlerin kesin izinlere başarılı bir şekilde erişmesine izin verir. Kapalı kapasite sızıntıları ise bazı açık arayüzler veya servisler dışındakilere izin verir. İmzalama anahtarının aynısıyla diğer uygulamanın aldığı izinleri aktarılır veya elde edilir. Sonuç olarak açık sızıntılar ciddi güvenlik hataları verir. Kapalı sızıntılar ise Android'te izin tabanlı güvenlik modelini yıkar ve uygulamaya yeteneklerini yanlış sunar. Woodpecker mimarisi Şekil-5'te gösterildiği gibidir [17].

PiOS (Privacy İOS), ilk olarak uygulamanın hassas veriye eriştiği ve sonradan ağ üzerinde verinin iletildiği yerde kod yollarının varlığı için uygulamaları kontrol edebilmek adına statik analiz yöntemini kullanmıştır. PiOS, binarilerde doğrudan analiz gerçekleştirmelidir çünkü kaynak kodu mevcut değildir. Statik analiz gibi ikili analiz yöntemi de zor bir yöntemdir. Bunun yanı sıra İOS uygulamalarının çoğu nesne tabanlı C programlama dili ile geliştirilmiş olup daha da karmaşıklaşmıştır. PiOS mimarisini Şekil-6'da görebiliriz [18].



Şekil-6 PiOS Mimarisi

Android uygulamalardaki bileşen kaçırma açıklıklarını tespit edebilmek için hataları otomatik olarak ayıklamak işlevini gerçekleştiren CHEX sistemi statik analiz metodu önerilmiştir. Bir veri akış analizi açısından bu açıklıkların modellenmesi, CHEX'in Android uygulamaları analiz etmesi ve sistemin bağımlı grafiklerinde erişilebilir testle yapması ile gerçekleşmektedir. Android'in özel programlama paradigması tarafından kullanıcıya zorla sunulan analiz zorluklarıyla mücadele etmek için uygulamada asenkron uygulamaların çoklu giriş noktaları modellenmiştir ve bu giriş noktası bileşenleri keşfetmek ve bütünlüğü korumak için yeni bir teknik kullanılmıştır [19].



Şekil-7 CHEX İş Akış Diyagramı

CHEX (Component Hijacking Examiner), Dalysis (Dalvik Bytecode Analysis) örnek alınarak geliştirilmiştir. Android uygulamasının baytkodunda analizlerin birçok tipini desteklemek için kurulan kapsamlı statik analiz çerçevesidir. CHEX'in çalışma prensibine göre elimizdeki Android uygulama alınarak önce veri akış özetine bölünür ve daha sonra bölünen parçalardan bitişik olanlar birbirlerine bağlanır. CHEX, 5486 Android uygulaması incelenmiş ve 254 tanesinde potansiyel bileşen kaçırma açıklığı bulunmuştur. CHEX'in bir uygulamada ortalama çalışma süresi 37.02 saniyedir. Bu yeterince hızlı olup test senaryoları ve uygulama incelemeleri için kullanılmaktadır. CHEX'in iş akışı Şekil-7'de görüldüğü gibidir [19].

IV. SİSTEM KARŞILAŞTIRMALARI

Sistem karşılaştırmaları sadece statik analiz metotu kullanan analiz metotları için monitörleme tipi (sistem çağrıları, ağ, olay günlüğü, talimatlar, izinler, program izleri, işlem kontrol blokları, API çağrıları, çekirdek seviyesi ve kullanıcı seviyesi), analiz tipi (uzman, makina öğrenmesi, sınıflandırma, bağımlı grafikler, istatistikler ve olasılık modelleri), teşhis tipi (anomali, suistimal ve spesifikasyon), monitörleme ve teşhiş yeri ve analiz yeri (yerel anahat, IRM, bulut, dağıtık, balküpü, bulut yineleme, sandbox, hibrit) olarak yapılacaktır. Bu yönemlerin karşılaştırmalı haline Tablo-1'de ulaşabilirsiniz.

Tespit Yaklaşımı	DENDROID	AppProfiler	RiskRanker	WoodPecker	CHEX	PiOS
Platform	Android	Android	Android	Android	Android	IOS
Monitörleme	Talimatlar	API Çağrıları, Program İzleri	Talimatlar, İzinler, API Çağrıları	Talimatlar, İzinler	Talimatlar	Yok
Analiz Tipi	İstatistiksel, Bağımlı Grafikler, Sınıflandırma	Uzman	Bağımlı Grafikler	Bağımlı Grafikler	Bağımlı Grafikler	Bağımlı Grafikler
Teşhis Tipi	Yok	Suistimal	Suistimal	Yok	Yok	Yok
Monitörleme ve Teşhis Yeri	Bulut	Yerel Anahat	Bulut	Bulut	Bulut	Bulut
Analiz Yeri	Bulut	Bulut,Yerel Anahat	Bulut	Bulut	Bulut	Bulut

Tablo-1 Statik Analiz Metotlarının Karşılaştırılması

Tablo-1'de görüldüğü gibi, tespit yaklaşımlarına göre yapılan karşılaştırmada Android platformunu kullanmayan tek araç PiOS aracıdır ve monitörleme ve teşhiş yaklaşımlarını kullanmamaktadır. Monitörleme yaklaşımında AppProfiler'ın Talimatlar tipini kullanmadığı, Analiz Tipi karşılaştırmasında AppProfiler dışındakilerin bağımlı grafiklerden faydalandığı, Teşhis Tipi yaklaşımını suistimal boyutunda AppProfiler ve RiskRanker'ın kullandığı, Monitörleme ve Teşhis Yeri karşılaştırmasında sadece AppProfiler'ın yerel anahat üzerinde çalıştığı diğerlerinin bulutta çalıştığı ve son olarak Analiz Yeri kıyaslanmasında ise hepsinin bulut kullandığı AppProfiler'da ek olarak yerel anahatta çalıştığı görülmektedir.

V. YAPILAN DEĞENLENDİRMELER VE SONUÇ

Bu çalışmada, dünya çapında kullanımı en yaygın olan mobil işletim sistemi Android için zararlı yazılım tespitinde kullanılan statik analiz metotları anlatılmış ve karşılaştırmalı olarak incelenmiştir. İnceleme sonucunda tespit yaklaşımlarına göre statik analiz yöntemiyle zararlı yazılım tespiti yapan araçların farklı yaklaşım yöntemleri olduğu kadar benzer yöntemlere sahip oldukları da görülmüştür.

Kıyaslama yapılan statik analiz araçları arasında makina öğrenmesi yaklaşımını kullanan bir metot bulunmamaktadır. İnceleme neticesinde herhangi bir Android uygulaması geliştirildikten sonra güvenlik açıklarının temelini oluşturan izin mekanizmaları uygulamanın kaynak kodları ile beraber analiz edilmesinin gerekli olduğu anlaşılmıştır. Ayrıca uygulamanın kurulumu öncesinde zararlı yazılım tespiti için uygulamanın kullanıcılarını yönlendirecek daha ayrıntılı zararlı yazılım tespit araçlarına da ihtiyaç olduğu yapılan araştırmalar sonucunda görülmüştür.

REFERENCES

[1] http://www.idc.com/prodserv/smartphone-os-market-share.jsp

[2] A. Schmidt, R. Bye, H. Schmidt, 1. Clausen, O. Kiraz, K. A. Yuksel, S. A. Camtepe, and S. Albayrak, "Static Analysis of Executables for Collaborative Malware Detection on Android", Proceedings of the 2009 IEEE international conference on Communications, 2009, pp. 63 1 - 63S.

[3] S. Seo, D. Lee, and K. Yim, "Analysis on maliciousness for mobile applications", Proceedings of 20 12 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), pp. 126-129.

[4] A. Shabtai, Y. Fledel, and Y. Elovici, "Automated Static Code Analysis for Classifying Android Applications Using Machine Learning", Proceedings of 20 I 0 International Conference on Computational Intelligence and Security, 20 10, pp. 329-333.

[5] A. Shabtai, "Malware Detection on Mobile Devices", Proceedings of 20 10 Eleventh International Conference on Mobile Data Management (MDM), 20 10, pp. 289 - 290.

[6] Y. Bulygin, "Epidemics of mobile worms," in Proceedings of the 26th IEEE International Performance Computing and Communications Conference, IPCCC 2007, April 11-13, 2007, New Orleans, Louisiana, USA. IEEE Computer Society, 2007, pp. 475–478.

[7] J. Oberheide, E. Cooke, and F. Jahanian, "Cloudav: N-version antivirus in the network cloud," in Proceedings

of the 17th USENIX Security Symposium (Security'08), San Jose, CA, July 2008. IEEE Criteria f or Class IE Electric Systems (Standards style), IEEE Standard 308, 1969.

[8] D. Samfat and R. Molva, "IDAMN: An Intrusion Detection Architecture for Mobile Networks," IEEE Journal on Selected Areas in Communications, vol. 15, no. 7, pp. 1373–1380, Sep. 1997.

[9] A. Bose, X. Hu, K. G. Shin, and T. Park, "Behavioral detection of malware on mobile handsets," in Proceeding of the 6th international conference on Mobile systems, applications, and services. Breckenridge, CO, USA: ACM, 2008, pp. 225–238.

[10] A. Shabtai, U. Kanonov, Y. Elovici, C. Glezer, Y. Weiss, "Andromaly": a behavioral malware detection framework for android devices, J. Intell. Inf. Syst. 38 (1) (2012) 161–190.

[11] Arp D, Spreitzenbarth M, Malte H, Gascon H, Rieck K. Drebin: effective and explainable detection of Android malware in your pocket. In: Symposium on network and distributed system security (NDSS); 2014. p. 23e6.

[12] D.-J. Wu, C.-H. Mao, T.-E. Wei, H.-M. Lee, and K.-P. Wu, "DroidMat: Android Malware Detection through Manifest and API Calls Tracing," in 2012 Seventh Asia Joint Conference on Information Security, 2012, pp. 62–69.

[13] Abdullah Talha KABAKUŞ, İbrahim Alper DOĞRU, Aydın ÇETİN, "Android kötücül yazılım tespit ve koruma sistemleri," in Erciyes Üniversitei Fen Bilimleri Enstitüsü Dergisi, 31(1):9-16

[14] G. Suarez-Tangil, J. E. Tapiador, P. Peris-Lopez, and J. B. Alis, "Dendroid: A text mining approach to analyzing and classifying code structures in android malware families," Expert Systems with Applications, 2013, in Press.

[15] M. Grace, Y. Zhou, Q. Zhang, S. Zou, and X. Jiang, "Riskranker: scalable and accurate zero-day android malware detection," in Proc. 10th int. conf. on Mobile systems, applications, and services. ACM, 2012, pp. 281–294.

[16] S. Rosen, Z. Qian, and Z. M. Mao, "Appprofiler: a flexible method of exposing privacy-related behavior in android applications to end users," in Proc. 3rd ACM conference on Data and application security and privacy. ACM, 2013, pp. 221–232.

[17] M. Grace, Y. Zhou, Z. Wang, and X. Jiang, "Systematic detection of capability leaks in stock android smartphones," in Proc. 19th Annu. Symp. on Network and Distributed System Security, 2012.

[18] M. Egele, C. Kruegel, E. Kirda, and G. Vigna, "Pios: Detecting privacy leaks in ios applications," in Proc. Network and Distributed System Security Symp., 2011.

[19] L. Lu, Z. Li, Z. Wu, W. Lee, and G. Jiang, "Chex: statically vetting android apps for component hijacking vulnerabilities," in Proc. 2012 ACM conf. on Computer and communications security. ACM, 2012,pp. 229–240.

Gülsüm KAYABAŞI, Ankara'da doğdu. İlk ve orta öğrenimini Ankara'da, yükseköğrenimini Eskşehir'de tamamladı. 2010 yılında Eskişehir Osmangazi Üniversitesi Bilgisayar Mühendisliği Bölümü'nden mezun oldu. 2011 yılından beri bir kamu kuruluşunda veritabanı yöneticisi olarak çalışmaktadır. 2013 yılında Gazi Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı'nda yüksek lisans eğitimine başladı. İlgi alanları; mobil uygulamalar, veri güvenliği, web yazılımı..

İbrahim Alper DOĞRU, 2004 yılında Atılım Üniversitesi Bilgisayar Mühendisliği bölümünden mezun olmuştur. Yüksek lisansını 2007 yılında Gazi Üniversitesi Bilgisayar Mühendisliği bölümünde tamamlamıştır. 2012 yılında Gazi Üniversitesi Elektronik-Bilgisayar eğitimi bölümünde doktorasını tamamlamıştır. Halen Gazi Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı'nda Yardımcı Doçent olarak görev yapmaktadır. İlgi Alanları; mobil ağ teknolojileri, mobil tasarsız ağlar, mobil güvenlik, ağlarda adli bilişim

PoS SİSTEMLERİNE YÖNELİK RAM KAZIMA SALDIRILARININ İSTATİSTİKSEL ANALİZİ VE SAVUNMA ÖNERİLERİ

Ecir Uğur Küçüksille, Bekir Eray Katı, Mehmet Ali Yalçınkaya

E.U. Küçüksille, Süleyman Demirel Üniversitesi Bilgisayar Mühendisliği Bölümü, Merkez, Isparta, Türkiye; (telefon: +90(246)2111478;e-posta: ecirkucuksille@sdu.edu.tr)
B.E. Katı, Süleyman Demirel Üniversitesi Bilgisayar Mühendisliği Bölümü, Merkez, Isparta, Türkiye (telefon: +90(537)5715271; e-posta: bekireraykati@gmail.com)
M.A. Yalçınkaya, Süleyman Demirel Üniversitesi Bilgisayar Mühendisliği Bölümü, Merkez, Isparta, Türkiye (telefon: +90(246)2111381; e-posta: mehmetyalcinkaya@sdu.edu.tr)

Özet — Kredi kartlarının yaygınlaşması ile birlikte PoS (Point of Sale) cihazları neredeyse bütün endüstriyel sektörlerde yaygın olarak kullanılmaya başlanmıştır. Kredi kartlarının kullanımın bu kadar yaygınlaşması, söz konusu kartlara ait verilerin elde edilmesi amacıyla gerçekleştirilen saldırılarda da doğru orantılı olarak artış meydana getirmiştir. Bu çalışmada kredi kartlarına ait verilerin elde edilebilmesi amacıyla gerçekleştirilen RAM kazıma saldırıları incelenerek, kurum ve kuruşlara söz konusu saldırılara karşı alınabilecek önlemler sunulmuştur. Ayrıca RAM kazıma saldırılarında kullanılan zararlı yazılımlar incelenmiş, söz konusu saldırıların farklı sektörlerde yıllara göre neden olduğu zararlar, araştırma raporlarından elde edilen istatistiki veriler ışığında analiz edilmiştir. Gerçekleştirilen çalışma RAM kazıma saldırılarına derinlemesine bir bakış sunarak, söz konusu saldırıların etkilerini minimuma indirmek adına çözümler sunmaktadır

Anahtar Kelimeler — PoS, Kredi Kartı, RAM Kazıma, Track, REG-EX Algoritması, Luhn Algoritması.

Abstract — With the expansion of credit cards, PoS devices have been used widely almost in all industrial sectors. Because of the widespread usage of credit cards, RAM scraping attacks which are performing in order to obtain data from said card, increased proportionally. In this study, we investigated the RAM scraping attacks that carried in order to obtain credit card data and we present different defense methods for companies to prevent these attacks. Also we examined malwares used in these attacks and damages caused in different sectors many years examined making use of the statistical data that obtained from research reports. This study provides an in-depth overview to RAM scraping attacks and offers solutions in order to minimize the effects of such attacks.

Keywords — PoS, Credit Card, RAM Scraping, Track, REG-EX Algorithm, Luhn Algorithm.

I. GİRİŞ

Günümüzde gerek gerçek hayatta gerekse sanal ortamda kredi kartlarının kullanım alanı oldukça yaygınlaşmıştır. İnsanlar beraberlerinde yüklü miktarda nakit para taşımak yerine bankamatik veya kredi kartı taşımayı tercih etmektedirler. Kredi kartlarının giderek yaygınlaşması, söz konusu kartları ve kartlar ile alışveriş işlemlerinde kullanılan diğer sistemleri açık birer hedef haline getirmiştir. Saldırganlar gerçekleştirdikleri saldırılarda, tüketiciler tarafından kullanılan kredi kartlarına ait verileri ele geçirmeyi amaçlamaktadır. Bu amaçla gerçekleştirilen RAM kazıma saldırılarında saldırganlar, kredi kartları ile gerçekleştirilen alışveriş işlemlerinde kullanılan PoS cihazlarının bağlı olduğu sunucuları hedef almakta ve birçok kredi kartı kullanıcısının track (manyetik bant verisi) bilgilerini ele geçirmektedir.

Uluslararası akademik dünyada RAM kazıma saldırıları konusunda yapılmış çeşitli çalışmalar mevcuttur. Fakat ülkemizde henüz POS cihazlarına yönelik gerçekleştirilen RAM kazıma saldırılarını inceleyen bir çalışma bulunmamaktadır. Literatürde yer alan yabancı kaynaklı çalışmalar incelendiğinde söz konusu çalışmaların RAM kazıma saldırılarının metodolojisi ve etkilerini detaylı olarak incelediği görülmektedir. Gerçekleştirilmiş olan bu çalışmanın literatürde yer alan diğer çalışmalardan farkı ise, RAM kazıma saldırıları konusunda, çeşitli araştırma raporlarından elde edilen en güncel istatistiksel verilerin toplanmış olması, böylelikle de söz konusu saldırıların tarihsel gelişimi, işletim sistemlerine göre görülme oranları, ülkelere göre görülme oranları, farklı endüstriyel alanlarda görülme oranları ve bu oranların yıllara göre değişimi gibi konularda istatistiksel bir bakış açısı sunmasıdır.

Bu çalışmada 2.Bölümde RAM kazıma saldırılarının genel metodolojisinin yanı sıra kredi kartlarına ait verilerin saklandığı çeşitli veri grup formatlarına değinilmiştir. Bu bölümde ayrıca RAM kazıma saldırılarında, saldırganların kredi kartlarına ait verileri elde etmek için kullandıkları çeşitli algoritmalara değinilmiştir. 3. Bölümde saldırganların RAM kazıma saldırılarında izledikleri saldırı senaryolarına değinilmiş ve RAM kazıma saldırılarının işletim sistemlerine göre gerçekleşme oranları incelenmiştir. 4.Bölümde saldırganlar tarafından RAM kazıma saldırılarında kullanılan zararlı yazılımların tarihsel gelişimi ve işleyişi incelenmiştir. Gerçekleştirilen çalışmanın 5. Bölümünde, POS cihazlara yönelik gerçekleştirilen RAM kazıma saldırılarının diğer saldırılara oranı ve söz konusu saldırıların ülkesel olarak dağılımı, ülkemizde gerçekleşmiş örnekleri sunularak, istatistiksel grafiklerle incelenmiştir. Gerçekleştirilen çalışmanın 6. Bölümünde RAM kazıma saldırılarına yönelik alınabilecek önlemler sunulmuş, 7. Bölümde ise gerçekleştirilen saldırıların sonuçlarına değinilmiştir.

II. RAM KAZIMA SALDIRILARININ METODOLOJİSİ VE KREDİ KARTI BİLGİLERİNİN ELE GEÇİRİLMESİ

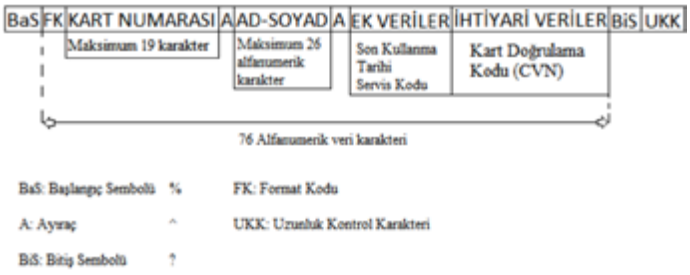
RAM kazıma saldırılarında, kredi kartıyla PoS cihazı kullanılarak bir işlem yapıldığında kart verileri öncelikle şifrelenmiş bir şekilde sunuculara iletilmektedir. Ödeme doğrulama aşamasında şifrelenmiş olan bu veriler çözümlenerek, sunucuda bulunan RAM üzerinde çok kısa bir süreliğine tutulmaktadır. PoS sunucusu üzerine, saldırganlar tarafından çeşitli saldırı metotları aracılığı ile yerleştirilen RAM kazıyıcı zararlı yazılımlar, RAM üzerinde tutulan bu kredi kartı verilerini çeşitli algoritmaları kullanarak tespit etmektedir. Elde edilen bu veriler içinde kart numarası, kullanıcı adları, adresler, söz konusu kartın son kullanım tarihi ve kart doğrulama kodu (CVN) gibi verileri barındıran Track veri grupları bulunmaktadır [1]. Gerçekleştirilen saldırıların ve elde edilen verilerin daha anlaşılır olması açısından kredi kartlarında kullanılan çeşitli veri kayıt tiplerine değinmek gerekmektedir.

Manyetik şerite sahip kredi kartlarında Track adı verilen 3 farklı veri grupları tutulabilir. Bu çalışma kapsamında incelenmiş olan ödeme amaçlı kullanılan kartlarda, Track 1 ve Track 2 veri grupları tutulmaktadır. Bu veri grupları Uluslararası Standartlar Teşkilâtı (ISO) ve Uluslararası Elektroteknik Kurulu (IEC) tarafından standart hale getirilmiştir [2].

Track 1 veri grubu ilk olarak Uluslararası Hava Taşıma Birliği (IATA) tarafından oluşturulmuştur. Track 1 veri grubunu kullanan kartlarda manyetik şerit üzerinde 79 alfa numerik karakter içeren 210 bit veri depolanabilmektedir.

Track 2 veri grubu ise Amerikan Bankacılar Birliği (ABA) tarafından oluşturulmuştur. Track 2 veri grubunu içeren kartlarda manyetik şerit üzerinde 40 alfa numerik karakter içeren 75 bit veri depolanabilmektedir [3]. Şekil 1' de Track 1 veri grubu formatları gösterilmektedir.

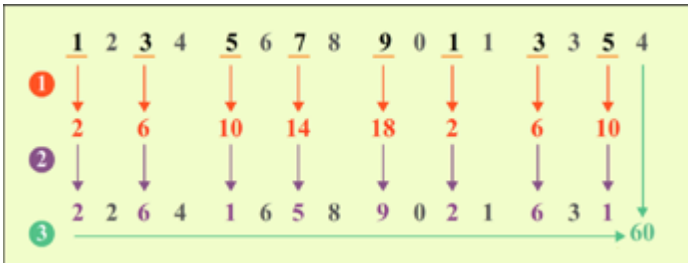
Kredi kartlarında ayrıca 3 veya 4 haneli olan CVN bulunmaktadır. Kart doğrulama kodu, kredi kartı markasına göre CAV, CID, CVC, CVV şeklinde isimlendirilebilmektedir. 1 Ocak 1997'den itibaren tüm Mastercard'larda, 1 Ocak 2001'den itibaren de tüm Visa'larda güvenlik kodu bulunması zorunlu hale gelmiştir. Kart basımı sırasında kart numarası ve kartın son kullanma tarihi bankanın belirlediği bir algoritmadan geçirilerek güvenlik numarası oluşturulmaktadır. Her kart için bu numaralardan 2 adet üretilmektedir. Bunların bir tanesi (CVV) manyetik bant içindeki Track veri grubu içerisinde saklanırken, diğeri (CVV2) genellikle kartın arka yüzeyinde imza bandının olduğu yerde bulunmaktadır. CVV2 numarası sanal harcamalarda kartı fiziksel olarak doğrulamak için kullanılmaktadır [4].



Şekil 1: Track1 veri grubu formatı

RAM kazıma saldırıların POS sunucularına saldırganlar tarafından çeşitli yöntemler ile yerleştirilmiş olan zararlı yazılımlar, RAM üzerinde yer alan kart numarasını ararken REG-EX (Regular-Expression) algoritmasını kullanmaktadırlar. Günümüzde kredi kartı numarasını doğrulamak için kullanılan olan ilk algoritma REG-EX algoritmasıdır. Her bir kredi kartı, kendi finansal şirketinin belirledikleri kurallar çerçevesinde bir kart numarasına sahiptir. Söz konusu kart numarasının uzunluğu 13 ile 16 karakter aralığındadır. İlk 4 hane kredi kartı türlerinin birbirinden ayırt edilmesinde kullanılmaktadır [5]. Regex algoritmasının işleyişine değinmeden önce farklı kart türlerine ait kart numaralarına değinmek gerekmektedir. Visa tipi kartlarda, kart numaraları her zaman 4 ile başlamaktadır. Daha sonra gelecek olan 12 adet rakam ise 0 ile 9 arasındaki rakamlardan oluşmaktadır. MasterCard tipi kartlarda tanımlı kart numaralarının ilk 2 hanesi 51 ile 55 arasındaki sayılardan oluşmaktadır. Geri kalan 14 hane ise 0 ile 9 arasındaki rakamlardan oluşmaktadır. Söz konusu kartlarda kart numarası toplamda 16 haneden oluşmaktadır. American Express tipi kartarda kart numarası 34 veya 37 ile başlamaktadır. Geri kalan 13 hane ise 0 ile 9 arasındaki rakamlardan oluşmaktadır. American Expres tipi kartlarda kart numaraları toplamda 14 haneye sahiptir.

Saldırganlar bahsedilen tiplerdeki kart numaralarını tespit edebilmek için REG-EX algoritmasını kullanmaktadır. Söz konusu algoritma kullanılarak gerçekleştirilen arama işlemlerinde, iki adet \b tagı arasına kart tipine ait tanımlama yerleştirilerek arama gerçekleştirilmektedir. Realgoritması kullanılarak Visa tipindeki bir karta ait kart numarasının tespit edilebilmesi için "\b4[0-9]{12}(?:[0-9]{3})?\b" bloğu kullanılmaktadır. Söz konusu blokta "\4" ifadesinden sonra yer alan "4" ifadesi, aratılacak kart numarasının ilk hanesinin, 4 olduğunu, "[0-9]{12}" ifadesi ise daha sonra gelen 12 rakamın 0 ile 9 arasındaki rakamlardan oluştuğunu belirtmektedir. Saldırganlar gerçekleştirdikleri aramalarda söz konusu bloğu, ilgili kart tipine göre değiştirerek, kart numarasını elde etmeyi amaçlamaktadırlar. Gerçekleştirilen RAM kazıma saldırılarında, kredi kartı numarasının tam olarak tespit edilmesinde regex algoritması tek başına yeterli değildir. Çünkü RAM' den kazınan verileri içerisinde geçersiz verilerde bulunabilmektedir. Saldırganlar regex algoritmasını kullanarak elde etmiş oldukları verilerin doğruluğunu onaylamak için Luhn Algoritması kullanmaktadırlar. Tespit edilen kart numaralarının Luhn algoritması kullanılarak geçerliliğinin doğrulanması işlemi Şekil 2' de gösterilmektedir.



Şekil 2: Kredi kartları için kullanılan Luhn algoritmasının gösterimi [6]

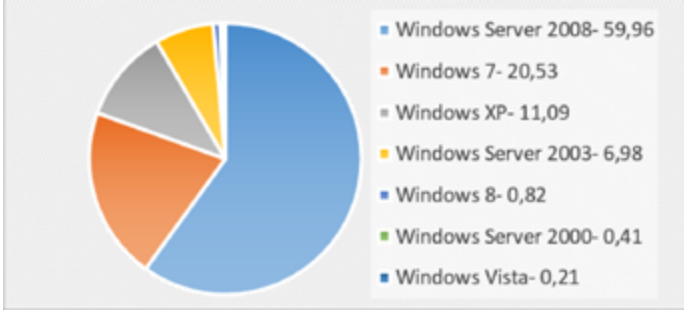
Luhn algoritması kullanılarak elde edilen kart numarasının doğrulanması işleminde ilk olarak, kart numarasında tek haneye denk gelen sayılar 2 ile çarpılmaktadır. Eğer elde edilen sonuç 2 basamaklı bir sayı ise, söz konusu sayının rakamları toplanmaktadır. Eğer elde edilen sonuç tek rakamlı ise, olduğu gibi bırakılmaktadır. Son olarak tespit edilen kart numarasında çift haneye denk gelen rakamlar ile birlikte, bir önceki adımdan elde edilen sonuçlar toplanmaktadır. Gerçekleştirilen tüm işlemler sonunda elde edilen toplam sonuç 10'a tam bölünebiliyorsa, RAM' den kazınmış olan kart numarası Luhn algoritmasına göre geçerli bir kart numarasıdır [6].

III. RAM KAZIMA SALDIRI KAYNAKLARI

Saldırganlar tarafından gerçekleştirilen RAM kazıma saldırılarında, şirket içi, sosyal mühendislik ve oltalama gibi birçok saldırı kaynağı bulunmaktadır. Söz konusu saldırı kaynaklarından ilki şirket içinden gerçekleştirilen saldırılardır. Şirket içinden gerçekleşecek bir saldırı, savunulması en güç saldırı türüdür. Şirket içinde çalışan kötü niyetli kişilerin doğrudan sistem sunucularına yazılımsal veya donanımsal olarak erişebilmesi mümkündür. RAM kazıma amacıyla kullanılacak zararlı yazılım, bir USB aracılığıyla kolaylıkla sunucu yazılımına bulaştırılabilir.

RAM kazıma saldırılarında karşılaşılan diğer saldırı kaynakları da yemleme ve sosyal mühendisliktir. Saldırganlar şirket

içinde çalışan yetkili kişilere e-posta yolu ile ulaşarak şirket içerisinde yer alan sistemlere zararlı yazılımlar bulaştırabilmektedir. Söz konusu saldırı türünde çoğunlukla kendisini bir banka yetkilisi olarak tanıtan bir saldırgan, şirket içinde yer alan çalışana sahte bir mail göndermektedir. Bu mail içerisinde, eklenmiş zararlı bir yazılım ya da zararlı yazılımı barındıran, kurbandan tıklaması istenen bir link içermektedir. Mail ekindeki ya da gönderilen linkteki zararlı yazılımın indirilip çalıştırılması ile saldırganlar, şirket iç ağında yer alan bir sistem üzerinde oturum elde edebilmektedir. Oturum elde edilen bu sistem, küçük boyuttaki bir şirkette yetersiz kaynaklardan dolayı hem kişisel bilgisayar hem de PoS sunucu olarak hizmet veriyor olabilir. Böyle bir durumda saldırgan doğrudan PoS sunucuya erişim sağlamış olmaktadır [7]. Büyük boyuttaki şirketlerde ise PoS sunucu olarak hizmet veren sistemler, kullanıcılara ait sistemlerden ayrı tutulmaktadır. Bu tarz durumlarda saldırganlar, oturum elde etmiş oldukları şirket iç ağındaki bir sistemden, ağ pivotlama (network pivoting) işlemi ile şirket içerisinde yer alan PoS sunuculara erişim sağlayabilmektedir.



Şekil 3: Yapılan RAM kazıma saldırılarının işletim sistemine göre dağılımı [8]

Ayrıca günümüzde kurumlar tarafından kullanılmakta olan birçok yazılımda veya işletim sisteminde güvenlik açıkları bulunmaktadır. Bu güvenlik açıkları gerek yazılım test uzmanları, gerekse kullanıcı geri bildirimleriyle incelenip uygun güncelleştirmeler ile giderilmektedir. Kurum içerisinde kullanılan yazılım ve işletim sistemlerinin sürekli olarak güncel tutulması, PoS saldırılarına karşı hayati önlem taşımaktadır. Şekil 3' te bugüne kadar gerçekleştirilmiş RAM kazıma saldırılarının işletim sistemlerine göre oranları verilmiştir. Söz konusu grafik incelendiğinde, toplam saldırıların neredeyse %60' ının Windows Server 2008 işletim sisteminin kullanıldığı sunuculara yönelik olarak gerçekleştiği görülmektedir [8], [9].

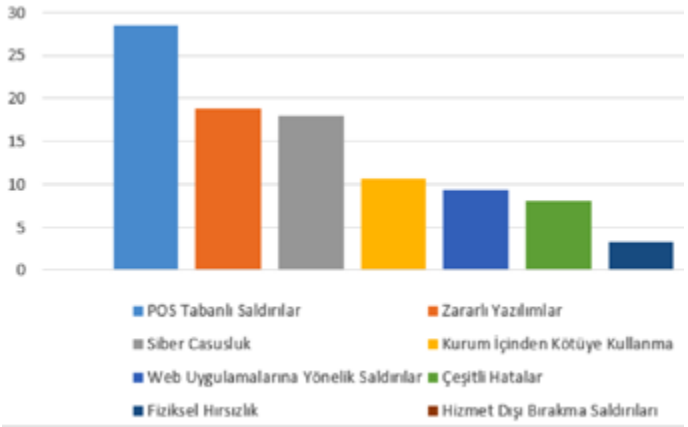
IV. RAM KAZIMA SALDIRINDA KULLANILAN ZARARLI YAZILIMLAR

Kredi kartı kullanımı ve PoS sistemlerinin yeni yeni yaygınlaşmaya başladığı yıllarda saldırganlar basit keyloggerlar kullanarak tek bir hedef cihaz üzerine çeşitli saldırılar gerçekleştirmişlerdir. Gelişen teknoloji ile birlikte kredi kartları ve PoS cihazları için köklü güvenlik değişimlerine gidildiğinden dolayı saldırganlar, güvenlik önlemlerini aşmak için yeni yollar aramaya, yeni yazılımlar geliştirmeye başlamışlardır. Bu bölümde PoS cihazlara yönelik RAM kazıma saldırılarında kullanılan zararlı yazılımlar incelenmiştir. 2012 yılına kadar PoS sunucularına yönelik olarak gerçekleştirilen büyük çaplı saldırılarda BlackPOS isimli zararlı yazılım rol almıştır. BlackPOS, 2010 yılında Rus hacker Sergey Taraspov tarafından geliştirilmiştir. Söz konusu zararlı yazılım, indirildiği sistem üzerinde yer alan bellek içerisindeki kart verilerini tespit etmektedir. Söz konusu arama işleminde BlackPOS, bir önceki başlık altında incelenmiş olan arama ve doğrulama algoritmaları yardımı ile eş zamanlı olarak hem kredi kartı bilgisine uygun olabilecek tüm verileri taramakta hem de bulduğu verilerin geçerliliğini test etmektedir. BlackPOS aynı zamanda, söz konusu bellekten elde ettiği ek verileri, ilerleyen süreçte çevrimdışı ortamda inceleyip filtrelemek amacıyla depolayabilmektedir. PoS sunuculara yönelik gerçekleştirilen saldırıların büyük kısmında etkin olarak kullanılan BlackPOS, diğer PoS zararlı yazılımlarının geliştirilmesinde önemli bir paya sahiptir [10]. BlackPOS zararlı yazılımının geliştirilmesi ile elde edilen POSCard, Chewbacca, Reedum ve BrutPOS zararlı yazılımları 2012 ve 2015 yılları arasında gerçekleştirilen saldırılarda etkin bir şekilde kullanılmıştır. Şekil 4' te RAM kazıma saldırılarında kullanılan zararlı yazılımların zamana bağlı oluşum grafiği ve söz konusu zararlı yazılımlar arasındaki ilişkiler gösterilmektedir.

Şekil 4: RAM kazıma saldırılarında kullanılan zararlı yazılımların tarihsel gelişimi [10]

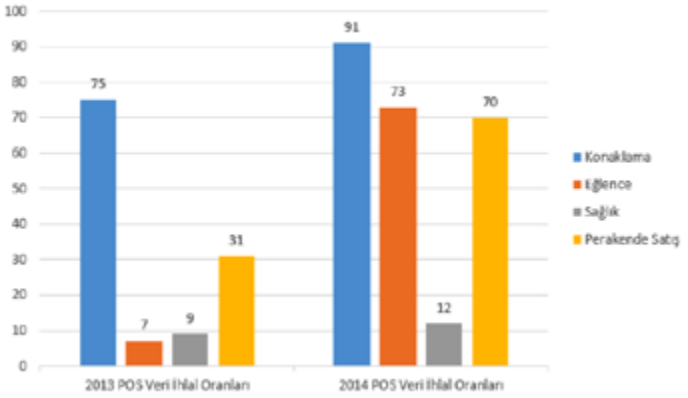
V. İSTATİSTİKSEL VERİLER İLE RAM KAZIMA SALDIRILARININ ETKİLERİNİN İNCELENMESİ

BlackPOS zararlı yazılımı kullanılarak gerçekleştirilen saldırılar sonrasında büyük firmalar hem maddi hem de prestij olarak ciddi zarara uğramışlardır. Bunun üzerine 2012 yılından itibaren PoS sistemlerin güvenliklerinin sağlanması, temel güvenlik politikaları içerisinde en önemli maddelerden biri haline gelmiştir ve söz konusu saldırıların neden olduğu risk değeri istenilen seviyelere düşürülmüştür. Bu düşük orana rağmen, Verizon 2015 veri ihlal raporuna göre, günümüz siber dünyasında gerçekleştirilen veri istismarı saldırıları içerisinde en büyük pay, 28.5% oranla PoS sunuculara yönelik saldırılardır. Elde edilen bu oran saldırganlar açısından kredi kartı verilerinin ne kadar değerli olduğunun önemini göstermenin yanı sıra, PoS sunuculara yönelik saldırıların neden olduğu kayıpları ispatlar niteliktedir. Şekil 5'te 2015 yılında gerçekleştirilen saldırı türleri ve oranları gösterilmektedir [11].



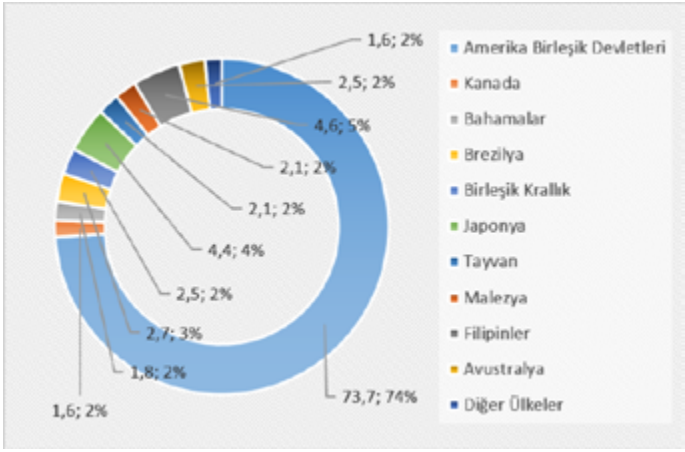
Şekil 5: 2015 yılı veri istismarı saldırı sınıflarının görülme oranları [11]

Kredi kartları günlük hayatta bireylerin alışveriş, konaklama, eğlence ve sağlık gibi birçok ihtiyacın karşılanmasında yardımcı olmaktadır. PoS sistemlerine yönelik gerçekleştirilen RAM kazıma saldırılarında saldırganlar çoğunlukla konaklama, eğlence, sağlık ve perakende satış gibi endüstriyel sektörlerde hizmet veren firmalar hedef alınmaktadır. Araştırma raporlarından elde edilen veriler incelendiğinde, 2013 yılında söz konusu alanlar içerisinde en fazla saldırıya konaklama sektöründe hizmet veren firmalar uğramıştır. 2014 yılında gerçekleştirilen saldırılarda ise yine en fazla konaklama sektörü hedef alınırken, perakende satış ve eğlence sektörlerinde hizmet veren firmalara yönelik saldırılarda büyük artış gözlemlenmiştir. Şekil 6' da 2013 ve 2014 yıllarında konaklama, eğlence, perakende satış ve sağlık sektörüne yönelik gerçekleştirilen RAM kazıma saldırılarının artış yüzdeleri gösterilmektedir [11].



Şekil 6: 2013 ve 2014 yıllarındaki endüstriyel alanlarda gözlemlenOne RAM kazıma saldırılarının artış yüzdesinin gösterimi [11]

Gerçekleştirilen saldırılarda çoğunlukla insan hayatı için zorunlu ihtiyaç veren sektörler hedef alındığından dolayı, birçok dünya ülkesi RAM kazıma saldırılarına maruz kalınmıştır. Şekil 7' de RAM kazıma saldırılarının ülkesel olarak dağılımı gösterilmektedir. Söz konusu şekilde görüldüğü gibi Amerika Birleşik Devletleri, gerçekleştirilen RAM kazıma saldırılarına maruz kalmada en yüksek paya sahiptir. Bu yüksek saldırı oranının en temel sebeplerinden biri, yüksek ekonomiden kaynaklanan pahalı malların alım satım oranının diğer ülkelere göre daha yüksek olmasıdır [12].



Şekil 7: PoS RAM kazıma saldırılarının ülkesel dağılımı [12]

Ülkemizde PoS sistemlere yönelik gerçekleştirilen RAM kazıma saldırılarından en önemlisinde, 2006 yılında GİMA Ticaret AŞ'ne 8 ile 22 Mayıs tarihleri arasında kredi kartı kullanarak alışveriş yapan müşterilerin kart bilgileri kopyalanmıştır [13]. Bankalar, bilgileri kopyalanmış olan kredi kartlarında anormal harcama değerleri tespit etmişlerdir. Gerçekleştirilen bu saldırı sonrasında Türkiye'deki kartlı ödeme sisteminde köklü bir çözüme gidilmiştir. Çipli kredi kartlarının üretilmesiyle, kartların sadece PoS cihazına okutulması zorunlu hale getirilmiştir. PoS cihazlarına bağlı bulunan kart pini giriş sistemi sayesinde bilgiler PoS üzerinde şifrelenip bankaya şifreli bir şekilde gönderilmeye başlanmıştır [14].

ABD ve bazı ülkelerde bankacılık ve ödeme sistemi ülkemizden farklıdır. Bu ülkelerde çipli kart sistemi yaygın olarak kullanılmadığı için PoS cihazı yerine bilgisayar üzerinde kullanılan PoS uygulaması kullanılmaktadır. Söz konusu uygulamada savunma düzeyi donanım katmanından uygulama katmanına geçtiği için saldırganlar tarafından daha rahat istismar edilmektedir. Visa Veri Güvenliği raporuna göre 2 Ekim 2008'de yaptığı uyarıda ABD'nin birçok bölgesinde geniş çaplı saldırıların gerçekleştirileceğini belirtmiştir. 2014 Ocak ayında ise gerçekleştirilen saldırılardan en çok etkilenen Target firmasının yaklaşık 70 milyon müşterisine ait kredi kartı verileri saldırganlar tarafından ele geçirilmiştir [15].

VI. POS SİSTEMLERE YÖNELİK RAM KAZIMA SALDIRILARINA KARŞI ALINABİLECEK ÖNLEMLER

Günümüzde kredi kartı işlemleri, kart verilerinin iletimi ve kart verilerinin güvenliği, Ödeme Kartı Endüstrisi (PCI) ve Veri Güvenlik Standartı (DSS) tarafından sağlanmaktadır. Belirlenen standartlar katmansal açıdan güvenliği pekiştirmek için oluşturulmuş standartlardır. Fakat günümüzde birçok şirket kaynak yetersizliğinden dolayı bu standartları doğru bir şekilde uygulayamamaktadır. Bunun sonucunda da söz konusu şirketler saldırganlar için açık bir hedef haline gelmektedir. Kurumlar RAM kazıma saldırılarına karşı etkili bir savunma sağlamak için ilk olarak, kurum ağı içerisinde yer alan tüm sistemler üzerindeki işletim sistemlerinin son güncel sürümlerine sahip olduklarından emin olunmalıdır. Kullanıcıların söz konusu güncelleştirme işlemlerini unutma ihtimaline karşı, söz konusu güncelleştirmeler otomatize bir şekilde gerçekleştirilmelidir. Ayrıca mümkün olduğunca

sistemler üzerinde korsan yazılım kullanımından kaçınılmalı ve lisanslı yazılımlar kullanılmalıdır. Kurumlar ayrıca dış ağdan, ya da sunuculara erişim yetkisi olmayan bir iç ağdan gelebilecek saldırıları engellemek amacıyla güvenlik duvarı kullanmalıdırlar. Kurumlar ayrıca iyi yapılandırılmış bir güvenlik duvarının yanında, beyaz liste mantığı ile çalışan ve sadece uzmanlar tarafından izin verilmiş uygulamaların sistemler üzerinde çalışmasını sağlayan güvenlik uygulamalarını tercih etmelidirler. Böylelikle şirket güvenlik uzmanı tarafından onaylanmış ve beyaz liste kapsamına alınmış uygulamalar dışında hiçbir uygulama ve yazılım, kritik sunucular üzerinde çalıştırılamayacaktır. Kurumlar ayrıca tam kapsamlı bir koruma sağlamak amacıyla, çalıştırılabilir olan exe, swf, pdf gibi uzantılara sahip tüm dosyaları, indirme sonrasında sistemler üzerinde çalıştırmadan, kendi bünyesinde sanal sistemler üzerinde çalıştıran ve analiz eden aktif ağ cihazlarını da tercih etmelidirler. Kart bilgilerinin geçici olarak da olsa tutulduğu sistemlere ait ağ trafiği kayıt altına alınmalı ve düzenli olarak takip edilmelidir. Şirket içinden gerçekleşebilecek saldırılara karşı önlem almak amacıyla PCI-DSS standartları kapsamında, sunuculara sadece güvenilir kişiler tarafından erişilmesi mümkün kılınmalıdır. Ayrıca sisteme erişim yetkisi olan tüm bireyler için olan bilgi güvenliği politikası oluşturulmalıdır [16].

Kredi kartı verilerinin elde edilmesi amacıyla gerçekleştirilen büyük saldırılar sonrasında, söz konusu bilgilerin güvenliğini sağlamak amacıyla Chip-Pin sistemi getirilmiştir. Fakat bu gelişme sonrasında, saldırılar söz konusu kartların zayıf halkası olan manyetik bant sistemi üzerinden gerçekleştirilmeye devam etmiştir. Bu durumun önüne geçmek için firmalar tarafından uygulanması gereken bazı önlemler bulunmaktadır. Chip-Pin sistemini desteklemeyen uygulamalar, dünya genelinde herhangi bir alış-veriş merkezi veya ATM sisteminde kullanılmamalıdır. Kullanılacak cihazlar Chip-Pin uygulamalarını ve getirdiği güvenlik önlemlerini desteklemelidir. Söz konusu kartlar ile ilgili bir sorun oluştuğunda, kart doğrulama işlemi için manyetik bant alanının kullanılmamasına özen gösterilmelidir. Chip-Pin sistemine sahip kartların, yan kanal analizi ve tersine mühendislik saldırılarına karşı güvenilir olduğunun bağımsız kuruluşlar tarafından sertifikalandırılması gerekmektedir. Türkiye'de kredi kartlarının ürün ve sistem güvenlik değerlendirmeleri için TS ISO/IEC 15408 standartı kullanılmaktadır. ISO/IEC 15408 sertifikası olmayan akıllı kartlar tercih edilmemelidir [17].

VII. SONUÇ

Araştırma raporlarından elde edilen veriler göstermektedir ki, günümüz siber dünyasında veri istismarı amacıyla gerçekleştirilen saldırılarda en büyük orana, PoS sistemlere yönelik RAM kazıma saldırıları sahiptir. Söz konusu saldırıların bu denli yaygın olmasının nedeni; konaklama, sağlık, perakende satış gibi insan hayatının temel ihtiyaçlarını karşılayan birçok sektörde kredi kartlarının kullanılıyor olmasıdır.

Bu çalışmada PoS sistemlere yönelik gerçekleştirilen RAM kazıma saldırılarının metodolojisi incelenmiştir. Söz konusu saldırıları daha iyi analiz edebilmek adına, kredi kartlarına ait verilerin tutulduğu veri grupları incelenmiştir. Ayrıca söz konusu veri gruplarının RAM' den elde edilebilmesi amacıyla saldırganlar tarafından arama ve doğrulama amaçlı kullanılan algoritmalara değinilmiştir. Çalışmada ayrıca, RAM kazıma saldırılarının başlangıç kaynaklarına değinilmiş, söz konusu saldırılarda kullanılan zararlı yazılımların tarihsel gelişimi incelenmiştir. Çalışma, çeşitli istatistiksel veriler ışığında, söz konusu saldırının etkinliğinin analiz edilmesi ile devam etmiş ve bünyesinde PoS sistemler barındıran firmaların alması gereken bir takım önlemlerin sıralanması ile tamamlanmıştır. Sonraki çalışmalarda temassız akıllı kartların geliştirilmesindeki süreç, güvenlik açıkları, bu kartlara yönelik saldırılar ve bunlara karşı alınabilecek önlemler ele alınacaktır.

KAYNAKLAR

[1] RAM Kazıyıcılar ve Diğer POS Zararlı Yazılımlar, Bağlantı: https://blog.kaspersky.com.tr/ram-kaziyicilar-ve-diger-satis-noktasi-zararli-yazilimlari/872/, Mayıs 2015

[2] Magnetic Stripe Card, Bağlantı: https://en.wikipedia.org/wiki/Magnetic_stripe_card, Haziran 2015

[3] Track Format of Magnetic Stripe Cards (Track 1 and 2), Bağlantı: http://www.acmetech.com/documentation/credit_cards/magstripe_track_format.html, Haziran 2015.

[4] Kredi Kartı Güvenlik Kodu: CVV2/CVC2/CID, Bağlantı: http://www.tuketicifinansman.net/2008/06/cvv2-cvc2-cid-guvenlik-kodu-kredi.html, Haziran 2015

[5] Validating Credit Card Numbers on Your Order Form, Bağlantı: http://www.regular-expressions.info/creditcard.html, Mayıs 2015

[6] Kredi Kartı Doğrulama – Luhn Algoritması, Bağlantı: http://www.yazilimdilleri.net/YazilimMakale-1830-Kredi-Karti-Dogrulama---Luhn-Algoritmasi.aspx, Mayıs 2015

[7] NitlovePOS Malware Uses Phishing Attacks TO Target POS Terminals, Bağlantı: http://www.bsminfo.com/doc/nitlovepos-malware-uses-phishing-attacks-to-target-pos-terminals-0001, Temmuz 2015

[8] POS RAM Scraper Malwares, http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-pos-ram-scraper-malware.pdf, Haziran 2015.

[9] CVE and Candidates as of 20150820, Bağlantı: https://cve.mitre.org/data/downloads/allitems.html, Temmuz 2015

[10] Point Of Sale (POS) Malware, Bağlantı: https://www.elevenpaths.com/wp-content/uploads/2015/06/TDS-PoS-Malware-Telefonica-2015-05.pdf, Temmuz 2015

[11] 2015 Data Breach Investigation Report, http://www.verizonenterprise.com/DBIR/2015/, Temmuz 2015

[12] A Look at Point of Sale RAM Scraper Malware and How It Works, Bağlantı: https://nakedsecurity.sophos.com/2013/07/16/a-look-at-point-of-sale-ram-scraper-malware-and-how-it-works/, Mayıs 2015

[13] Markette Korsan Çıktı Gima Yeni Sisteme Geçti, Bağlantı: http://www.hurriyet.com.tr/ekonomi/4494369_p.

asp, Mart 2015

[14] RAM Casusluğu, Bağlantı: https://www.mertsarica.com/ram-casuslugu/, Mayıs 2015

[15] In Home Depot Breach, Investigation Focuses on Self-Checkout Lanes, Bağlantı: http://krebsonsecurity.com/tag/target-data-breach/, Mart 2015

[16] PCI-DSS Nedir? , Bağlantı http://www.logsign.net/blog/index.php/pci-dss/ , Haziran 2015

[17] Manyetik Şeritli Kartlar ve CHIP&PIN Uygulaması, Bağlantı:https://www.bilgiguvenligi.gov.tr/donanim-guvenligi/manyetik-seritli-kartlar-ve-chip-pin-uygulamasi-3.html, Temmuz 2015

Doç. Dr. Ecir Uğur Küçüksille - 1976 yılında Isparta'da doğdu. Lisans eğitimini Gazi Üniversitesi Teknik Eğitim Fakültesi Bilgisayar Sistemleri Öğretmenliği Bölümü'nde tamamladı. Yüksek Lisans Eğitimini Süleyman Demirel Üniversitesi Makine Eğitimi Ana Bilim Dalında yaptı. Doktora Eğitimini Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü İşletme/Sayısal Yöntemler Ana Bilim Dalında tamamladı. Halen Süleyman Demirel Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü'nde öğretim üyesi olarak görev yapmaktadır. Bilgisayar, güvenlik ve yapay zeka alanlarında çalışmaları bulunmaktadır.

Bekir Eray Katı - 1992 yılında Karaman' da doğdu. Lisans eğitimini Süleyman Demirel Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümün'nde tamamladı. Halen Süleyman Demirel Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı' nda yüksek lisans eğitimine devam etmektedir. Araştırma konuları arasında, veri tabanı güvenliği ve sızma testleri yer almaktadır.

Mehmet Ali Yalçınkaya - 1990 yılında Isparta' da doğdu. Lisans eğitimini Süleyman Demirel Üniversitesi Teknik Eğitim Fakültesi Bilgisayar Sistemleri Öğretmenliği' nde tamamladı. Yüksek Lisans Eğitimini Süleyman Demirel Üniversitesi Bilgisayar Mühendisliği Ana Bilim Dalında yaptı. Halen Süleyman Demirel Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü' nde araştırma görevlisi olarak görev yapmakla birlikte, Süleyman Demirel Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı' nda doktora eğitimine devam etmektedir. Araştırma konuları arasında, bilgi güvenliği ve sızma testleri yer almaktadır.

SALDIRI TESPİT SISTEMİNİN BULUT BİLİŞİMDE KULLANIMI VE ETKİLERİ

Fatma Didem ÖĞRETMEN, Muhammed Ali AYDIN, Ahmet SERTBAŞ

Fatma Didem ÖĞRETMEN, Harran Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 63000, Şanlıurfa, Türkiye. (e-mail: fdidemogretmen@harran.edu.tr).
Muhammed Ali AYDIN, İstanbul Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 34320, İstanbul, Türkiye. (e-mail: aydinali@istanbul.edu.tr).
Ahmet SERTBAŞ, İstanbul Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 34320, İstanbul, Türkiye. (e-mail: asertbas@istanbul.edu.tr).
Bu çalışma aynı zamanda İstanbul Üniversitesi, Fen Bilimleri Enstitüsü'nde Güvenli Bulut Bilişim İçin Saldırı Tespit Sistemi Kullanımı adlı yüksek lisans tezinin bir parçasıdır.

Özet — Teknolojinin hızla gelişmesiyle, bu hıza ayak uydurabilmek için tamamen internet tabanlı olarak geliştirilen bulut bilişimin dünyada olduğu gibi ülkemizde de kullanımı giderek yaygınlaşmaktadır. Bulut bilişimin kullanıcılarına sunduğu başta ekonomik fayda olmak üzere birçok faydanın yanı sıra, bu yeni modelle yeni güvenlik sorunları da ortaya çıkmıştır. Bulut bilişimin maruz kaldığı güvenlik tehditlerine, açıklıklarına, yetersizliklerine karşı çeşitli güvenlik mekanizmaları geliştirilmesi üzerine çalışmalar yapılmaktadır. Bu çalışmada, bulut bilişimde sanallaştırma ve sanal makine güvenliğinin sağlanması için Sunucu-Tabanlı Saldırı Tespit Sistemi (Host Based Intrusion Detection System – HIDS) kullanılması ve böylelikle güvenli bulut bilişim sağlanması amaçlanmıştır.

Anahtar Kelimeler — Bulut bilişim, Güvenlik, Saldırı tespit sistemi, Sanallaştırma.

Abstract — With the rapid development of technology, this model that is developed entirely based on internet to keep up the development speed is increasingly being used in our country as well as over the world. Cloud computing provides number of benefits out of which economic benefit is main benefit. With the lots of benefits, this new model has emerged as new security problems. There are many studies on improving a variety of security mechanisms for security threats, vulnerabilities and insufficiencies in cloud computing. This study is aims to use host-based intrusion detection system (HIDS) for enabling security of virtualization and virtual machine in cloud computing and thus providing secure cloud computing.

Index Terms — Cloud computing, Security, Intrusion detection system (IDS), Virtualization.

I. GİRİŞ

Bilişim teknolojileri, bilgi çağı olarak nitelendirilen modern çağın getirdiği yenilikler ve kullanıcıların hızla değişen ihtiyaçları nedeniyle sürekli bir değişim ve gelişim göstermektedir. Kullanıcıların ofis bağımlılığı olmadan çalışabilme olanaklarının artması, dinamik yapıdaki ofislerin yaygınlaşması, daha az kaynak ile daha fazla hizmet sunma gerekliliğinin ortaya çıkması ve ekonomik nedenlerden dolayı kurumlarda kısıtlı miktardaki kaynakların daha etkili kullanımı önem kazanmış ve tüm iş gruplarının kendi içlerinde yeniden yapılanması gerekliliği ortaya çıkmıştır. Bilişim sektöründe ortaya çıkan bu ihtiyaçlara yönelik olarak fiziksel sistemlerinin sanal ortamlara taşınması ve bir fiziksel sistemin üzerinde birçok sanal sistem kullandırılması çözümleri geliştirilmiştir. Zamanla kullanıcıların uygulamalarını mekân, zaman ve platformdan bağımsız olarak kullanabilme yönündeki artan talepleri doğrultusunda "bulut bilişim (cloud computing)" kavramı ortaya çıkmış, sanallaştırmanın gelişmesi ve bilgi teknolojilerinde hızla yer edinmesinden sonra, yeni bir bilişim teknolojisi olarak sanallaştırma alt yapısının üzerine yapılandırılmıştır.

Bulut bilişim, kullanıcılarına ekonomik faydası başta olmak üzere birçok fayda sağlamaktadır. Buna karşın dağıtık servise dayalı mimarisi, çoklu-kullanıcılar ve çoklu-domain altyapıları nedeniyle, tehditlere ve savunmasızlıklara karşı dayanıksız olarak görülmektedir. Bulut bilişimde güvenlik sorunu kritik bir sorundur. Bu sorun nedeniyle kullanıcılar bulutları kullanmaya tereddüt etmektedirler [1]. Güvenlik sorunları, servisleri barındıran bulut sağlayıcılarını daha fazla ilgilendirmektedir. Antivirüs yazılımları, güvenlik duvarları, bekçi sistemleri ve özellikle saldırı tespit sistemleri gibi mekanizmaları kullanılarak güvenlik sorunlarının üstesinden gelinmeye çalışılmaktadır. Güvenlik mekanizmaları ele alınırken bulut bilişimin doğası gereği kaynak kullanım verimliliği konusu göz önünde tutulmaktadır.

II. BULUT BİLİŞİM

Bulut bilişim, kullanıcılara veriye daha az maliyetle ve daha hızlı bir şekilde ulaşma imkânı sağlayan, veri ve uygulamaları muhafaza etmek, işlemek ve kullanmak için internet ve merkezi uzak sunucuları kullanan servis tabanlı bir teknolojidir. Yeni nesil bilişim konularından olan bulut bilişimin literatürde birçok tanımı olmasıyla birlikte, ilgili kaynaklarda sıklıkla atıf yapılan ve en çok benimsen Amerika Birleşik Devletleri Ulusal Standartlar ve Teknoloji Enstitüsü (National Institute of Standards and Technology - NIST) [2] tarafından yapılan tanıma göre bulut bilişim, yapılandırılabilir bilişim kaynaklarından (bilgisayar ağları, sunucular, veri depolama, uygulamalar ve servisler vb.) oluşan ortak bir havuza, uygun koşullarda ve isteğe bağlı olarak her zaman, her yerden erişime imkân veren bir modeldir. Söz konusu kaynaklar asgari düzeyde yönetimsel çaba ve servis alıcı-servis sağlayıcı etkileşimi gerektirecek kolaylıkta tedarik edilebilmekte ve elden çıkarılabilmektedir [2].

III. SANALLAŞTIRMA

Sanallaştırma, bulut bilişimin gelişiminde önemli bir teknoloji olanağıdır. Sanallaştırma, donanım ve işletim sistemi arasında yer alan ve üzerinde uygulamaların çalıştırıldığı bir yazılım soyutlama katmanıdır. Genel anlamda bilgisayar kaynaklarının kullanıcılardan soyutlanması olarak tanımlanabilir. Soyutlamanın gerçekleştirilmesi kaynakların paylaşılması veya birleştirilmesiyle yapılmaktadır.

Sanallaştırma, 1960'lı yıllarda IBM şirketinin anaçatı (mainframe) sistemlerinde "Zaman Paylaşımı" fikrini ortaya çıkardığı ve büyük bir anaçatı bilgisayarı birkaç mantıksal örneğine ayırması amacıyla geliştirdiği günden bu yana bilişim dünyasında yer almaktadır [3]. İlk çıktığı günden bu yana sanallaştırma kavramı önemli ölçüde olgunlaşmış ve hafıza, depolama, işlemciler, yazılım, ağ servisleri gibi bilişim teknolojilerinin tüm yönlerine uygulanmıştır.

Sanallaştırma; fiziksel sınırlamaların ortadan kalkmasının sağlanması, tek bir merkezden birden çok sunucunun yönetilebilmesi ile yönetim yükünün en aza indirgenmesi, alt yapı maliyetlerinin büyük ölçüde azaltılması, fiziksel sunuculara oranla yeni sunucuların kullanıma alınması işleminin oldukça kısa zaman alması, aynı makine üzerinde birbirinden farklı birden fazla işletim sisteminin yürütülebilmesi gibi birçok fayda sağlamaktadır. Bu sayede bilişim teknolojilerinde oldukça yaygın kullanım alanına sahip bir teknolojidir.

Bir sanallaştırma ortamında aşağıda yer alan bileşenler bulunmaktadır:

Hipervizör (Hypervisor): Sanallaştırmayı sağlayan yazılım katmanıdır. Sanal Makine Denetleyicisi (Virtual Machine Monitor – VMM) olarak da bilinen hipervizör, misafir sanal makinelerin üzerinde işletileceği sanal ortamın oluşturulmasından sorumludur. Misafir sistemleri denetler ve kaynakların misafir sanal makinelere gerektiği şekilde tahsis edilmesini sağlar.

Misafir (Guest) veya Sanal Makine (Virtual Machine – VM): Hipervizörün üstünde sanallaştırılan uygulama veya işletim sistemidir. Fiziksel makinenin sanallaştırılmış bir temsilidir. Her bir sanal makine (VM) işlemci, hafıza, ağ bağdaştırıcısı, çıkarılabilir aygıtlar ve çevresel aygıtları taklit ederek ayrı bir bilgisayar gibi davranan kendine yeten bir operasyon ortamıdır. Aynı fiziksel makinede farklı işletim sistemli birkaç VM eş zamanlı olarak işletilebilmektedir, fakat her bir misafir işletim sistemi için hipervizör tarafından tek bir donanım sunumu vardır.

IV. BULUT BİLİŞİMDE GÜVENLİK

Dağıtık yapısı nedeniyle bulut bilişim ortamları olası güvenlik açıklarını arayan saldırganlar/davetsiz misafirler için bir hedeftir. Çoğu çalışmalar göstermiştir ki, istemcilerin veri mahremiyetini, gizliliğini garantilenmesi için bulut bilişim sağlayıcılarına güvenmek zor bir konudur [4]. Bununla birlikte bulut sağlayıcısı veya yöneticisinin de her zaman güvenilir olacağının garantisi yoktur.

Bulut bilişimde kritik konu olan güvenlik konusunda karşılaşılan sorunlara genel olarak iki açıdan yaklaşılmaktadır. Birincisi, bulut servis sağlayıcısının sağladığı servislerin güvenli olduğunu garanti edebilmesi ve kullanıcının kimlik yönetimini başarabilmesi; ikincisi ise, kullanıcının kullandığı servislerin yeteri kadar güvenli olduğundan emin olabilmesidir.

Bulut bilişimde karşılaşılan güvenlik riskleri veri gizliliği ve mahremiyetinin korunması, yönetim yetersizliği, yönetim arayüzündeki olası güvenlik açığı, bulut çalışanlarının kötü niyetli davranışları, kullanılabilirliğin garantilenememesi, izolasyon başarısızlığı, uyum ve yasal riskler olarak belirtilmektedir [5].

A. Bulut Bilişimde Güvenlik Mekanizmaları

İnternet kullanımının yaygınlaşması ile birlikte bilişim sistemlerine karşı güvenlik tehditlerinde önemli artışlar ve saldırı türlerinde genişlemeler olmuştur. Karşılaşılan tehditler ve saldırılar sebebiyle yeni mekanizmaların geliştirilmesi zorunluluğu ortaya çıkmıştır. Bilişim sistemlerinde güvenliğin sağlanması amacıyla güvenlik duvarları (firewall), güvenlik açığı tarayıcıları (vulnerability scanner) ve saldırı tespit sistemleri kullanılmaktadır. Bu güvenlik mekanizmalarının hiçbirinin tek başına kullanılması güvenlik açısından tam olarak yeterli görülmemektedir; çünkü her biri farklı açılardan güvenlik konularına odaklanmıştır. Sistemde güvenliğin sağlanması, bu mekanizmaların birbirini destekleyecek şekilde beraber kullanılmasını gerektirmektedir.

Bulut bilişim sistemleri de güvenliğin sağlanması amacıyla çeşitli yönetim modellerine odaklanmıştır [6]. Bu modellerde çoğunlukla Saldırı Tespit Sistemi, Güvenli Bilişim, veri şifreleme gibi mekanizmalar yer almaktadır.

Saldırı Tespit Sistemi (Intrusion Detection System – IDS)

Bir kaynağın veya verinin güvenilirliğini, bütünlüğünü, gizliliğini veya erişilebilirliğini engellemeye yönelik tüm eylemler saldırı (intrusion) olarak tanımlanmaktadır. Saldırı tespit sistemi (Intrusion Detection System – IDS), bir bilgisayar sistemi veya bilgisayar ağında meydana gelen olayların izlenmesini otomatik hale getiren, bu sistemlerde oluşan kötü niyetli faaliyetlerin ve bilgisayar güvenlik politikaları, kabul edilebilir kullanım politikaları veya standart güvenlik politikaları ihlallerinin analiz edilmesini ve yönetim birimine raporlanmasını sağlayan yazılım veya donanım sistemidir [7]. IDS, bir tür alarm sistemi olarak düşünülebilir. Saldırıların tespit edilebilmesi tetikleme mekanizmaları ile gerçekleştirilir. Bir IDS, birkaç bileşenden oluşmaktadır [8]:

Algılayıcılar (Ajanlar): Güvenlik olaylarını oluşturur.
Monitör: Olayları ve uyarıları izlemek ve algılayıcıları kontrol etmek için kullanılır.
Merkezi Motor: Algılayıcılar tarafından günlüğe kaydedilen kayıtları bir veri tabanında tutar ve bir güvenlik olayı alındığında uyarıları oluşturmak için bir kurallar sistemini kullanır.

Saldırı tespit sistemleri genel olarak Sunucu-Tabanlı IDS (Host-Based IDS)'ler ve Ağ-Tabanlı IDS (Network-Based IDS)'ler olmak üzere iki sınıfa ayrılmaktadır.

Sunucu-Tabanlı Saldırı Tespit Sistemi (Host-Based IDS - HIDS)

Sunucu-Tabanlı saldırı tespit sistemi (host-based IDS – HIDS), hedef sistemin bireysel bilgisayarların [8] olduğu, tasarlanan ilk saldırı tespit yazılım türüdür. HIDS, sadece bilgisayar sisteminden gelen ve giden paketleri izler ve şüpheli etkinlik 1 olarak tespit edilirse kullanıcı veya yönetici uyarır. Özellikle önemli sunucu sistemler üzerinde gizli ve kritik bilgileri korumak amacıyla kullanılmaktadır. HIDS'ler, belirli bir makinede meydana gelebilecek saldırıları önlemek üzere sunuculara ya da çalışma istasyonlarına yerleştirilmiş algılayıcılardan (ajanlardan) oluşmaktadır. HIDS sistem durumu izlemek için işletim sistemi denetim rotalarından ve sistem log kayıtlarından faydalanarak karar verebilmektedir. Hangi kaynakların hangi programlara eriştiğini algılayabilmektedir.

HIDS, kullanıcıya özel olaylar; zararlı bir kodun çalıştırılması ve bellek taşması gibi kod analizlerini, bütünlük ve erişimi içerecek şekilde dosya sisteminin izlenmesini, kullanıcı kayıtları gözden geçirilirken meydana gelen kayıt analizlerini ve son olarak ağ ayarı yapılandırmalarındaki değişiklikleri izlemekte ve tespit etmektedir.

Ağ-Tabanlı Saldırı Tespit Sistemi (Network-Based IDS – NIDS)

Ağ tabanlı saldırı tespit sistemleri (Network-Based IDS – NIDS), belirli bir sunucudan ziyade ağın kendisine odaklanmaktadır. Ağ üzerinden geçen trafiği veri kaynağı olarak görüntülemektedir. NIDS, ağ segmenti veya anahtarlama cihazını dinleyerek, bu ağ segmentine bağlı birden çok hostu etkileyen ağ trafiğini izleyebilmektedir.

NIDS'lerde temelde ağda dolaşan paketlerden, ağda belirli noktalarda yer alan algılayıcılar üzerinden geçen paketleriyle ilgilenilmektedir. Algılayıcıya gelen paket sistemdeki mevcut imzalarla karşılaştırılarak paketin analizi yapılır. Başlangıç düzeyindeki filtre hangi paketlerin kabul edilip hangilerinin atılacağını veya paketin saldırı tanıma modülüne gönderilip gönderilmeyeceğini belirlemektedir. Saldırı tespit edilirse cevap modülü saldırıya karşılık olarak alarm üretim mekanizmasını tetiklemektedir.

Güvenilir Bilişim (Trusted Computing)

Güvenilir bilişim (Trusted Computing – TC), Trusted Computing Group (TCG) tarafından geliştirilen ve desteklenen bir metodolojidir [9]. Güvenilir bilişim, donanım iyileştirmeleri ve buna bağlı yazılım değişiklikleri yoluyla bilgisayar güvenlik sorunlarını çözmek için teknoloji ve öneriler sunan geniş bir kavramdır. Bilişim sistemlerini oluşturan bileşenlerin her biri arasında gizlilik, bütünlük, erişebilirlik ve kurtarılabilirlik gereksinime dayanan bir "güven ilişkisi" planlar. Birçok büyük donanım üreticisi ve yazılım sağlayıcı firmalar TCG ile işbirliği yapmaktadır [9].

Yetkisiz değişikliklere ve tehditlere karşı bilgisayar kaynaklarını korumak için TC yaklaşımının ana parçası olan Trusted Platform Module (TPM) kullanılır. TPM, anakart üzerindeki tümleşik bir devredir ve sistem üzerinde çalışan yazılım tarafından iyi tanımlı olan veya olmayan komutların ve örneklerin bütünlüğünü kontrol eder [10]. TPM, temelinde şifreleme anahtarlarından yararlanılarak oluşturulan, güvenlikle ilgili temel işlemleri sağlamak amacıyla tasarlanmıştır. TPM'de yer alan Platform Configuration Registers (PCRs)'daki işletim durum platformu hakkında bilgi depolar. Geçerli platformdan gelen platform isteklerini doğrular. TC uzaktan doğrulama (remote attestation), mühürlü depolama (sealed storage), güvenilir önyükleme (trusted boot) gibi teknolojileri içermektedir [10].

V. GÜVENLİ BULUT BİLİŞİM İÇİN SALDIRI TESPİT SİSTEMİ KULLANIM ÖRNEĞİ

Bulut bilişimde sistem güvenliğini arttırmak için en popüler yöntem sistemin sürekli izlenmesidir. Bunun için IDS kullanılması tercih edilen, iyi bir yöntemdir. IDS, bir sistemin hesaplama ve ağ kaynaklarını hedef alan zararlı faaliyetleri tanımlamaya çalışır. Literatürde bulut bilişim için birçok IDS mekanizması tanımlanmıştır.

Bu çalışmada güvenli bulut bilişim için farklı IDS kullanım modellerinin oluşturulması amacıyla bulut bilişimde kullanılabilecek IDS tabanlı bir hibrid yaklaşım sunan AdjointVM [11] Koruma Modeli ile AdjointVM yaklaşımıyla sunulan güvenlik mekanizmasının zayıf yönlerinin üstesinden gelmek amacıyla birtakım iyileştirmeler ekleyerek eksikliklerinin giderilmesini hedefleyen U. Oktay ve Ark. [12]'nın öneri olarak literatüre sunduğu İyileştirilmiş AdjointVM Koruma Modeli, IDS kullanım mekanizmaları açısından ele alınarak bulut bilişimde IDS kullanımının nasıl olması gerektiğinin ortaya konulması için test ortamında gerçekleştirilmiş ve bulgular karşılaştırılmıştır.

A. İlgili Çalışmalar

AdjointVM Koruma Modeli

J.Kong [11], AdjointVM adında güvenilir olmayan bulut sağlayıcılarına yönelik önemli bir çalışma olarak görülen bir IDS mekanizması geliştirmiştir. Bu mekanizma, güvensiz sağlayıcılardan gelebilecek iç saldırılara karşı ve gerçek veya sanal ağ üzerinden gelebilecek dış saldırılara karşı IDS tabanlı VMM barındırmaktadır. Her iki saldırı türüne karşı koruma sağlayan hibrid bir mimari geliştirilmiştir.

AdjointVM Koruma Modeli'nde sanallaştırma ortamı olarak açık kaynak kodlu bir platform olan Xen Hypervisor [13] seçilmiştir.

Xen hipervizörü, sanallaştırma katmanının ince ve minimal olması düşüncesine dayanarak geliştirilmiştir. Bu tasarım fikrine dayanarak, asıl sanallaştırma işi hipervizörün bir seviye üstüne devredilmesi gerekmektedir. Bu nedenle sıradan VM'lerden daha fazla ayrıcalıklara sahip Domain 0 (Dom0) olarak adlandırılan özel bir VM yer almaktadır. Dom0, Xen hipervizör üzerinde çalışan ve yönetim araçlarını üzerinde bulunduran güvenli VM'dir. Dom0 hipervizörün boot aşamasında otomatik olarak çalışmaya başlamakta ve fiziksel donanıma doğrudan erişerek diskler, ağ bağdaştırıcısı gibi aygıtlarla iletişim kurmakta, diğer VM'ler için sanal aygıtların yönetilmesinden, karmaşık işlemlerin gerçekleştirilmesinden sorumlu olmaktadır. Xen terminolojisinde tüm VM'lere domain denilmektedir. Yönetimden sorumlu VM Dom0 iken, diğer tüm VM'ler de DomUs olarak adlandırılmaktadır.

AdjointVM modelinde her bir VM, AdjointVM denilen başka bir VM tarafından izlenmekte ve korunmaktadır. AdjointVM'in kurulması ile Xen, bir VM'nin adres alanının diğer bir VM tarafından korunması için eşleştirilmesine yardımcı olur. Kullanıcı kesme noktalarını belirleyebilir, daha sonra bir olay izleme ve günlükleme daemon'u ile korunmuş VM'nin bellek alanını takip eder. Bir saldırı belirlendiğinde daemon, bunu kullanıcıya birkaç yolla raporlar.

VM'in güvenliğinin sağlanması amacıyla bu modelde hibrid bir IDS yapısı oluşturulmuş olup, sunucu tabanlı bir IDS olan Opearting System Security (OSSEC) HIDS [14] ile işletim sistemi çekirdeğini izleyen Kernel Monitor Daemon (KMD) [11] kullanılmıştır. OSSEC sunucusu AdjointVM üzerine kurulur, ajanı ise korunmuş VM üzerindedir. Saldırılara ait imzalar OSSEC sunucusu üzerinde bulunmakta ve ajanlar saldırı bilgilerini sunucu bünyesindeki saldırı imza veri tabanından almaktadır.

İyileştirilmiş AdjointVM Koruma Modeli

AdjointVM, bulut ortamı için iyi bir model olarak sunulmasına karşın bazı güvenlik sorunları öngörülmektedir. AdjointVM'yi hedef alan herhangi bir saldırı olursa ve servislerini manipule ederse sisteme doğru pozitif oranlı uyarı verememektedir, bu nedenle korunan VM, iç ve dış saldırılara karşı savunmasız hale gelir. AdjointVM'de, bir VM çiftinde iki tür VM mevcuttur: koruyan VM ve korunan VM. Koruyan VM'nin görevi, korunan VM'nin güvenliğini sağlamaktır, ancak koruyan VM saldırılara karşı savunmasızdır.

AdjointVM modelindeki bu eksikliklere çözüm olarak U. Oktay ve Ark., AdjointVM Modeli'nin İyileştirilmesi [12] yaklaşımına göre yeni bir model önermişlerdir. Önerilen modelde, AdjointVM modelinde koruyan VM'nin güvenliği nasıl sağlanacak sorusuna cevap bulunmaktadır. Bu modelde, AdjointVM çiftindeki her iki VM de aynı anda hem koruyan hem de korunandır. VM'ler hem haritalama, hem de her birinin çekirdeğini herhangi bir rootkit'e karşı izleme yeteneğine sahiptir. Ayrıca modelde iki OSSEC sunucusu ve ajanı yer alır, birer sunucu ve ajan korunan VM'de kurulur, diğerleri de koruyan VM'de kurulur, böylece her iki VM de bir diğerini korur. Koruyan VM'ye herhangi bir saldırı olursa Korunan VM bunu engelleyebilir.

Geliştirilen sistem web tarama, sshd brute force, ftp scan, çoklu spam saldırıları, SQL enjeksiyonu gibi web saldırıları için ve knork ya da vlogger gibi çekirdek rootkitleri için dayanıklıdır [15].

B. Gerçekleştirme Ortamı ve Kullanılan Bileşenler

Gerçekleştirme ortamında, referans model olarak ele alınan J. Kong'un yaptığı AdjointVM Koruma Modeli uygulamasının [11] gerçekleştirme ortamına uygunluk açısından Intel Virtualization Technology (Intel®VT) [16] ve hyper-therading özelliğine sahip Intel Core i7-2630QM 2.00 GHz CPU ve 4 GB DDR3 rastgele erişimli bellek (Random Access Memory – RAM) kapasiteli bir bilgisayar üzerinde sistemler inşa edilmiştir. Sistemde bahsi geçen referans modele uygunluk açısından hem ana makine hem de oluşturulan sanal makineler üzerinde 64 bit mimariye sahip Fedora 20 (Linux 3.19.5-100.fc20.x86_64) işletim sisteminin kullanılması tercih edilmiştir. Sanallaştırma platformu olarak referans modeldeki Xen Hypervisor seçilmiştir. Saldırı tespiti için hem referans modelinde belirtildiği üzere hem de yapılan literatür araştırmasında da etkinlik ve performans açısından tercih edilen açık kaynak kodlu sunucu tabanlı saldırı tespit sistemi (HIDS) olan OSSEC HIDS v2.8 kullanılmıştır.

C. Uygulama

Bulut bilişim sistemlerinde sistemin içerisinden, yani sanal altyapıda kullanılan hipervizör katmanından, bulut sistem yöneticilerinin veya operatörlerinin yetkisiz erişimlerine karşı bulut kullanıcıları tarafından alınabilecek güvenlik mekanizmalarının eksikliğin giderilebilmesi adına yapılan bu çalışmada, sistem üzerindeki istenmeyen ve zararlı aktivitelere karşı HIDS'lerin nasıl konumlandırılması ve yapılandırılması gerektiği ile ilgili iki güvenlik mekanizması mimarisi oluşturulmuştur. Bunlar;

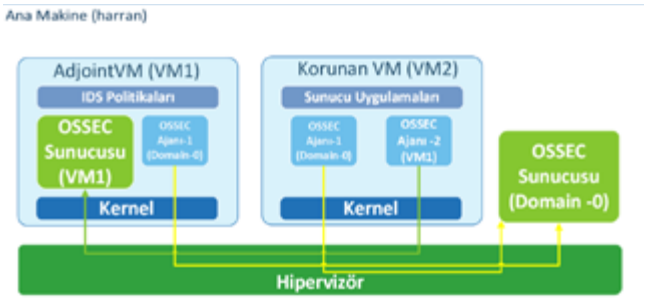
1. AdjointVM Koruma Mekanizması,
2. İyileştirilmiş AdjointVM Koruma Mekanizması.

Birinci mekanizmada, sistem içerisinde yer alan VM'lerin güvenliğinin sağlanması için her bir VM'yi koruyacak AdjointVM olarak adlandırılan eşlenik bir VM oluşturulmuştur. Oluşturulan bu AdjointVM, korunan VM'nin güvenliğinden sorumludur. OSSEC'in sunucu istemci mimarisinden dolayı OSSEC sunucusu eşlenik olarak ortama eklenen VM üzerinde, istemci durumundaki OSSEC ajanları da asıl güvenliğinin sağlanması amaçlanan VM üzerinde teşkil edilmiştir. Saldırılara ait imzalar OSSEC sunucusu üzerinde bulunmakta ve ajanlar saldırı bilgilerini sunucu bünyesindeki saldırı imzaları veri tabanından almaktadır. Bu nedenle HIDS olarak OSSEC sunucusu AdjointVM üzerine kurulmuş, bu sunucuya ait OSSEC ajanı (agent) ise korunan VM üzerine kurulmuştur. Bu iki VM bir koruma çifti olarak sistemde yer almaktadır.

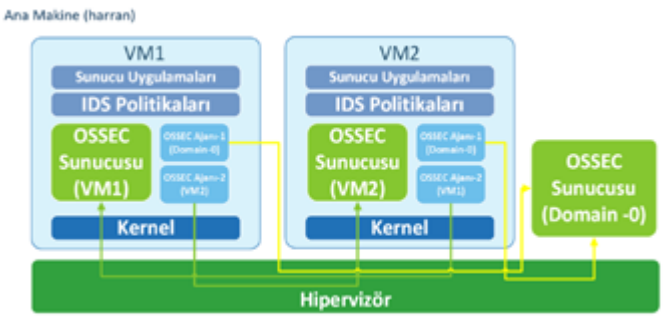
İkinci mekanizmada ise koruma çiftlerindeki her bir VM hem OSSEC sunucusu hem de OSSEC ajanı barındırmakta, karşılıklı olarak birbirlerinin güvenliğinin sorumluluğunu almaktadırlar. AdjointVM korunan VM'yi izlerken, korunan VM üzerindeki OSSEC sunucusu da AdjointVM'yi izlemektedir. Bu şekilde sanallaştırma ortamında tüm VM'ler çiftler halinde yer almaktadır.

Gerçekleştirme ortamında ana makine "harran" olarak adlandırılmaktadır. AjdointVM Koruma Modelinin saldırı tespit mekanizmasının uygulanması için J. Kong [11]'un çalışmasında belirttiği üzere sistemde en az iki sanal makineye ihtiyaç vardır. Oluşturulan gerçekleştirme ortamında korunan VM olarak "fed20-2" isimli VM, bu sanal makineyi koruyacak olan eşlenik VM yani AdjointVM ise "fed20" isimli VM yapılandırılmıştır.

Şekil 2'de AdjointVM Koruma Mekanizmasına ait mimari, Şekil 3'de ise İyileştirilmiş AdjointVM Koruma Mekanizmasına ait mimari gösterilmektedir.



Şekil 2. AdjointVM Koruma Mekanizması mimarisi



Şekil 3. İyileştirilmiş AdjointVM Koruma Mekanizması mimarisi

D. Bulgular

Bu çalışmada oluşturulan tüm güvenlik mekanizmalarının test işlemlerinde bulut bilişimde çok önemli bir kriter olan kaynak kullanım miktarları açısından elde edilen bulguların değerlendirilmesi ve böylelikle sistemlerin verimliğin ölçülmesi hedeflenmiştir.

Hem her bir güvenlik mekanizmasının çalıştırılması durumunda, hem de sistemde hiçbir güvenlik mekanizmasının yer almadığı durumda donanım kaynakları olarak sistemdeki işlemci (CPU) ve hafıza (RAM) kullanım oranları elde edilmiş ve birbiriyle kıyaslanmıştır. Böylelikle hangi güvenlik mekanizmasının tercih edileceği hakkında bulut kullanıcılarına yol göstermesi amaçlanmıştır.

Güvenlik mekanizmalarının denenmesi sırasında ana makine ve VM'lerin kullandığı CPU kaynaklarının kullanım miktarlarının ölçülebilmesi için Xen sanallaştırma platformunun sunduğu "xentop" isimli sanallaştırma ortam monitör araç yazılımı kullanılmıştır. xentop, bir Xen sistemi hakkında gerçek zamanlı bilgi vermektedir. Bu sayede sanallaştırma platformunda yer alan tüm çalışan VM'ler ve sistem ile ilgili anlık bilgilerin izlenebilmesini sağlamaktadır. Ana makine ve VM'lerin tükettiği hafıza yani RAM kaynaklarının kullanım miktarlarının ölçülebilmesi için Fedora OS ile varsayılan olarak sunulan "top" isimli sistem araç yazılımı kullanılmıştır. top, sistem üzerinde çalışmakta olan işlemlerin gerçek zamanlı listesini görüntüler. Aynı zamanda sistem çalışma süresi, anlık olarak mevcut CPU ve hafıza kullanımı, ya da çalışan işlemlerin toplam sayısı hakkında ek bilgileri görüntüler ve kullanıcıların işlem listesi üzerinde sıralama ya da çalışan bir işlemi öldürme gibi eylemleri gerçekleştirmesine imkân verir. Ana makinede top yazılımının çalıştırılmasıyla mevcut donanıma ait hafıza kaynağının ne kadarını kullandığı ölçülürken her bir VM üzerinde top yazılımı ayrı ayrı çalıştırılarak VM'nin kendisine tahsis edilen hafıza bölümünün ne kadarını kullandığı tespit edilmiştir.

xentop ve top yazılımında varsayılan olarak 3 saniye aralıklarla donanım kullanım miktarları gösterilmektedir. Sistemlerden elde edilen değerlerin daha güvenilir ve doğruluğunun daha yüksek olabilmesi amacıyla farklı zaman dilimlerinde elde edilen verilerin çeşitliğini arttırmak için her iki yazılımı –d opsiyonuyla çalıştırarak donanım kullanım durumu güncelleme sıklığı 1 saniye ve 10 saniye olarak değiştirilmiş olup, her bir zaman aralığı için 50 iterasyon yapılarak veriler elde edilmiştir. Bu işlem 1, 3 ve 10 saniye zaman dilimleri için birden fazla kez tekrarlanmıştır. Bu şekilde her bir durum için ayrı ayrı 1, 3 ve 10 saniye aralıklarla ölçümler yapılarak sistem üzerinden CPU ve RAM kullanım miktarları yüzde olarak alınmış; hem bu farklı frekans aralıklarındaki ölçüm değerlerinin ayrı ayrı ortalaması hem de tüm değerlerin genel ortalaması hesaplanmış olup, ilk üç durum için elde edilen tüm sonuçlar detaylı olarak Tablo 1, Tablo 2, Tablo 3 ve Tablo 4'te gösterilmiştir.

Güvenlik Mekanizmaları	CPU Kullanımı (%)								
	1 sn aralık			3 sn aralık			10 sn aralık		
	Domain-0	fed20	fed20-2	Domain-0	fed20	fed20-2	Domain-0	fed20	fed20-2
OSSEC HIDS Kullanılmazken	22,816	5,330	0,346	22,218	1,956	0,336	22,418	8,714	0,372
AdjointVM Koruma Mekanizması	22,736	9,792	7,978	**22,597**	8,968	7,734	**22,922**	9,768	7,932
İyileştirilmiş AdjointVM Koruma Mekanizması	**24,213**	**10,751**	**10,669**	22,393	**11,658**	**12,574**	22,601	**12,647**	**12,381**

Tablo 1. - Farklı frekanslarda cpu kullanımlarına göre güvenlik mekanizmalarının kıyaslanması

Güvenlik Mekanizmaları	RAM Kullanımı (%)								
	1 sn aralık			3 sn aralık			10 sn aralık		
	Domain-0	fed20	fed20-2	Domain-0	fed20	fed20-2	Domain-0	fed20	fed20-2
OSSEC HIDS Kullanılmazken	54,578	59,975	39,669	54,946	57,106	39,358	54,910	57,720	39,436
AdjointVM Koruma Mekanizması	56,114	**73,125**	41,117	55,977	**73,095**	40,984	56,127	**72,696**	40,606
İyileştirilmiş AdjointVM Koruma Mekanizması	**58,996**	62,571	**44,850**	**59,234**	63,596	**44,183**	**59,924**	63,805	**44,215**

Tablo 2. Farklı frekanslarda ram kullanımlarına göre güvenlik mekanizmalarının kıyaslanması

Güvenlik Mekanizmaları	CPU Kullanımı (%)		
	Domain-0	fed20	fed20-2
OSSEC HIDS Kullanılmazken	22,484	5,333	0,351
AdjointVM Koruma Mekanizması	22,752	9,509	7,881
İyileştirilmiş AdjointVM Koruma Mekanizması	**23,069**	**11,685**	**11,875**

Tablo 3. Ortalama cpu kullanımlarına göre güvenlik mekanizmalarının kıyaslanması

Güvenlik Mekanizmaları	RAM Kullanımı (%)		
	Domain-0	fed20	fed20-2
OSSEC HIDS Kullanılmazken	54,811	58,267	39,487
AdjointVM Koruma Mekanizması	56,073	**72,972**	40,902
İyileştirilmiş AdjointVM Koruma Mekanizması	**59,385**	63,324	**44,416**

Tablo 4. Ortalama ram kullanımlarına göre güvenlik mekanizmalarının kıyaslanması

E. Elde Edilen Bulguların Karşılaştırılması

AdjointVM Koruma Mekanizması ve İyileştirilmiş AdjointVM Koruma Mekanizmasından elde edilen bulgular göstermiştir ki genel olarak IDS kullanımı sistemdeki kaynak tüketim miktarını arttırmaktadır. İyileştirilmiş AdjointVM Koruma mekanizması, AdjointVM Koruma Mekanizmasına göre VM'lere daha fazla iş yaptırmış, bu nedenle CPU ve RAM kullanım miktarları kıyaslandığında daha fazla kaynak tüketimine neden olmuştur.

VI. SONUÇ

Bu çalışmada güvenli bulut bilişim için sunucu-tabanlı saldırı tespit sisteminin (HIDS) nasıl kullanılması gerektiği konusunda araştırmalar yapılmış, literatürde yer alan farklı güvenlik modellerine göre saldırı tespit mekanizmaları oluşturularak özellikle donanım kullanım maliyetleri bakımından karşılaştırılması yapılmış, böylelikle kullanıcılara en optimum güvenlik mekanizmasının tercih edilmesi konusunda fayda sağlamak amaçlanmıştır.

Öncelikle sistemde hiçbir güvenlik mekanizması kullanılmadığı zaman donanım kullanım miktarları tespit edilmiştir. İkinci durumda güvenlik modeli olarak literatürde yer alan AdjointVM yaklaşımına ait sunucu-tabanlı saldırı tespit sistemi kullanım mimarisi yönünden model ele alınarak AdjointVM Koruma Mekanizması gerçekleştirilmiştir. Oluşturulan mekanizmanın donanım kullanım miktarları incelenmiştir.

Üçüncü durumda literatürde öneri olarak yer alan, AdjointVM IDS yaklaşımına bazı ek iyileştirmeler ekleyerek eksikliklerinin giderilmesiyle oluşturulan İyileştirilmiş AdjointVM Koruma Mekanizması sunucu-tabanlı saldırı tespit sistemi mimarisi gerçekleştirilerek donanım kullanım miktarları incelenmiştir. Yapılan çalışmalar kapsamında oluşturulan güvenlik mekanizmalarının birinci durumdan son duruma doğru donanım kullanım miktarları karşılaştırıldığında güvenlik

seviyesinin yükseltilmesinin, daha iyi ve daha dayanıklı saldırı tespit sistemi yapılandırılmasının sistemdeki donanım kullanım miktarlarını arttırdığı gözlemlenmiştir. Bulut bilişimde kullanıcıların güvenlik endişelerini giderilmesine katkı sağlamak maksadıyla yapılan bu çalışmanın öncelikli hedefi sistemin güvenliğinin en üst düzeye çıkarılması olduğu için sisteme sunduğu güvenlik katkılarıyla donanım kullanım miktarlarındaki artışın tolore edilebileceği sonucuna varılmıştır.

Gelecek çalışma olarak İyileştirilmiş AdjointVM Koruma Mekanizmasının eksik yönleri tespit edilerek bu eksikliklerin giderilmesini sağlayan, böylelikle sistemin direncinin arttırılmasını ve esnek bir güvenlik politikasının oluşturulmasını hedefleyen yeni bir güvenlik mekanizması önerilecektir.

KAYNAKLAR

[1] D. Teneyuca, "Internet cloud security: The illusion of inclusion", Information Security Technical Report, doi:10.1016/j.istr.2011.08.005, 2011.

[2] P. Mell, and T. Grance, "The NIST Definition of Cloud Computing, NIST Special Publication 800-145 (SP800-145)", National Institute of Standards and Technology, September 2011.

[3] D. Marshall, S. S. Beaver, J. W. McCarty, "VMware ESX: Essentials in the Virtual Data Center", CRC press, 2009.

[4] M.R. Farcasescu, "Trust Model Engines in Cloud Computing", 14th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), 2012, pp. 465-470.

[5] "Security for Cloud Computing 10 Steps to Ensure Success", (2012, August), Cloud Standards Customer Council [Online], Erişilebilir: http://www.cloud-council.org/Security_for_Cloud_Computing-Final_080912.pdf.

[6] A. Kumar, V. Kumar, P. Singh, & A. Kumar, "A Novel approach: Security measures and Concerns of Cloud Computing", International Journal of Computer Technology and Applications, vol. 3(3), 2012, pp. 1008 -1014.

[7] K. Scarfone, and P. Mell, "Guide to Intrusion Detection and Prevention Systems (IDPS)", NIST Special Publication 800-94 (SP800-94), National Institute of Standards and Technology, Gaithersburg, 2007.

[8] V. Marinova-Boncheva , "A short survey of intrusion detection systems", Problems of Engineering Cybernetics and Robotics, 58, 2007, pp. 23–30.

[9] Trusted Computing Group [Online], Erişilebilir: http://www.trustedcomputinggroup.org/.

[10] J. Kong, "Protecting the confidentiality of virtual machines against untrusted host", International Symposium on Intelligence Information Processing and Trusted Computing (IPTC), China, 2010, pp. 364.

[11] J, Kong, "AdjointVM: a new intrusion detection model for cloud computing", Energy Procedia, vol. 13, 2011, pp. 7902-7911.

[12] U. Oktay, M. A. Aydın, O. K. Sahingoz, "A circular chain intrusion detection for cloud computing based on improved AdjointVM approach", Computational Intelligence and Informatics (CINTI), 2013 IEEE 14th International Symposium, IEEE, November, 2013, pp. 201-206.

[13] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, A. Warfield, "Xen and the art of virtualization", ACM SIGOPS Operating Systems Review, 37, 5, 2003, pp. 164-177.

[14] OSSEC HIDS [Online], Erişilebilir: http://www.ossec.net/.

[15] U. Oktay, M. A. Aydın, and O. K. Sahingoz, "Circular Chain VM Protection in AdjointVM", International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE), 2013, pp. 93–97.

[16] "Intel Virtualization Technology for Directed I/O Architecture Specification Rev. 1.3", Intel Corporation, 2011.

BİLGİ HASATLAMASI YÖNTEMLERİ VE KİŞİSEL BİLGİ HASATLAMASI

Celal Turan Ulus, Eyüp Burak Ceyhan, Şeref Sağıroğlu

Celal Turan Ulus, TÜBİTAK, turanulus@gmail.com
Eyüp Burak Ceyhan, Gazi Üniversitesi, ebceyhan@gazi.edu.tr
Prof. Dr. Şeref Sağıroğlu, Gazi Üniversitesi, ss@gazi.edu.tr

Özet — Bu makalede bilgi hasatlamasına dair genel bir bakış ortaya konmaktadır. Bununla birlikte bilgi hasatlamasının daha detaylı bir başlığı olan kişisel bilgi hasatlaması ile ilişkilendirme yapılmıştır. Bilgi hasatlaması bilgi çıkarma ve bilgi çekme olmak üzere iki temel başlıkta incelenmiştir. Bilgi çıkarma işleminde sonuç belli bir sorgu yardımıyla oluşmaktayken bilgi çekme işleminde hedeflenen belli alanlar doğrudan ya da belli örüntüler ile çekilmektedir. Bir sonraki bölümde, yapılan hasatlama çalışmaları olarak bilgi hasatlamasında yapılan bazı çalışmalardan detaylıca bahsedilmiştir. Bilgi hasatlaması terörist organizasyonların takip edilmesinden Wikipedia üzerinden bilgi çekilmesine kadar birçok alanda kullanılmaktadır. Altıncı bölümde bilgi hasatlaması internet ağı üzerinden kişisel bilgi çekilmesi açısından ele alınmıştır. Sosyal ağlarda ve arama motorlarında kişisel bilgilerin çekilmesi ile birlikte ilgili tehlikelerden bahsedilmiştir.

Anahtar Kelimeler — Bilgi hasatlama, internet ağı hasatlaması, bilgi çekme, bilgi çıkarma, kişisel bilgi

Abstract — This paper provides an overview of information harvesting. Besides the overview, information harvesting is associated to personal information harvestiong. Information harvesting is reviewed under two main topics: information retrieval and information extraction. The result occurs by the help of a query in the information extraction whereas the targeted fields of information extraction is harvested directly or by the help of patterns. In the following topic, some studies are mentioned as previous information studies in detail. Information harvesting is used for tracking terrorist organizations, Wikipedia harvesting, etc. In the sixth topic, information harvesting is handled according to personal information harvesting. Personal information harvesting from social networks, search engines and dangers related with them are mentioned.

Index Terms — Information harvesting, web harvesting, information extraction, information retrieval, personal information

I. GİRİŞ

İnternet ağının hızlı büyümesiyle birlikte yüksek miktarda veri, internet kullanıcıları tarafından erişilebilir halde bulunmaktadır. Düşük maliyet, yüksek erişilebilirlik ve özgürce yayın yapabilmek internet ağının karakteristikleri arasında yer almaktadır. Bu durum internet ağının popüleritesini arttırmaktadır. İnternet sayfaları aslında karmaşık metinlerden oluşmaktadır. Metinin ve multimedya bileşenlerinin yanında bağlantılar, HTML etiketleri, tanımlayıcı veri (meta-data) gibi özellikler barındırırlar. Birçok araştırmada internet sayfalarının metin bileşenleri internet ağı hasatlamada en önemli bilgiyi sağladığı varsayılır. Metin olmayan diğer bileşenlerin hasatlama performansını iyileştirdiği varsayılır [1].

İnternet ağı üzerinde veriler yapısal ve yapısal olmayan bir şekilde bulunur. Yapısal veriler alanları, başlıkları, etiketleri belli ve düzenli bir halde bulunur. Bunun sonucunda bilgisayar tarafından kolaylıkla kullanılabilirler. Fakat yapısal olmayan veriler belli bir düzeni olmayan verilerdir. Yapısal olmayan verileri çekmek, okumak ve işlemek daha zordur. Bilgisayardaki ve internetteki bilgilerin çoğu yapısal olmayan veridir [2].

İnternet ağı hasatlamasında, bir ya da iki siteden olan internet sayfaları içeriklerine göre daha önceden tanımlanmış kategorilere atanır. İnternet sayfaları düz metin dokümanlarından daha fazlası olduğu için internet ağı hasatlama metotları diğer içeriklerin niteliğini kullanmayı dikkate almalıdır. Yapılan bir araştırmada internet sayfalarını hasatlamak için düz metin haricinde internet sayfasının başlığı ve sayfa içinde bulunan bağlantıyı belirten sözcük dikkate alınmıştır [1].

Bilgi hasatlama işleminin aşamalarından biri olan bilgi çekme işleminde hasatlanan her bir metin topluluğu genellikle doküman olarak adlandırılır. Dokümanların kendine özgü yapıları vardır. Bir yazılım aracı bu yapıyı belli format işaretlerine ve anahtar sözcüklere göre işaretleyebilir. Fakat tüm bu durumlarda bulunan yapı ilgili kitabın anlamsal içeriğini değil organizasyonel yapısını gösterecektir. Yazılım aracı "bölüm 1", "şekil 1" gibi alanları kolaylıkla bulabilir. Fakat "bilgi çekme" ile ilgili bir başlığı bulmak çok daha zor ve daha belirsiz bir problemdir [3].

Bilgi hasatlama işleminin aşamalarından diğeri bilgi çekme işlemidir. Bilgi çıkarmada belli bir konuda ya da kişi hakkında gereken bilgi doküman içinden çıkartılır. Konu ya da gereken bilgi kullanıcı tarafından belirlenen bir sorgu ile çıkartılır. Belirlenen sorgu tarafından karşılanan dokümanlar kullanıcı tarafından konu ile ilgili, karşılanmayanlar ise ilgisiz olarak nitelenir. Bir bilgi çıkarma motoru dokümanı sınıflandırmak için sorguyu kullanabilir. Sorgu sınıflandırma kriterlerini karşılayan dokümanları sonuç olarak döner [3].

Bu makalenin bundan sonraki bölümünde bilgi hasatlamasından genel itibariyle bahsedilmiştir. Bilgi hasatlamasının yöntemlerinden biri olan bilgi çıkarması işlemlerinden üçüncü bölümde bahsedilmiştir. Dördüncü bölümde yine bilgi hasatlamasının diğer bir yöntemi olan bilgi çıkarması ele alınmıştır. Beşinci bölümde bilgi hasatlamasında internet ağı üzerinde yapılan çalışmalardan bahsedilmiş olup altıncı bölümde kişisel bilgi hasatlamasında güvenliğin önemi örneklerle vurgulanmıştır. Son bölümde ise bilgi hasatlamasının aşamaları hakkında özet bilgilerden bahsedilmiş olup kişisel bilgi hasatlamasının bilgi güvenliği ile ilgisi ortaya konmuştur.

II. BİLGİ HASATLAMASI

İnternetin yükselişi ile birlikte sosyal uygulamalar, bloglar, e-postalar ve çeşitli internet uygulamaları hayatımızın içinde yer almaktadır. Bu uygulamalarda kişisel bilgilerimiz bulunmaktadır ve iyi korunmayan bilgiler risk oluşturmaktadır. Bu bilgiler üzerinde bilgi hasatlaması yapılarak kullanıcıların bilgileri tahmin edilmekte ve ele geçirilmektedir.

Birçok sosyal ağ uygulamaları kişilerin ilişkileri hakkında ilginç bilgileri açığa çıkarmaktadır. Örneğin blog yorumcularının gönderilerini ya da yer imi benzerliklerini analiz edilerek insanlar arasındaki bağlantılar ortaya çıkarılabilir. Birçok kaynaktan bu bilgilerin çıkarımını yaparak ve birleştirerek kişisel ve kurumsal sosyal ağların kapsamlı bir resmi ortaya çıkarılabilir. İki kişinin bir makalede yazarlığı varsa aynı zamanda bir sosyal ağda bağlantısının bulunduğunu birçok çalışma göstermektedir. Yeni çalışmalarda bu iki kişinin bağlantısına kanıt olarak e-posta ve internet sayfaları olarak da gösterilmektedir [4].

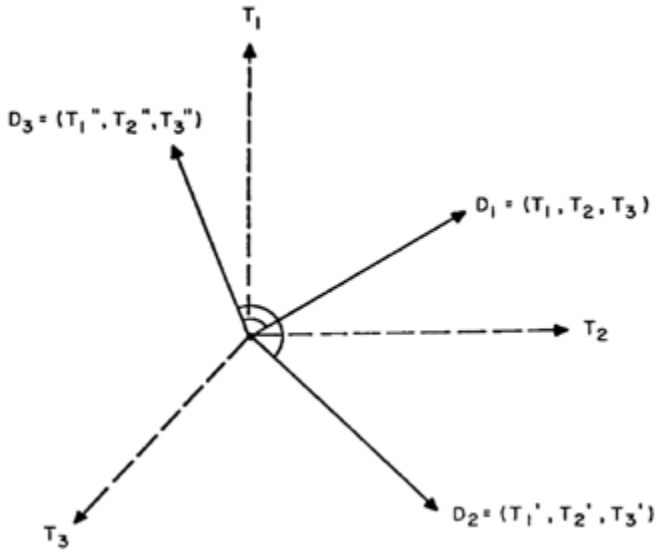
Genellikle bilgi hasatlama teknikleri internet ağı hasatlamasında da kullanılır. İnternet ağı hasatlaması bilgi hasatlamanın genişletilmiş halidir. Bilgi hasatlama çevrimdışı bir işlem iken internet ağı hasatlaması çevrimiçi bir işlemdir. Bilgi Hasatlama verisi çevrimdışı olarak veri ambarında saklanır. İnternet ağı hasatlama verisi sunucu veri tabanında saklanır [5].

III. BİLGİ ÇIKARMA YÖNTEMLERİ

A. Vektör Uzay Modeli

Vektör uzay modeli (VSM) 1975'te Salton ve çalışma arkadaşları tarafından SMART sistemi için geliştirilmiştir. Vektör ağırlık modeli bilgi çıkarımı için kullanımının basitliğinden dolayı en çok kullanılan modellerden biridir. Bir VSM'de her bir doküman bir vektör olarak kabul edilir. Her boyut bir terimi ifade eder. Bir terim bir dokümanın içinde yer alıyorsa o terimin vektördeki değeri sıfır değildir [6].

Bir grup Di dokümanından oluşan doküman uzayı olsun. Her biri bir veya birden fazla indeks Tj terimi ile tanımlansın. Terimler kendi önemlerine göre 0 ve 1 arasından ağırlıklandırılır. Tipik bir üç boyutlu vektör uzayı Şekil 1'de gösterilmiştir. t farklı terim indeksi varsa üç boyutlu örnek t boyuta kadar genişletilebilir [7].



Şekil.1 Doküman uzayının vektörel gösterimi [7]

Terim frekansı (tf) vektör uzay modeli ile birlikte en çok kullanılan yöntemlerden birisidir. Bir çeşit sözcüğün kaç defa bir doküman içinde geçtiğinin sayısıdır. Doküman frekansı (df) ise bir kelimeyi en az bir kere içeren doküman sayıdır. Terim frekansı 0 ve N arasında bir tam sayıdır. Doküman frekansı ise 0 ve D arasında bir tam sayıdır. Bilgi çıkarımında doküman frekansları ters doküman frekanslarına çevrilir. Ters doküman frekansı (IDF) terim ağırlıklandırmada önemli bir rol oynar [8]. IDF(t) t terimini içeren dokümanın sahip olduğu bilgi biti sayısı olarak yorumlanabilir.

$$IDF(t) = -\log_2 \frac{df(t)}{D} \quad (1)$$

Bir dokümandaki her terime sayısal ağırlık belirlenebilir. Böylece ilgili terimin ilgili doküman içindeki yararlılığı sayısal olarak ölçülmüş olur. Yararlılıktan kastedilen durum belli dokümanı diğer dokümanlardan ne kadar ayırt edici olabildiğidir. Belli bir terim farklı dokümanlar içinde farklı ağırlıklara sahip olabilir. Çünkü bir terim, bir doküman için diğer dokümanlara olduğundan daha iyi bir tanımlayıcı ya da ayırt edici olabilir [3].

Doküman uzayında her doküman D, dokümanın içinde yer alan terimlerin ağırlıkları ile tanımlanır. Terim uzayında her doküman bir boyuta karşılık gelir. Terim uzayında vektör bir terimdir. Bir terimin koordinatları, içinde geçtiği dokümanla ilişkili ağırlığı olarak yorumlanır [3]. Burada önemli bir ayrıntı terim ağırlığının nasıl bulunacağıdır. Terim ağırlığını bulurken en çok kullanılan yöntem aşağıdadır.

$$w = tf*IDF \quad (2)$$

Bir terimin bir dokümanın içindeki ağırlığı bulunurken (2) kullanılır. Dokümanın içindeki tf terim frekansını göstermektedir [3]. İki dokümanın benzerliğini bulurken (3) kullanılır. Bu eşitliğe kosinüs benzerliği denir. İki doküman arasındaki kosinüs açısına göre benzerlikleri belirler [9].

$$Sim(v_1, v_2) = \frac{\sum_{i=1}^{n} v_1.v_2}{\sqrt{\sum_{i=1}^{n} v_{1i}^2}\sqrt{\sum_{i=1}^{n} v_{2i}^2}} \quad (3)$$

v metin vektörüdür. v1i , v1 vektöründe bulunan i. sözcüğün tf-idf ağırlığıdır.

İkili İşlem Modeli

İkili işlem modelinde sorgular oluşturulur. Sorgu, ikili işlem operatörlerine dayanarak oluşturulur. Genel bir ikili işlem sorgusu AND, OR ve NOT operatörlerinden oluşur. Örneğin t1 ve t2 terimlerini D1 dokümanı içeriyorsa "t1 AND t2" sorgusu D1 dokümanı tarafından karşılanır. Benzer şekilde t1 ve t2 terimlerinden birini D1 dokümanı içeriyorsa "t1 OR t2" sorgusu karşılanır. "t1 AND NOT t2" sorgusunda ise t1 terimi varsa ve t2 terimi D1 dokümanında yoksa D1 dokümanı tarafından sorgu karşılanmış olur. Daha karmaşık ikili işlem sorguları oluşturulup bunların sonuçları ilgili doküman açısından sorgulanabilir. Klasik bir ikili işlem modeli sonuç olarak true ya da false değeri döner. Dolayısıyla bir doküman üzerinde işletilen bir ikili işlem sorgusu sonucunda o doküman sorgu içindeki terimle ya alakalı ya da alakasız bir doküman olur. Dokümanlar arasında herhangi bir sıralama yapılmaz. Klasik bir ikili işlem modelinde terim ağırlıkları kullanılmaz. Bir terimin ağırlığı ya birdir (terim vardır) ya da sıfırdır (terim yoktur) [3].

IV. BİLGİ ÇEKME YÖNTEMLERİ

A. İsimlendirilmiş Varlık Tanımlaması

NER isimlendirilmiş varlıkların önceden tanımlanmış türlerinin tespit edilmesi ve sınıflandırılması problemidir. İsimlendirilmiş varlıklara kuruluşlar (Dünya Sağlık Örgütü), kişiler (Muammer Kaddafi), yer isimleri (Baltık Denizi) örnek olarak verilebilir. NER metinden tespit edilen varlıklar hakkında tanımlayıcı bilgi çekilebilir. Örneğin kişi durumu için ünvan, mevki, milliyet, cinsiyet ve kişinin diğer nitelikleri hakkında bilgi çekilebilir [10].

NER'in başarısı, standart varlıklar için %95'e ulaşmaktadır [11]. NER için kullanılan metotlar farklı boyutlara göre sınıflandırılabilirler. Bu boyutlardan birisi manuel-otomatik zıtlığıdır. Bazı yöntemler kaynakları manuel olarak kullanırken diğerleri işaretlenmiş eğitim verisinden otomatik olarak bir model üretmek için öğrenme algoritmalarını kullanır. Diğer boyutta ise varlık modelinin özelliği belirtilir: Bazı modeller sembolik iken diğerleri sayısaldır. Sembolik modellerde kaynaklar açık ve anlaşılabilirdir. Sembolik modellere örnek olarak kural tabanlı sistemler verilebilir. Sayısal modellere ise istatistiksel model örnek verilebilir. Kural tabanlı metotlar ilgili varlığın içeriğini temsil eden kuralları kullanırlar. Örneğin kişinin ismini algılamak için kişinin adının bir ünvandan sonra gelecek şekilde kodlayan aşağıdaki gibi bir kural tanımlanabilir [12] : "Mr.+ capitalized_word"

Bu kurallar kural kümesinin sorumluluğunu alacak uzmanlar tarafından geliştirilir. Kural tabanlı sistemler bilgi çekmenin ilk yıllarında üretilmesine rağmen hala gerçek dünya sistemler etkin bir şekilde kullanılmaktadır. Öğrenme tabanlı metotlar varlıkları tanımlayacak modeli öğrenmek için işaretlenmiş veriyi kullanırlar. İşaretlenmiş veri içinde tipleri belirtilmiş şekilde bulunan varlık bulunduran dokümandır. Öğrenilen modeller sembolik ya da istatistiksel olabilir. Sembolik modeller kural öğrenen bir algoritmayı kullanırlar. Kural öğrenen algoritma isimlendirilmiş varlıkların tanımlanması için kural kümesi oluşturur. İstatistiksel modeller, standart istatistiksel makine öğrenme algoritmalarıdır. Makine öğrenme algoritmaları içeriği ya da varlığın niteliğinin gösterimine dayanır. Bu modellerde NER sınıflandırma görevi görür [12].

B. Çoklu Referans Çözümlemesi

Bu yöntemi uygulayabilmek için metin içinde aynı varlığı tanımlayan birden fazla referansın bulunması gerekir. Metin içinde birden fazla bulunma durumu aşağıdaki örneklerle açıklanabilir [10] :

İsimlendirilmiş Varlık: Varlığın isimlendirilmesi durumudur. 'General Electric' ve 'GE' aynı metin içinde aynı varlığı işaret edebilir.
Zamir Durumu: Varlığın zamirle ifade edilmesi durumudur. 'John bought food. But he forgot to buy drinks.' Zamir 'he', 'John'u işaret etmektedir.
Temsili Durum: Varlığın temsili bir sözcük ile ifade edilmesi durumudur. 'Microsoft revealed its earnings. The company also unveiled future plans.' 'The company' sözcüğü 'Microsoft'u temsil etmektedir.

C. İlişki Çekimi

Metinde bulunan varlıklar arasında önceden tanımlanmış ilişkileri tespit etme ve sınıflandırma yöntemidir [13]. Aşağıda ilişki çekimi ile ilgi bazı örnekler verilmiştir:

Çalışan (Bill Gates, Microsoft): Bir kişi ve bir kurum arasındaki ilişkidir. 'Bill Gates works for Microsoft' cümlesinden çıkarılmıştır.
Konum (Uslu, Washington): Bir kişi ve bir konum arasındaki bir ilişkidir. Bu ilişkide kişinin hangi konumda olduğu ilişkisi belirtilir.'Mr. Uslu talked at the conference in Washington' cümlesinden çıkarılmıştır.
AltŞirket (ARK, Seyhan Holding): İki şirket arasındaki ilişkiyi gösterir. Bir şirket diğer şirketin alt şirketidir. 'Listed broadcaster ARK said its parent company, Seyhan Holding, is considering various options for the potential sale'. Genellikle çıkartılabilecek ilişki sayısı limitsiz olmasına rağmen sonuca yönelik çıkartılan ilişki kümesi önceden tanımlanmış, sınırlı ve sabit olmalıdır.

D. Olay Çekimi

Metin içinden olayları tespit ettikten sonra bilgileri detaylı ve yapısal bir biçimde kullanıcıya sunma işlemidir. Yapısal bilgi içerisinde 'kimin kime ne yaptığı', ne zaman, nerde, neler kullanarak ve nasıl yaptığı gibi bilgiler tespit edilebilir. Genelde olay çekimi birkaç varlık ve bu varlıkların arasındaki ilişkilerin çekimini kapsar. Örneğin bir metinden terörist saldırısı ile ilgili bilgi çekme yapılabilir. 'Masked gunmen armed with assault rifles and grenades attacked a wedding party in US, killing at least 44 people.' Belirtilen cümleden olayın failleri (masked gunmen), kurbanlar (people), öldürülen/yaralanan sayısı (at least 44), kullanılan silah ve cephaneler (rifles and grenades), konum (US) bilgileri çıkartılabilir [10].

V. YAPILAN HASATLAMA ÇALIŞMALARI

A. Ağırlıklandırma Değerleri İle Hasatlama

Amerikan Homeland Security'nin yapmış olduğu bir çalışmada bilgi hasatlama yöntemi ile radikal fikirlere sahip internet sitelerini tespit edilmektedir. Yapılan çalışmada iki çeşit analiz yapılmıştır. Biri link analizi diğer ise içerik analizidir. Radikal grupların amaçlarını anlayabilmek için nitelik tabanlı bir metodoloji geliştirilmiştir. Bunlar; iletişim, para toplama, ideoloji paylaşımı, iç propaganda, dış propaganda, sanal topluluk, komuta kontrol, eleman toplama ve eğitim olarak sıralanır. Seçilen nitelikler 13 yıl tecrübeli bir CIA istihbarat analistinin tecrübesi vasıtasıyla seçilmiştir [14].

Yüksek seviyeli her bir nitelik düşük niteliklerin bileşiminden oluşmaktadır. Örneğin iletişim e-posta bağlantısı, telefon bağlantısı, multimedya dosyaları, online geribildirim formu ve dokümantasyondan oluşmaktadır. Detaylardaki düşük seviyeli nitelikleri tanımlayan kodlama şeması geliştirilmiştir. Kodlama şeması aracı para toplama ya da propaganda gibi belli kaynakları ayırma örüntüsünü bulur. Böylece izlenen grupların internet ağını nasıl kullandığı öğrenilmiş olur. Düşük seviyeli bir niteliğe ağırlık atamak belli amaçlar için kullanım seviyesini ölçmeye yarar.

B. Kural ve Örüntü Tabanlı Metodlar İle Yapılan Hasatlama

[15]'te yapılan çalışmada çeşitli bilgi çekme yöntemleri tartışılmıştır. Arama motorlarının işlevselliğinin anlamsal seviyesini yükselten başlıca eğilimler vardır. Amaç otomatik bir şekilde isimlendirilmiş varlıklarla alakalı kapsamlı bir bilgi tabanı, anlamsal sınıflarını ve ortak ilişkilerini yüksek oranda başarı ile oluşturmaktır. Yapılan diğer çalışmada [11] hasatlamanın ilk seviyesinde tüm varlıklar toplanmıştır. Bu varlıklar kişiler, şirketler, şehirler ve ürünler gibi varlıklardır. Varlıklar anlamsal sınıflara ayrılmıştır. Örneğin bu sınıflar sanatçılar, bilim insanları, moleküler biyologlar vs. olabilir. Belli bir varlık birden fazla sınıfa dahil olabilir. Örneğin Angela Merkel politikacı, bilim insanı, başbakan gibi birden fazla sınıf içine dahil olabilir.

WordNet [16] İngilizce kelimelerin sözlüksel anlamlarını barındıran bir çatıdır. YAGO (Yet Another Great Ontology) [17] isimli çalışmada WordNet ile Wikipedia üzerinde çalışılmıştır. YAGO kendi bünyesinde barındırdığı sınıfları WordNet'ten içeri aktarmıştır. YAGO'nun Wikipedia'da bulduğu her varlık YAGO'nun belirlediği sınıflardan birine atanmalıdır. Eğer varlık atama işlemi başarısız olursa varlık bilgi tabanına dahil edilmez. Varlıkların sözlüksel anlamları ve dahil oldukları sınıflar ile ilgili bilgi elde edildikten sonra varlıklar hakkında ilişkisel bilgiler elde edilebilir. İlişkilerden ikili ilişkiler ele alınabilir. Örneğin dogum tarihi ⊆ kişi x şehir, evlilik ⊆ kişi x kişi, mezuniyet ⊆ kişi x üniversite gibi ilişkiler ikili ilişkilere örnek verilebilir [15].

Doğal dillerin kısıtlamaları vardır. Bunlardan birisi olarak isim kelimesi sadece bazı fiiller ile birlikte kullanılabilir. Örneğin meyve suyu içilebilir ya da üretilebilir. Fakat yenilemez ya da sürülemez. İsim kelimeleri belli fiil kelimelerine göre kümelenebilir. Heaest örüntüleri POS (Part Of Speech) ile zenginlendirilmiş düzenli ifadelerdir [11]. Hearst örüntüleri [18] serbest formatlı metin ifadelerinden gelen önceden tanımlı ilişkisel modelin örneklerini bulmayı amaçlar. Örneğin instanceOf ilişkisi için isim örnekleri otomatik olarak aşağıdaki örüntüden tespit edilebilir:

$$NP_0 : \{NP_1, NP_2 ... (\text{and} \mid \text{or})\}\ NP_n \quad (4)$$

(4) numaralı örüntüde belirtilen NP özel isimler için bir POS etiketidir. Hearst örüntüleri yüksek duyarlılık oranına sahiptir. Fakat düşük duyarlılık değerine sahiptir. Hearst örüntüleri el ile yazılır. Otomatik olarak üretilmez. Bunun için örüntüleri üretmek zordur.

Elle üretilen örüntülerde yüksek duyarlılık değeri mümkün olabilmekte fakat genellikle düşük duyarlılık değerleri çıkmaktadır. Bunun için elle üretilen örüntülerin yanında otomatik örüntü üreten çalışmalar da olmuştur. Örneğin KnowItAll [19] çalışmasından düşük hata positif oranı (false-positive rate) ile birlikte yüksek çağrı değeri elde edilmektedir.

C. İnternet Ağı Üzerindeki İlişkisel Tablolardan Yararlanarak Yapılan Hasatlama

İnternet sayfaları üzerinde veri tutan birçok listeler vardır ve bu listelerden bilgi hasatlaması yapılabilir. İnternet sayfası üzerinde bulunan listeler çok kolonlu tablolara dönüştürülebilir. Tablo bilgilerinden öncelikle kolon bilgilerinin çekilmesi gerekmektedir. Bunun için tabloda bulunan alanların kalitesi ölçülmektedir. Kaliteyi ölçen metrik alan kalite skorudur (FQ) [20].

$$FQ(f) = a_{ts} \times S_{ts}(f) + a_{lms} \times S_{lms}(f) + a_{tcs} \times S_{tcs}(f) \quad (5)$$

(5)'te Sts (f) tip desteğidir. Slms dil model desteğidir. Stcs tablo metin desteğidir. Her bilgi kaynağı bir ağırlığa atanır. Bu ağırlıklar sırasıyla a_{ts}, a_{lms} , ve a_{tcs}'dir [20].

Tip Destek Skoru (Sts): Tip destek skoru herhangi bir alanın ayrık tablo kolonlarında sık sık bulunup bulunmadığını anlamaya yarar. Yapılan çalışmada sayısal değerler, tarih değerleri, URL'ler, e-postalar ve telefon numaraları alan olarak belirlenebilir. f'in tipi belirlenirse $S_{ts}(f)$ değeri 1'e eşitlenir aksi takdirde 0'a eşitlenir [20].

Dil Model Destek Skoru (Slms): Dil modeli sözcük dizilerinin oluşma ihtimalini belirler. İki çeşittir. Biri içsel uyum skorudur. Diğeri dışsal uyum skorudur. İçsel uyum skoru S_{ic} ile gösterilir. Dışsal uyum skoru S_{ei} (f) ile gösterilir. Her iki skor (6) ve (7)'da gösterilmiştir [20].

$$S_{ic}(f) = \frac{\sum_{h=1}^{m-1} Pr(w_{i+h} | w_i, \ldots, w_{i+h-1})}{m-1} \quad (6)$$

$$S_{ei}(f) = \frac{2}{Pr(w_i | w_{i-1}) + Pr(w_{i+h+1} | w_{i+h})} \quad (7)$$

(6)'da $Pr(w_i | w_1, \ldots, w_{i-1})$, w_i'nin $(w_1, \ldots, w_i)$ sözcük dizisini takip etme ihtimalidir. m ise bir satırdaki sözcük sayısını ifade eder. (7)'de $Pr(w_i | w_{i-1})$, f'de bulunan ilk sözcüğün son sırada bulunan sözcüğü takip etmesi ihtimalidir. Yine (7)'de bulunan $Pr(w_{i+h+1} | w_{i+h})$, bir sonraki alanda bulunan ilk sözcüğün f'deki son sözcüğü takip etmesi ihtimalidir. Dil model skoru içsel ve dışsal uyumun ağırlıklı ortalamasıdır [20].

$$S_{lms}(f) = a_{ic} \times S_{ic}(f) + a_{ei} \times S_{ei}(f) \quad (8)$$

a_{ic} ve a_{ei} 0 ve 1 aralığında bulunan değerlerdir ve $a_{ic} + a_{ei} = 1$'dir.

Tablo metin destek skoru (Stcs): Tablo metin destek skoru f'in internet sayfasında bulunan tablolarındaki metinlerde ne kadar desteklendiğini gösterir. tc_support değeri bir tabloda f'in kaç kere hücre değeri olarak bulunduğunu göstersin. tc_support değeri min_tc_support değerinden küçükse Stcs değeri 0, büyükse 1 olarak atanır.

Tablo 1'de yapılan çalışmaya ListExtract adı verilmiştir ve RoadRunner [21] ile karşılaştırma yapılmıştır.

Uygulama	Duyarlılık	Çağrı	Ağırlıklı Ölçüm
ListExtract	0.64	0.63	0.63
RoadRunner	0.39	0.28	0.32

Tablo I - Listextract ve roadrunner karşılaştırılması [20]

D. Wikipedia'dan Bilgi Hasatlama Çalışması

Yapılan bir çalışmada [22] Timely YAGO (T-YAGO) isimli bir bilgi tabanı ortaya konmuştur. Timely YAGO Wikipedia'dan başlıklardan, listelerden ve kategorilerden zamansal olguları çıkarabilmektedir. Aynı zamanda zamansal olgular sorgulanabilmektedir. T-YAGO zamansal bilgi tabanına dayanarak SPARQL isimli bir sorgu dili imkânı sağlar. Olgular özne, nitelik ve nesne üçlüsü halinde gösterilir. Kullanılan zamansal koşulları gösteren sözcükler on, since ve until olarak kullanılmıştır. Zamansal olgular arasındaki ilişkileri gösteren sözcükler şunlardır [2] : before, after, equal, during, overlaps, sameYear

Yukarıda kullanılan sözcüklere benzer ilişkisel başka sözcükler de kullanılmaktadır. Fakat temel olarak kullanılanlar yukarıdadır. Örnek olarak David Beckham'la aynı takımda ve aynı zaman aralığında oynayan oyuncuları sorgulayan sorgu aşağıdadır[22].

```
?id1: "David Beckham" playsForClub ?x .
?id2: ?a playsForClub ?x .
?id1 since ?t1 . ?id1 until ?t2 .
?id2 since ?t3 . ?id2 until ?t4 .
[?t1-?t2] overlaps [?t3-?t4] .
?a notEqual "David Beckham"
```

Şekil.2 SPARQL sorguları[22]

[?t1-?t2] zaman aralığını temsil eder. Sorguda overlaps koşulu iki zaman aralığının örtüşüp örtüşmediğini belirler.

VI. KİŞİSEL BİLGİ HASATLAMASI

Bilgi hasatlamasının ilgi alanı içinde kişisel bilgilerin hasatlaması da yer almaktadır. Dolayısıyla herkese açık olan sosyal medya servisi, blog ve diğer internet sitelerinden kişisel bilgilerin hasatlaması yapılabilir.

Kişisel bilgi olarak internet sitelerinden örneğin bir kişinin işten atılıp atılmadığı anlaşılabilir. Örneğin Türkçe olarak "işten atıldım", "işten kovuldum" gibi cümle parçacıkları aranarak ilgili cümle içinde kişi ismi de varsa eşleşme yapılabilir [23]. Diğer bir sorun TC Kimlik numarasının internette excel, pdf, word gibi dosyalarda açıkça bulunmasıdır. Kişisel bilgi hasatlamasında kişi ismi ile istenen kişinin TC kimlik numarası bulunabilir.

2009 yılında yapılan bir çalışmada MySpace'de kişilerin benzerlik oranları çıkarılmıştır. Arkadaş olanların birbirine yakın yaşlarda aynı dini görüşlere sahip olduğu ve çocuklara karşı benzer yaklaşımları olduğu çıkarılmıştır. Fakat cinsiyet konusunda benzerliğe dair herhangi bir sonuç bulunamamıştır [24].

Sosyal ağlarda birçok kullanıcının e-posta, okul, iş gibi kişisel bilgileri tüm arkadaşlarına açıktır. Yapılan bir çalışmada kişilerin facebook vb. sosyal ağlarda oturumları çalınarak tüm arkadaş bilgileri çekilmiştir. Arkadaş bilgileri çekildikten sonra bu arkadaşlar taklit edilmiştir. Oturum çalma işleminden sonra kullanıcının doğum günü ya da konum bilgisi ile de saldırı gerçekleştirilebilir [25]. E-mail adresleri, anlık mesajlaşma bilgileri gibi hassas kişisel bilgiler spam mesajları tarafından inandırıcılığı arttırılabilmesi için kullanılabilmektedir. Sosyal ağlardaki kişisel bilgi havuzuna erişim sağlamak ve bir sosyal ağ kullanıcısını taklit etmek çözülmesi kolay olmayan bir zorluktur [26]. İlk yapılan çalışmalar [27,28] sosyal ağlarda bilgi çıkarımına dair bir farkındalık oluşturdu. Çünkü bu çalışmalardan sonra sosyal medya üzerinde yapılan çalışmalar artmıştır. Yapılan diğer bir çalışmada çeşitli sosyal ağlarda bulunan kullanıcıların bilgilerini çekebilen ve bu çekilen bilgileri kullanarak sahte profil üretebilen iCloner isimli bir sistem geliştirilmiştir. iCloner birden fazla sahte profil üretirken CAPTCHA'yı analiz ederek çalışmaktadır.[29].

VII. SONUÇ

İnternet ağı üzerinden bilgi hasatlaması en önemli araştırma konularından birisidir. Bilgi hasatlaması araştırmasında öne çıkan iki temel konu vardır. Birisi bilgi çıkarma (information retrieval) diğeri ise bilgi çekme (information extraction) dir. Bilgi hasatlaması ile ilgili detaylı bir literatür taraması yapılmıştır.

Bilgi çıkarmada en çok kullanılan yöntemlerden biri vektör uzay modeli (VSM)'dir. Vektör uzay modelinde dokümanlar arasındaki benzerlik kosinüs açıları ile bulunur. Aralarındaki açı küçük olan dokümanlar birbirine daha çok benzer olarak kabul edilir. Vektör uzay modeli ise boyutu büyük olan dokümanlarda düşük performans gösterir. Çünkü ilgili dokümanda her bir terim için terim frekansı ve ters doküman frekansı hesaplanır [30]. Vektör uzay modeli incelendikten sonra ikili işlem modeli incelenmiştir. İkili işlem modeli dokümanlar üzerinde çalışırken ilgili terim model için ya vardır ya da yoktur [31]. Bundan dolayı ya çok az ya da çok fazla doküman üzerinde çalışır.

Makalede incelenen bilgi çekme yöntemleri isimlendirilmiş varlık (NER) tanımlaması, çoklu referans çözümlemesi, ilişki çekimi ve olay çekimidir.

Bilgi hasatlamasında en büyük iki engel çok anlamlılık (polysemy) ve eş anlamlılıktır (synonymy). Çok anlamlılık bir kelimenin birden fazla anlamı olmasıdır. Eş anlamlılık birden fazla kelimenin bir anlamı olmasıdır [32].

Bilgi hasatlaması bilgi çekme ve bilgi çıkarma olarak kategorik incelendikten sonra bilgi hasatlaması kişisel bilgi hasatlaması özelinde olarak incelenmiştir. Bilgi hasatlamasında kişisel bilgilerin ele geçirilebileceği görülmüş olup araştırmanın sonunda bu konu üzerinde durulmuştur. Özellikle sosyal medyada kişisel bilgilerin internet ağına ne kadar açıkta olduğunun farkında olunması lazımdır. Çünkü bilgileri hasatlanan kişilerin bilgileri kullanılarak sosyal mühendislik saldırıları gerçekleştirilebilmektedir. Bununla birlikte özellikle Facebook başta olmak üzere sosyal ağları tarayan internet robotları bulunmaktadır. Gizlilik özelliklerini yeterince kullanmayan sosyal medya kullanıcıların bilgileri hiç kullanmamış oldukları sahte sosyal ağlarda bulunabilmektedir. Yine arama motorları yardımıyla kişisel bilgileri içeren dosyalar internet üzerinde bulunabiliyorsa bu dosyaların arama motorları tarafından indekslenmesi engellenmelidir.

KAYNAKLAR

[1] A. Sun, E. P. Lim, and W. K. Ng, "Web classification using support vector machine," in Proceedings of the 4th international workshop on Web information and data management, 2002, pp. 96–99.

[2] Y. Bassil, "A Survey on Information Retrieval, Text Categorization, and Web Crawling," J. Comput. Sci. Res., vol. 1, no. 6, pp. 1–11, 2012.

[3] E. Greengrass, "Information retrieval: A survey," Information Retrieval. p. 224, 2000.

[4] I. Guy, M. Jacovi, E. Shahar, N. Meshulam, V. Soroka, and M. Carmel, "Harvesting with SONAR - The Value of Aggregating Social Network Information," Soc. Networks, pp. 1017–1026, 2008.

[5] K. Sharma, G. Shrivastava, and V. Kumar, "Web mining: Today and tomorrow," 2011 3rd Int. Conf. Electron. Comput. Technol., vol. 1, pp. 399–403, 2011.

[6] L. Duan, S. Oyama, M. Kurihara and H. Sato, 'Establishing Relationships between Emotion Taxonomies Using the Vector Space Model', Lecture Notes in Engineering and Computer Science, vol. 2215, no. 1, pp. 19-24, 2015.

[7] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," Communications of the ACM, vol. 18, no. 11. pp. 613–620, 1975.

[8] M. Yamamoto and K. W. Church, "Using Suffix Arrays to Compute Term Frequency and Document Frequency for All Substrings in a Corpus," Computational Linguistics, vol. 27, no. 1. pp. 1–30, 2001.

[9] X. Li and W. Cao, "A method for person name disambiguation based on Baidu Encyclopedia," in Transportation, Mechanical, and Electrical Engineering (TMEE), 2011 International Conference on, 2011, pp. 423–426.

[10] J. Piskorski and R. Yangarber, "Information extraction: Past, present and future," in Multi-source, multilingual information extraction and summarization, Springer, 2013, pp. 23–49.

[11] R. Besancon, G. de Chalendar, O. Ferret, F. Gara, O. Mesnard, M. Laib, and N. Semmar, "LIMA : A Multilingual Framework for Linguistic Analysis and Linguistic Resources Development and Evaluation," in Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), 2010.

[12] S. Elloumi, A. Jaoua, F. Ferjani, N. Semmar, R. Besançon, J. Al-Jaam, and H. Hammami, "General learning approach for event extraction: Case of management change event," J. Inf. Sci., p. 0165551512464140, 2012.

[13] N. Bach and S. Badaskar, "A review of relation extraction," Lit. Rev. Lang. Stat. II, 2007.

[14] Y. Zhou, E. Reid, J. Qin, H. Chen, and G. Lai, "US domestic extremist groups on the Web: link and content analysis," IEEE Intell. Syst., vol. 20, no. 5, 2005.

[15] G. Weikum and M. Theobald, "From information to knowledge: harvesting entities and relationships from web sources," Proc. twenty-ninth ACM SIGMOD-SIGACT-SIGART Symp. Princ. database Syst., pp. 65–76, 2010.

[16] P. University, 'About WordNet - WordNet - About WordNet', Wordnet.princeton.edu, 2015. [Online]. Available: https://wordnet.princeton.edu. [Accessed: 12- Jul- 2015].

[17] F. M. Suchanek, G. Kasneci, and G. Weikum, "Yago," in Proceedings of the 16th international conference on World Wide Web - WWW '07, 2007, p. 697.

[18] M. A. Hearst and M. A. Hearst, "Automatic Acquisition of Hyponyms from Large Text Corpora," in Proceedings of the 14th International Conference on Computational Linguistics, 1992, pp. 539–545.

[19] O. Etzioni, M. Cafarella, D. Downey, S. Kok, A.-M. Popescu, T. Shaked, S. Soderland, D. S. Weld, and A. Yates, "Web-Scale Information Extraction in KnowItAll (Preliminary Results)," in WWW'04 Proceedings of the 13th international conference on World Wide Web, 2004, pp. 100–110.

[20] H. Elmeleegy, J. Madhavan, and A. Halevy, "Harvesting relational tables from lists on the web," VLDB J., vol. 20, no. 2, pp. 209–226, 2011.

[21] V. Crescenzi, G. Mecca, and P. Merialdo, "RoadRunner: automatic data extraction from data-intensive web sites," in Proceedings of the 2002 ACM SIGMOD international conference on Management of data, 2002, p. 624.

[22] Y. Wang, M. Zhu, L. Qu, M. Spaniol, and G. Weikum, "Timely YAGO : Harvesting , Querying , and Visualizing Temporal Knowledge from Wikipedia," Proc. 13th Int. Conf. Extending Database Technol. (EDBT), Lausanne, Switzerland, March 22-26, pp. 697–700, 2010.

[23] D. Wilkinson and M. Thelwall, "Researching Personal Information on the Public Web: Methods and Ethics," Social Science Computer Review, vol. 29, no. 4. pp. 387–401, 2011.

[24] M. Thelwall, "Homophily in MySpace," J. Am. Soc. Inf. Sci. Technol., vol. 60, no. 2, pp. 219–231, 2009.

[25] M. Huber, M. Mulazzani, E. Weippl, G. Kitzler, and S. Goluch, "Friend-in-the-middle attacks: Exploiting social networking sites for spam," IEEE Internet Comput., vol. 15, no. 3, pp. 28–34, 2011.

[26] M. Huber, M. Mulazzani, E. Weippl, G. Kitzler, and S. Goluch, "Exploiting social networking sites for spam," Proc. 17th ACM Conf. Comput. Commun. Secur. - CCS '10, p. 693, 2010.

[27] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, "Social phishing," Communications of the ACM, vol. 50, no. 10. pp. 94–100, 2007.

[28] H. Jones and H. Soltren, "Facebook : Threats to

Privacy," Soc. Sci. Res., vol. December 1, pp. 1–76, 2005.

[29] L. Bilge, T. Strufe, D. Balzarotti, E. Kirda, and S. Antipolis, "All Your Contacts Are Belong to Us : Automated Identity Theft Attacks on Social Networks," Www 2009, pp. 551–560, 2009.

[30] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," Information Processing & Management, vol. 24, no. 5. pp. 513–523, 1988.

[31] P. Castells, M. Fernández, and D. Vallet, "An adaptation of the vector-space model for ontology-based information retrieval," IEEE Trans. Knowl. Data Eng., vol. 19, no. 2, pp. 261–272, 2007.

[32] M. W. Berry, Z. Drmac, and E. R. Jessup, "Matrices, Vector Spaces, and Information Retrieval," SIAM Review, vol. 41, no. 2. pp. 335–362, 1999.

HONEYTHING: NESNELERİN İNTERNETİ İÇİN TUZAK SİSTEM

Ö. Erdem, Dr. M. Kara, A. İkinci

Ö. Erdem TÜBİTAK BİLGEM Siber Güvenlik Enstitüsü, 41470 Gebze/Kocaeli TÜRKİYE (e-mail: omer.erdem@ tubitak.gov.tr).
Dr. M. Kara TÜBİTAK BİLGEM Test ve Değerlendirme Başkan Yardımcılığı, 41470 Gebze/Kocaeli TÜRKİYE (e-mail: mehmet.kara@ tubitak.gov.tr).
A. İkinci Honeynet Projesi Türk Chapter Kurucu Üyesi. VizyonArge Ürün Yöneticisi (e-mail: ali.ikinci@vizyonarge.com.tr).

Özet — Teknolojinin gelişmesiyle birlikte internete bağlı cihaz sayısı gün geçtikçe artmaktadır. Günümüzde kişisel, sosyal, sağlık gibi birçok alanda bu cihazların kullanımının yaygınlaşması teknoloji gelişimine bağlı olduğu kadar kullanıcılara sağladığı güvenlik ve mahremiyet yetenekleri ile de ilgilidir. Nesnelerin interneti cihazları bilgisayar, sunucular gibi güçlü donanıma sahip olmadıklarından bu cihazların bünyesinde saldırı tespiti için klasik yöntemler kullanılamamaktadır. Son yıllarda çıkan açıklıklar ve potansiyel kurban sayısının giderek artması saldırganların bu alana yönelmesine neden olmuştur. TR-069, cihazların uzaktan yönetimi için yaygın olarak kullanılan protokollerden biridir. Bu makalede TR-069 protokolünü kullanan nesnelerin interneti cihazlarında saldırı tespiti için tuzak sistem kullanımı ele alınmış ve bu cihazlardan ADSL (Asymmetric Digital Subscriber Line) modem/yönlendiriciler için bir tuzak sistem uygulaması geliştirilmiştir.

Abstract — The number of devices connected to the Internet is increasing day by day with the development of technology. Nowadays, widespread use of these devices in many areas like personal, social, health etc. depends on as well as technology development, it is also related to security and privacy capabilities that provide to users. The conventional Intrusion Detection Systems (IDS) can not be used at internet of thing devices because of limeted hardware (CPU, RAM etc.) and software resources. The vulnerabilities that are found in recent years and the gradual increase in the number of potential victims have led attackers to tend to this field. TR-069, one of the widely used protocols to manage these devices remotely. In this paper, the use of honeypot is presented for intrusion detection on the internet of things devices that use TR-069 protocol and a honeypot application has been developed for IoT.

Anahtar Sözcükler — Nesnelerin interneti, tuzak sistem, TR-069, modem/yönlendirici, RomPager

I. GİRİŞ

İnternetin yaşamımıza girdiği ilk yıllardan itibaren kullanıcı sayısı her geçen gün artmaktadır. Özellikle son yıllarda yaşanan teknolojik gelişmeler ve 1999 yılında ortaya atılan "Nesnelerin İnterneti (Internet of Things-IOT)" kavramı ile birlikte çevremizdeki birçok eşyanın birbirleriyle iletişim kurması, internete bağlanmasına olanak sağlanmıştır [1].

Bu gelişmelerin sosyal yaşamda sağladığı kolaylıkların kullanıcılar arasında hızla yayılması, bu alana daha fazla yatırım yapılması ve dikkate değer bir pazar haline gelmesine neden olmuştur. Ancak farklı türdeki nesnelerin bilgi paylaşımında bulunması kullanıcı gizliliği ve mahremiyeti konusunda çeşitli problemleri beraberinde getirmiştir. Ayrıca son yıllarda farklı türdeki cihazlarda çıkan açıklıklar ve olası açıklık durumunda potansiyel kurban sayısının çok fazla olması saldırganlar için cezbedici bir ortam oluşturmuştur.

Nesnelerin interneti cihazları arasında buzdolabı, su ısıtıcısı, ütü, televizyon vb. olmak üzere günlük yaşamda aktif olarak kullandığımız birçok farklı türde cihaz sayılabilir. Ev ya da küçük ofis kullanıcılarının internete bağlanmak için kullandığı modem/yönlendirici cihazlar bunlardan biridir. Son 10-15 yıllık zaman dilimi ile birlikte artık herkesin evinden internete bağlandığı düşünüldüğünde bu cihaz sayısında da önemli artışlar olmuştur. Bu durum cihazlarla uğraşan saldırgan, araştırmacı sayısının artmasına ve çeşitli açıklıkların ortaya çıkarılmasına neden olmuştur. Günümüzde hâlâ aktif olan bazı açıklıklarda 2004 yılında yayınlanan ve bu türdeki cihazların uzaktan yönetimini sağlayan TR-069 protokolü kullanılmaktadır. Nesnelerin interneti kullanıcılarının çoğunluğunun teknik olarak bilgi sahibi olması beklenmediğinden açıklıkların kapanması için çeşitli yamalar yayınlansa da bunun tüm cihazlara uygulanması ve yama yönetimi zor olmaktadır. Böylece üretici, sağlayıcı firmanın getirdiği çözümler her cihaza uygulanamamakta ve saldırganların hedef alabileceği kurban sayısı önemli ölçüde kalmaya devam etmektedir.

Cihazların fiziksel ve ağ güvenliğinin sağlanmasına yönelik çeşitli çalışmalar yapılmaktadır. Ancak saldırı tespiti noktasında bazı problemler bulunmaktadır. Bunlardan en önemlisi nesnelerin interneti cihazlarının kısıtlı bantgenişliği, hafıza, hesaplama yeteneği ve enerjiye sahip olmasından dolayı üzerlerinde yüksek işlem gücü gerektiren klasik saldırı tespit sistemlerini kullanmanın imkânsız olmasıdır. Tuzak sistemler, bilgi sistemlerine gerçekleştirilen saldırıların tespitinde kullanılan önemli mimarilerden biridir. Temel amacı hedef sistem gibi davranarak saldırganların dikkatini çekmek ve olası saldırı durumunda bütün aktiviteleri kaydetmektir. Bu uygulamalar doğrudan hedef sistem üzerinde çalışmadığından sistemin sahip olduğu donanımsal eksikliklerden etkilenmemektedir. Güncel olarak SMB, HTTP, FTP, SSH gibi birçok protokolün ve çeşitli işletim sistemlerinin benzetimini yapan tuzak sistemler bulunmakta ve saldırı tespit noktasında aktif olarak kullanılmaktadır. Ancak önemli bir hedef haline gelen nesnelerin interneti eşyalarından modem ve yönlendiricilere gelen saldırıların tespiti için geliştirilmiş bir tuzak sistem bulunmamaktadır.

Makalenin 2. bölümünde çalışmanın ve geliştirilen uygulamanın anlaşılmasını sağlamak amaçlı farklı başlıklar altında nesnelerin interneti, tuzak sistemler, TR-069 protokolü ile ilgili temel bilgiler, literatür taraması detaylıca ele alınmıştır. 3. bölümde geliştirilen uygulama sistem tasarımından, test ortamına ve kullanım senaryolarına kadar tanıtılmıştır. Bölüm 4'te ise çalışma ile ilgili sonuç ve ileriye dönük çalışmalar için önerilere yer verilmiştir.

II. TEKNOLOJİLER VE LİTERATÜR TARAMASI

A. Nesnelerin İnterneti

Nesnelerin interneti kavramının terminolojideki tanımı üzerine birçok farklı görüş vardır. Bu farklığın sebebi aslında kavramı oluşturan iki sözcükten gelmektedir. Çeşitli ticari şirketler ve araştırma kurumları kendi altyapılarına, ilgi alanlarına göre ya internet kısmına ya da nesne kısmına

ağırlık vererek tanım oluşturmuşlardır. Ayrıca kavram anlam bilimsel olarak incelendiğinde ortaya çıkan anlamsal tarafı vardır [2]. Böylece tanımlama yapılırken internet, nesne ve anlamsal olmak üzere 3 yaklaşım esas alınmıştır. Genel olarak nesnelerin interneti "çeşitli iletişim protokollerini kullanarak birbirleri ile haberleşen, bilgi üreten, oluşturdukları ağ sayesinde çevresiyle bilgi alış verişi yapabilen akıllı cihazların oluşturduğu bir topluluk ve pazardır.

Nesnelerin interneti kavramı ilk olarak 1999 yılında MIT Auto-ID Center kurucularından olan Kevin Ashton tarafından Procter & Gamble (P&G) şirketinde tedarik zinciri yönetimini konu aldığı bir sunumun başlığı olarak kullanılmıştır [3]. 2005 yılında International Telecommunication Union (ITU) tarafından yayınlanan "ITU Internet Report 2005: Internet of Things" raporu ile birlikte "Nesnelerin İnterneti" kavramı resmi olarak duyurulmuştur [4]. 2009 yılına gelindiğinde Avrupa Birliği "Nesnelerin İnterneti – Avrupa için Eylem Planı" başlıklı bir eylem planı yayınlayarak konuya verdiği önemi göstermiştir [5]. Cisco IBSG (Internet Business Solutions Group) tarafından 2011 yılında yayınlanan rapora göre 2003 yılında 500 milyon cihaz internete bağlı ve kişi başına düşen cihaz sayısı 0,08 iken 2010 yılında cihaz sayısı 12.5 milyara ve kişi başına düşen cihaz sayısı ise 1,84'e çıkmıştır. Yapılan çalışmalar sonucunda her 5 yılda bu oranın 2 katına çıkacağı öngörülmektedir. 2020 yılına gelindiğinde dünya nüfusunun 7.6 milyar, internete bağlı cihaz sayısının 50 milyar olacağı tahmin edilmektedir [6].

Günlük hayat incelenip gelecek öngörüleri düşünüldüğünde buzdolabı, araba, televizyon, su ısıtıcısı, fırın, ütü, kitap, kamera, klima, modem, yönlendirici benzeri akla gelebilecek birçok cihazın kablosuz ağ ya da RFID teknolojisi sayesinde birbirleri ile iletişim kurup internete bağlanarak yaşamımızı kolaylaştıracağı ve bazı alanlarda işleri daha verimli hale getireceği görülmektedir. Nesnelerin interneti kapsamında geliştirilen uygulamalar ağ erişilebilirliği, kapsam, yinelenebilirlik, taşınabilirlik, kullanıcı bağımlılığı ve etkisi türlerine göre sınıflandırılabileceği gibi kullanıldığı alanlar bakımından şu şekilde gruplandırılabilir [7]:

- Akıllı Ortam
- Sağlık Hizmetleri
- Ulaşım ve Lojistik
- Kişisel ve Sosyal
- Enerji ve Madencilik

Nesnelerin interneti çözümlerine gizlilik, bütünlük, erişilebilirlik, kimlik doğrulama, yetkilendirme vb. gibi güvenlik özellikleri entegre edilerek uygulamaların kullanılabilirliği arttırılabilir. SANS enstitüsü tarafından 2013 yılında yapılan ve kamu, askeri, sağlık, eğitim gibi birçok farklı sektörden yaklaşık 400 kurumun katıldığı araştırmaya göre katılımcıların %48,8'i nesnelerin interneti uygulamalarının günümüzde diğer sistemlerde karşılaşılan güvenlik problemleriyle aynı seviyede olduğunu belirtmiştir [8]. Karşılaşılan problemlerin çözümü için öncelikle tehdit kaynaklarının tespiti ve saldırı vektörlerinin belirlenmesi gerekmektedir. Ayrıca potansiyel hedef sayısının çok fazla olduğu bu alanda saldırı tespiti de önemli bir konu haline gelmiştir. Potansiyel tehditler ve saldırılar incelendiğinde kötü niyetli kullanıcı, kötü niyetli üretici ve dış saldırganlar olmak üzere tehditlerin kaynağı 3 farklı grupta toplanabilir [9]. Nesnelerin interneti uygulamalarının kullanım alanları çok farklı olduğundan bu uygulamalara yönelik saldırı vektörleri çeşitlilik göstermektedir. Bu durum saldırganların işini kolaylaştırırken savunma tarafındakiler için ele alınması gereken birçok parametre anlamına gelmektedir. OWASP adlı topluluk tarafından 2014 yılında yayınlanan çalışmayla nesnelerin interneti için 10 saldırı vektörü belirlenmiştir. Bunlar arasında web arayüzleri, ağ servisleri, şifreleme eksikliği ve cihaz yazılımları en önemlilerindendir [10].

Günümüzde bilgi sistemlerine gelen saldırıların tespitinde ağ servisi, işletim sistemi veya tüm ağın benzetimini yapabilen tuzak sistemler ve saldırı tespit sistemleri kullanılmaktadır. Nesnelerin interneti cihazları güçlü donanım özelliklerine sahip olmadığından klasik saldırı tespit sistemlerini kullanmak olanaksızdır. Bununla birlikte nesnelerin interneti uygulamalarında saldırı tespiti için çeşitli akademik çalışmalar yapılmaktadır. Raza S. ve arkadaşları 6LoWPAN ağı için geliştirdikleri ve adını SVELTE olarak belirledikleri saldırı tespit sistemi temel olarak sahte bilgi, seçmeli iletim ya da tuzak yönlendirme (sinkhole) gibi saldırıların tespitini amaçlamaktadır [11]. Yine EC FP7 (European Commission 7th Framework Programme) tarafından desteklenen "ebbits" projesi kapsamında hem kablosuz duyarga ağları hem de internet ağından gelebilecek saldırılara karşı savunmasız olan 6LoWPAN cihazları için saldırı tespit sistemi çalışma yapısı önerilmiştir [12]. Literatürde nesnelerin interneti uygulamalarına gelebilecek saldırıların tespiti için benzer çalışmalar yürütülse de cihazların çalışması veya yönetiminde kullanılan herhangi bir protokol için geliştirilmiş bir tuzak sistem çalışması bulunmamaktadır.

B. Tuzak Sistemler

Tuzak sistem (honeypot - bal küpü) bilgi sistemlerine gerçekleştirilen saldırıları tespit etmek amaçlı geliştirilen mimarilerden biridir. Hedef sistem gibi davranarak saldırganların dikkatini çekmek ve olası saldırı durumunda bütün aktiviteleri kaydetmek üzere tasarlanmıştır. Tuzak sistemlerin temel özellikleri arasında ağ servislerinin, işletim sistemlerinin ya da tüm ağın benzetimini yaparak saldırıları üzerine çekmek, zararlı yazılım örneklerini toplamak, saldırı yönteminin özellikleri ve tekniği hakkında bilgi sağlamak, gerçek sistemlere gelebilecek potansiyel saldırı riskini düşürmek sayılabilir [13]. Tuzak sistemler gerçek bir ağa ait gibi görünse de ele geçirilmesi durumunda gerçek sistemlerin etkilenmesini engellemek amaçlı izole edilmiş bir ağ ortamında çalışırlar. Ayrıca tuzak sistemlerin sahip olduğu IP adresleri duyurulmamış yani herhangi bir yere kaydettirilmemiş, herhangi bir adresle ilişkilendirilmemiş olduğundan kendisine gelen tüm trafik şüpheli olarak düşünülür.

Tuzak sistemler kullanım amacı, üstlendikleri rol, geliştirildikleri donanım türü ve saldırgan ile olan etkileşimlerine göre çeşitli gruplara ayrılırlar. Saldırgan ile olan etkileşimlerine göre tuzak sistemler düşük etkileşimli, orta etkileşimli ve yüksek etkileşimli olmak üzere 3'e ayrılır [14]. Etkileşim saldırganın tuzak sistemle gerçekleştirdiği aktivitelerle ölçülür. Hangi tuzak sistemin ne zaman kullanılacağı çeşitli faktörlere bağlıdır. Etmenler ve tuzak sistem türlerinin bunlarla ilişkisi Tablo 1'de verilmiştir.

Etmenler	Düşük Etkileşimli	Orta Etkileşimli	Yüksek Etkileşimli
Bulaşma Derecesi	Düşük	Orta	Yüksek
Gerçek işletim sistemi	Yok	Yok	Var
Kurulum	Kolay	Zor	Çok zor
Bakım	Kolay	Kolay	Zaman alıcı
Risk	Düşük	Orta	Yüksek
Ele geçirilme beklentisi	Yok	Yok	Var
Kontrol gereksinimi	Yok	Yok	Var
Çalıştırmak için gerekli bilgi	Düşük	Düşük	Yüksek
Geliştirmek için gerekli bilgi	Düşük	Yüksek	Orta-Yüksek
Veri Toplama	Kısıtlı	Orta	Kapsamlı
Etkileşimli	Servis benzetimi	İsteklere göre	Tam kontrol

Tablo I - Saldırgan ile olan etkileşimlerine göre tuzak sistemlerin karşılaştırılması

Düşük etkileşimli tuzak sistemler herhangi bir servisin ya da komple bir işletim sisteminin benzetimini yaparlar. Fakat servisler kullanılarak sistem ele geçirilemez. Orta etkileşimli tuzak sistemler, düşük etkileşimli tuzak sistemler gibi gerçek bir işletim sistemine sahip değildir. Ancak saldırgan ile daha çok etkileşime geçebilmesi ve daha karmaşık saldırıları üzerine çekebilmesi yönüyle düşük etkileşimli tuzak sistemlerden farklıdır. Yüksek etkileşimli tuzak sistemler saldırgan ile olan etkileşimi en yüksek olan tuzak sistemlerdir. Herhangi bir servisin benzetimini yapmak yerine gerçek işletim sistemleri üzerinde açıklık barındıran gerçek ağ servisleri sunarlar.

Tuzak sistem kavramının 1990 yılında Clifford Stoll'un "The Cuckoos Egg" ve Bill Cheswick'in "An Evening with Berferd" yayınlarıyla bilgi güvenliğinde kullanılmaya başlanmasıyla birlikte geliştirilen ve günümüzde aktif olarak kullanılan bazı önemli, açık kaynak kodlu, farklı türdeki tuzak sistemler arasında honeyd, dionaea, kippo, conpot, glastopf, thug sayılabilir [15]. Honeyd, dionaea ve kippo SMB, HTTP, FTP, TFTP, MSSQL ve SSH vb. protokollerin, conpot endüstriyel kontrol sistemlerin kullandığı protokollerin benzetimini yaparken, glastopf web uygulamaları ile ilgili açıklıkların benzetimini yapmaktadır. Thug ise istemci taraflı bir tuzak sistemdir. Bununla birlikte ADSL modem ve yönlendirici cihazlarda kullanılan TR-069 protokolü ile ilgili geliştirilmiş herhangi bir tuzak sistem türü bulunmamaktadır.

C. TR-069

TR-069 (Technical Report 069), Broadband Forum tarafından Mayıs 2004'te yayınlanmış ve CWMP (CPE WAN Management Protocol - Müşteri Tarafı Cihazı Geniş Alan Ağı Yönetim Protokolü) olarak adlandırılan teknik raporun kısa adıdır. İnternete bağlı modem, yönlendirici, ağ tabanlı depolama aygıtları, VoIP telefonlar vb. son kullanıcı cihazlarının uzaktan yönetimi için uygulama seviyesi protokolü tanımlar. Metin tabanlı çalışan TR-069 protokolünde mesajlar ACS (Auto Configuration Server - Otomatik Yapılandırma Sunucusu) ve CPE (Customer Premises Equipment - Müşteri Tarafı Cihazı) arasında transfer edilir. ACS genellikle internet servis sağlayıcı ya da kullanılan cihazı tedarik eden kurum tarafında bulunan sunucu iken CPE son kullanıcı tarafındaki yönlendirici, VoIP telefon gibi herhangi bir cihazdır. Nesnelerin interneti kullanımının yaygınlaşmasıyla protokolü kullanan cihaz sayısının giderek artacağı öngörülmektedir.

TR-069 protokolünün temel kullanım amaçları arasında otomatik yapılandırma, dinamik hizmet tedariki, aygıt yazılımı ve modül yönetimi, durum ve performans izleme, hata tanımlama sayılabilir. Protokol çift yönlü olarak SOAP/HTTP (Simple Object Access Protocol/Hypertext Transfer Protocol) üzerinde çalışmaktadır. Mesajlar XML (Extensible Markup Language) formatında RPC (Remote Procedure Call) yöntemiyle taraflara iletilmektedir [16].

CPE / ACS Yönetim Uygulaması
RPC Metodları
SOAP
HTTP
SSL / TLS
TCP / IP

Tablo II - Tr-069 protokol yığını

ACS ile CPE arasında oturumun kurulması aşamasında oturum her zaman CPE tarafından başlatılır. Oturumun başlatılmasında iki farklı senaryo vardır [17]. İlk olarak CPE herhangi bir nedenle ACS'ye bağlanabilir ki bu durumda CPE istemci iken ACS sunucu durumundadır. İkinci senaryoya göre ise ACS ilk yapılandırma, değişiklik, yazılım güncelleme vb. bir amaçla CPE'ye kendisine bağlanması için istek gönderir. CPE isteği işleyerek ACS'ye bağlantı kurar. Bu durumda CPE sunucu durumundayken ACS istemci durumuna geçmiştir. Ayrıca ikinci durum için cihaz üzerinde açık bir port bulunması gerekir. Bu port aynı zamanda saldırganlar için açık bir kapı anlamına gelmektedir.

ACS ve CPE iletişiminde kullanılan önemli bazı komutlar şunlardır [18]:

- **Inform:** CPE'den ACS'ye her oturum öncesi gönderilen komuttur. Oturum nedenini içerir.
- **GetRPCMethods:** CPE ya da ACS'nin desteklediği komut listesini öğrenmek amaçlı kullanılır.
- **GetParameterNames:** CPE'nin desteklediği parametrelerin listesini almak için kullanılır.
- **GetParameterValues:** İstenilen bir ya da daha fazla parametrenin güncel değerini döner.
- **SetParameterValues**: Bir veya birden fazla parametrenin değerini değiştirir.
- **Download:** CPE'ye belirtilen bir URL'den aygıt yazılımı, yapılandırma dosyası vb. bir dosyanın indirilmesini için kullanılır.
- **Reboot:** CPE'nin kapanıp açılmasını sağlar.

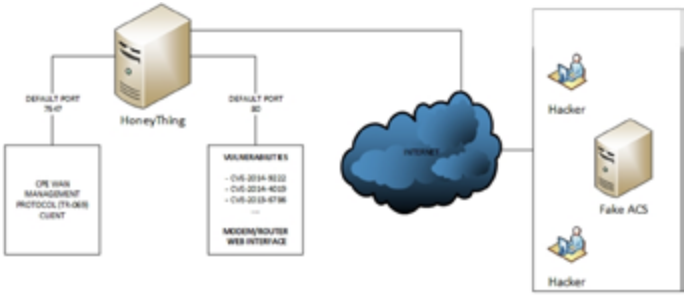
TR-069 protokolü HTTP basic, HTTP digest ya da sertifika tabanlı olmak üzere çift yönlü kimlik doğrulamayı gerektirir. CPE ACS'yi bağlantı isteğinde doğrularken, ACS CPE'yi oturumun başlatılması sırasında doğrular [19]. İletişimde HTTPS kullanılarak olası aradaki adam saldırılarının önüne geçmek hedeflenmiştir. Ancak protokolün işletimini sağlayan ACS/CPE üzerindeki uygulamalarda çeşitli açıklıklar bulunmaktadır. Ayrıca CPE'ler yönetimi için web uygulamalarına sahip olduğundan web tabanlı saldırılara karşı da hedef halindedir.

III. HONEYTHING

Bu bölümde nesnelerin interneti cihazlarından modem ve yönlendiricilere gelen saldırıların tespiti için geliştirilmiş düşük etkileşimli tuzak sistem olan HoneyThing, tasarım ve geliştirme, kullanım önerileri ve test başlıkları altında detaylıca incelenmiştir.

A. Tasarım ve Geliştirme

HoneyThing'in temel amacı modem ve yönlendiriciler için son yıllarda çıkmış popüler bazı açıklıklara karşı savunmasız, TR-069 protokolünü destekleyen bir sistem sunmak ve uygulama ile olan tüm etkileşimlerin detaylı bir şekilde kaydını tutmaktır. Bu çalışma kapsamında geliştirilen uygulama iki bölümden oluşmaktadır. Birinci kısımda bazı açıklıkların benzetimi yapılarak bir modem web arayüzü sunulmuştur. İkinci kısımda ise TR-069 protokolünün istemci tarafı komutlarının işletilmesini sağlayan uygulama HoneyThing'e entegre edilerek birlikte çalışması sağlanmıştır.

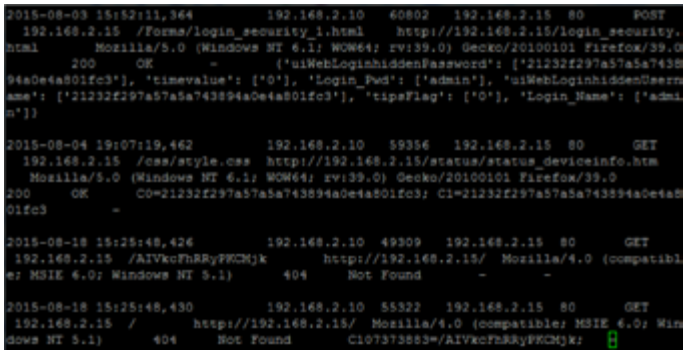


Şekil 1. HoneyThing tuzak sisteminin yapısı

İlk bölümün geliştirilmesi amaçlı farklı marka ve modelden birçok cihaz için açıklıklar araştırılmış ve günümüzde yaygın olarak kullanılan 3 açıklık tespit edilmiştir. Bu açıklıkların ortak yönü Allegro firması tarafından geliştirilen gömülü web sunucusu RomPager uygulamasında çalışmasıdır. TR-069 protokolünün çalıştığı sunucuların %52'sini oluşturan RomPager ve %52 içerisinde %98'lik dağılıma sahip 4.07 versiyonu, günümüzde halen yaklaşık 12 milyon cihaz tarafından kullanılmaktadır [20]. Açıklıklardan en önemlisi Aralık 2014 tarihinde Check Point firması araştırmacıları tarafından bulunan ve cihaz üzerinde yönetici hakkı elde etmeyi sağlayan "Misfortune Cookie"'dir (CVE-2014-9222) [21]. İstenilen yetkiyi elde eden saldırgan, DNS ayarlarını değiştirerek kullanıcı trafiği arasına girebilir, port yönlendirme ile modeme bağlı cihazlara erişebilir ve hassas kullanıcı verilerini ele geçirebilir. ROM-0 (CVE-2014-4019) açıklığında ise saldırgan cihaza ait yapılandırma bilgilerini içeren yedek (backup) dosyasını yetkilendirilmesi yapılmamış bir URL üzerinden indirebilmekte ve çeşitli yöntemlerle bu bilgilere ulaşabilmektedir [22]. Saldırganın sunucu üzerinde olmayan bir URL'ye gönderdiği özel istek sayesinde URL yönlendirme ve siteler arası betik çalıştırmayı sağlayan CVE-2013-6786 bir diğer önemli açıklıktır [23]. Sisteme benzer şekilde yeni açıklıklar eklenebilir. Sonuç olarak birinci bölüm, farklı açıklıklara sahip RomPager web sunucusunun benzetimini yapmakta ve kullanıcının giriş yapıp çeşitli sayfaları görüntüleyebildiği bir web uygulaması sunmaktadır.

İkinci bölümde TR-069 protokolünün istemci tarafını gerçekleyen çeşitli uygulamalar araştırılmış ve Google çalışanları tarafından geliştirilmiş, açık kaynak kodlu "Catawampus" uygulaması HoneyThing'e entegre edilmiştir [24]. Bu kısımda hedef, TR-069 protokolü kullanılarak yapılabilecek bilinmeyen saldırıların tespiti, saldırgan davranışının kayıt altına alınması ve bu protokole yönelik saldırı miktarı vb. istatistiklerin çıkarılmasıdır.

HoneyThing 3 adet kayıt dosyası tutmaktadır. "http.log", benzetimi yapılan web sunucusunun HTTP iletişimi ile ilgili kayıtları, "tr-069.log" ise TR-069 protokolü haberleşmesi ile ilgili kayıtları tutmaktadır. "honeything.log" dosyasında uygulamanın içsel hata ve bilgilere ait kayıtlar yer almaktadır. Tüm kayıtlar ayrıştırılmasını kolaylaştırmak amaçlı "tab" karakteriyle ayrılmış olarak yazılmakta ve dosyalar metin belgesi formatında tutulmaktadır. Bu format ile gelecekte saldırılar veri tabanlarına aktarılarak detaylı analizler yapılabilecektir.



Şekil 2. HoneyThing HTTP kayıt dosyası içeriği

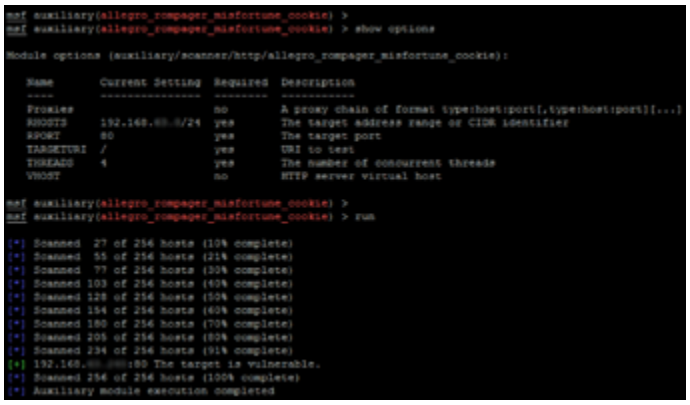
Uygulama Python programlama dili ile geliştirilmiştir. Bu nedenle taşınabilir ve Python çalıştıran herhangi bir işletim sistemi üzerine kolayca kurulumu yapılabilmektedir. Kurulum sonrası port bilgileri, web uygulaması yetkilendirme bilgileri ve modem/yönlendirici cihaza ait çeşitli bilgiler yapılandırma dosyası aracılığıyla değiştirilebilmektedir. Varsayılan olarak uygulama tüm ağ arayüzlerini dinlemekte, HTTP için 80, TR-069 için varsayılan bağlantı isteği portu olan 7547'yi kullanmaktadır.

B. Kullanım Önerileri

HoneyThing düşük etkileşimli bir tuzak sistem de olsa izole bir ağ ortamında kullanılması tavsiye edilmektedir. Örneğin TR-069'un "Download" komutu ile sisteme indirilecek herhangi bir zararlı yazılımın çalıştırılması ağdaki diğer makinalara zarar verebilir. Tuzak sistemin farklı lokasyonlarda çalıştırılması durumunda o ülkeye ait ISP'lerin sunduğu modem/yönlendirici cihazlarında bağlantı isteği için kullanılan port'un TR-069 portu olarak ayarlanması sistemin kullanılabilirliğini arttıracaktır. Ayrıca çoklu kullanımda saklanan kayıt dosyaları ayrıştırılabilir formatta olduğundan merkezi bir kayıt sunucuda toplanarak Kibana, Splunk benzeri açık kaynak kodlu bir uygulama ile izlenebilir.

C. Test

HoneyThing'in birinci kısmına ait açıklıklar manuel olarak ve Metasploit sızma testi aracı ile test edilmiştir. Metasploit aracı üzerinde bulunan tarayıcılardan "allegro_rompager_misfortune_cookie" modülü ile HoneyThing'in bulunduğu ağ taranmış ve HoneyThing için "Vulnerable" sonucu döndüğü gözlemlenmiştir [25].



Şekil 3. HoneyThing tuzak sisteminin Metasploit uygulamasıyla test edilmesi

İkinci kısımda ise TR-069 protokolünün çalışması için gerekli olan ACS "VMware Workstation" üzerinde sanal olarak hazırlanmıştır. ACS için OpenACS'nin devamı olan LibreACS adlı açık kaynak kodlu uygulaması kullanılmıştır. ACS (LibreACS) ve CPE (HoneyThing) çalıştırıldığında protokolün başarılı bir şekilde gerçeklendiği ve iletişimin kayıt altına alındığı izlenmiştir [26].

IV. SONUÇ VE ÖNERİLER

Nesnelerin interneti kavramının giderek önem kazandığı günümüzde bu kavram kapsamına giren cihazlara gelen saldırılarda da önemli artışlar olmuştur. Donanımsal olarak yetersiz olan bu cihazlarda klasik saldırı tespit sistemleri kullanılamamaktadır. Saldırı tespiti için önemli bir yöntem olan tuzak sistemler günümüzde birçok protokol için geliştirilmiş olsa da nesnelerin interneti cihazlarından TR-069 protokolünü kullanan modem/yönlendiriciler için geliştirilmiş bir sistem bulunmamaktadır. HoneyThing bu eksikliği gidermekle birlikte, bu çalışma farklı türde birçok cihazı içine alan nesnelerin interneti cihazlarında saldırı tespiti için tuzak sistem kullanımını önermektedir.

Nesnelerin interneti için geliştirilecek tuzak sistem uygulamaları normal tuzak sistem uygulamalarından farklı olarak sadece protokolün benzetimini yapmak yerine cihaza özel özellikleri de yansıtması gerekmektedir. Cihazların kullandığı port, komut seti ve benzeri bilgiler tedarikçi firmaya göre değişeceğinden tuzak sistem, hedef alınan kapsama göre yapılandırılabilir olmalıdır.
Geliştirilmesi devam eden HoneyThing'e yeni açıklık modülleri eklenebileceği gibi, kayıtların veritabanı, syslog vb. yerlere yazılması, Honeynet topluluğun veri besleme protokolü olan "hpfeeds"'in desteklenmesi, saldırganın komut satırına düşmesi durumunda belli başlı bazı kabuk komutlarının benzetiminin yapılması benzeri birçok özellik eklenerek daha verimli bir kullanım hedeflenmektedir.

KAYNAKÇA

[1] Y. Liu, G. Zhou, "Key Technologies and Applications of Internet of Things", Fifth International Conference on Intelligent Computation Technology and Automation (ICICTA), p. 197-200, 2012.

[2] L. Atzori, A. Iera, G. Morabito, "The Internet of Things: A survey", The International Journal of Computer and Telecommunications Networking vol.54,p. 2787-2805, 2010.

[3] K. Ashton, "That 'Internet of Things' Thing", RFID Journal, 2009 [Online], Avaliable: http://www.rfidjournal.com/articles/pdf?4986

[4] The Internet of Things, International Telecommunication Union, November 2005 [Online], Avaliable: http://www.itu.int/osg/spu/publications/internetofthings/InternetofThings_summary.pdf

[5] Internet of Things - An action plan for Europe, Commission Of The European Communities, Brussels, 278 final, 18.6.2009 COM(2009) [Online], Avaliable: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52009DC0278&from=EN

[6] D. Evans, "The Internet of Things How the Next Evolution of the Internet Is Changing Everything", Cisco IBSG, 2011 [Online], Avaliable: http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

[7] A. Gluhak, S. Krco, M. Nati, D. Pfisterer, N. Mitton, T. Razafindralambo, "A Survey on Facilities for Experimental Internet of Things Research", Communications Magazine, IEEE vol. 49, p. 58–67, 2011.

[8] J. Pescatore, G. Shpantzer, "Securing the 'Internet of Things' Survey", A SANS Analyst Survey, 2014 [Online], Avaliable: https://www.sans.org/reading-room/whitepapers/analyst/securing-internet-things-survey-34785

[9] A. Atamli, A. Martin, "Threat-Based Security Analysis for the Internet of Things", International Workshop on Secure Internet of Things (SIoT), p. 35-43, 2014.

[10] OWASP Internet of Things Top Ten Project, 2014 [Online], Avaliable: https://www.owasp.org/index.php/OWASP_Internet_of_Things_Top_Ten_Project

[11] S. Raza, L. Wallgren, T. Voigt, "SVELTE: Real-time intrusion detection in the Internet of Things", Journal Ad Hoc Networks, vol.11, Issue 8, p. 2661–2674, 2013.

[12] P. Kasinathan, G. Costamagna, H. Khaleel, C. Pastrone, M. A. Spirito, "An IDS Framework for Internet of Things Empowered by 6LoWPAN", 20th ACM Conference on Computer and Communications Security (CCS), 2013.

[13] Deliverable D5.3: Case Study: Malicious Activity in the Turkish Network, Information & Communication Technologies Trustworthy ICT, Seventh Framework Programme, SysSec, February 2013 [Online], Avaliable: http://www.syssec-project.eu/m/page-media/3/syssec-d5.3-TurkishNetworkCaseStudy.pdf

[14] Abhishek Mairh, "Honeypot in Network Security: A Survey", Department of Computer Sc. Engg. International Institute of Information, p. 600-605, 2005.

[15] R. C. Joshi (Editor), Anjali Sardana (Editor), "Honeypots: A New Paradigm to Information Security", ISBN-13: 978-1578087082, ISBN-10: 1578087082, p. 1-6, 2011.

[16] L. Zheng, Y. Hu, S. Chen, "Research and Application of CWMP in Distributed Network Management System", International Conference on Computer Science and Service System (CSSS), p. 647-650, 2012.

[17] JPM. Rojas, "Split Management of TR069 Enabled CPE Devices", Master of Science Thesis, POLITECNICO DI TORINO, 2011 [Online], Avaliable: http://repository.javeriana.edu.co/bitstream/10554/7075/1/tesis537.pdf

[18] TR-069 CPE WAN Management Protocol, Issue: 1 Amendment 5, November 2013 [Online], Avaliable: https://www.broadband-forum.org/technical/download/TR-069_Amendment-5.pdf

[19] J. Walls, T. Sheehan, "TR-069 - A Crash Course", Interoperability Laboratory, University of New Hampshire, 2009 [Online], Avaliavble: https://www.iol.unh.edu/sites/default/files/knowledgebase/hnc/TR-069_Crash_Course.pdf

[20] S. Tal, L. Oppenheim, "The Internet of TR-069 Things: One Exploit to Rule Them All", RSA Conference, 2015 [Online], Avaliable: https://www.rsaconference.com/writable/presentations/file_upload/hta-r04-the-internet-of-tr-069-things-one-exploit-to-rule-them-all_final.pdf

[21] Check Point's Malware and Vulnerability Research Group, Misfortune Cookie Vulnerability, 2014 [Online], Avaliable: http://mis.fortunecook.ie/

[22] T. Hlaváček, "Impact of 'rom-0' vulnerability", 2014 [Online], Avaliable: https://ripe69.ripe.net/presentations/61-rom0-vuln.pdf

[23] A. V. Blanco, CVE-2013-6786, 2013 [Online], Avaliable: http://osvdb.org/ref/99/rompager407.pdf

[24] Catawampus, TR-069 management for a CPE device in Python, 2012 [Online], Avaliable: https://code.google.com/p/catawampus

[25] J. Hart, L. Oppenheim, "Allegro Software RomPager 'Misfortune Cookie' (CVE-2014-9222) Scanner", 2015 [Online], Avaliable: http://www.rapid7.com/db/modules/auxiliary/scanner/http/allegro_rompager_misfortune_cookie

[26] LibreACS, The fork of OpenACS a still open source TR-069 CWMP server, 2015 [Online], Avaliable: http://sourceforge.net/projects/libreacs

Ö. Erdem, 2012 yılında İstanbul Ticaret Üniversitesi Bilgisayar Mühendisliği Bölümünden mezun oldu. 2013 yılında başladığı İstanbul Şehir Üniversitesi Bilgi Güvenliği Mühendisliği yüksek lisansı tez aşamasındadır. 2012 yılından bu yana TÜBİTAK BİLGEM Siber Güvenlik Enstitüsü bünyesinde araştırmacı olarak görev yapmaktadır. Saldırı tespit ve önleme sistemleri, ağ güvenliği, unix/linux sistemler, web teknolojileri konularında çalışmış olmakla birlikte çeşitli açık kaynak kodlu yazılımlar hakkında tecrübe sahibidir.

M. Kara, 1993 yılında Yıldız Teknik Üniversitesi Elektronik ve Haberleşme Mühendisliği Bölümünden mezun oldu. 1996 yılında Yüksek Lisansını, 2002 yılında da Doktorasını Kocaeli Üniversitesi Fen Bilimleri Enstitüsü Elektronik ve Haberleşme Anabilim dalında tamamladı. Doktora tezini Bilgisayar Ağlarında Çok Yollu Yönlendirme konusunda yaptı. 1994-2000 yılları arasında Kocaeli Üniversitesi Bilgisayar Mühendisliği bölümünde Araştırma Görevlisi olarak, 2000-2001 yıllarında Armada Bilgisayar AŞ'de Sistem Mühendisi olarak çalıştı. 2001'den beri TÜBİTAK BİLGEM'de çalışmaktadır. Bulanık mantık, siber güvenlik, ağ ve sistem, protokol güvenlik analizi, kritik altyapı güvenliği, güvenli yazılım geliştirme, Ortak Kriterler, sistem, yazılım/donanım güvenlik testleri konularında çalışmaktadır. Ulusal ve uluslararası dergi ve konferanslarda yayınları bulunmaktadır.

A. İkinci, 2007 yılında Mannheim Üniversitesinden Bilgisayar Yüksek Mühendisi olarak mezun oldu. 2006 yılından bu yana Siber Güvenlik alanında çalışmalar yapmakta ve özellikte tuzak sistemler konusuyla ilgilenmektedir. 2012 yılında Honeynet Projesinin Türkiye bölümünü kurmuştur. 2007'beri zararlı yazılım analizi, sandboxing ve zararlı içerik barındıran web siteleri konularında çalışmaktadır.

SİTELER ARASI KOMUT DİZİSİ (XSS) VE SQL ENJEKSİYONU SALDIRILARINA KARŞI GÜVENLİK ÖNLEMLERİNİN İNCELENMESİ

Halil İbrahim ULUS, Mehmet DEMİRCİ
Gazi Üniversitesi Bilgisayar Mühendisliği Bölümü

Özet — Bilgi sistemlerinin dış dünyaya açılan yüzü web uygulamalarıdır. Özellikle Web 2.0'ın ortaya çıkışı ile birlikte kullanıcı etkileşiminin yanı sıra sistemlerin güvenlik riskleri de artmıştır. Kötü niyetli saldırganlar web uygulamalarını istismar ederek sistemlere giriş yaptıktan sonra sunuculara ve ağın tamamına sıçrama yapabilmektedirler. Sonuçta yüksek yetkilere sahip olarak bilginin gizliliğine, bütünlüğüne ve erişilebilirliğine zarar verebilmektedirler. Bu çalışmada web uygulamalarına karşı yapılan saldırılardan yoğun bir şekilde kullanılan Siteler Arası Komut Dizisi (XSS) ve SQL Enjeksiyonu saldırıları hakkında literatürdeki çalışmalar esas alınarak bilgi verilmiş, bu saldırıların örnek senaryoları gerçekleştirilerek log kayıtları incelenmiş, bunun sonucunda çözüm önerileri ortaya konmuştur.

Anahtar Kelimeler — XSS, Siteler Arası Komut Dizisi, SQL Enjeksiyonu, Web Güvenliği

Abstract — Web applications are the windows of information systems to the outside. The advent of Web 2.0 has led to better user interaction for web applications, as well as more serious security risks. Malicious attackers may exploit the vulnerabilities in web applications to gain an entry point to systems and then control the entire network. In the end, they can compromise the confidentiality, integrity and availability of information. In this study, we present literature-based information about Cross Site Scripting (XSS) and SQL Injection attacks, realize example scenarios of these attacks and collect information by analyzing log records. In addition, we discuss defense methods against these attacks and list our suggestions for stronger defense.

Index Terms — XSS, Cross-Site Scripting, SQL Injection, Web Security

I. GİRİŞ

İnternet servislerinin ve web uygulamalarının kullanımı gün geçtikçe artmaktadır. Diğer taraftan, siber tehditlerin sayısı da web uygulamalarının sayısının artması ve bunların güvenliğine zarar vermenin giderek daha az teknik bilgi gerektirmesi nedeniyle hızla artmaktadır. Web uygulamalarının açıklıklarını sömürmek ve gizli bilgilere ulaşmak, büyük kazançlar sağlayabilecekleri için saldırganların dikkatini çekebilmektedir. Web uygulamalarına yapılan saldırılar verinin/sistemin gizlilik, bütünlük ve erişilebilirlik özellikleri başta olmak üzere çeşitli güvenlik gereksinimlerini hedef almaktadır. Bu saldırıları durdurmak için açıklık tarayıcıları, saldırı tespit sistemlerini, şifreleme cihazlarını ve güvenlik duvarlarını içeren birçok güvenlik mekanizması kullanılarak çok büyük çaba harcanmaktadır.
Web 2.0'ın ortaya çıkışı ile birlikte her ne kadar kullanıcı ile etkileşim sağlansa da aynı zamanda sistemlere kullanıcıdan veri girişine imkan tanındı. Günümüzde saldırganlar tarafından bu imkan kötü niyetli olarak istismar edilmeye başlanmıştır. Web uygulama seviyesinde kontrol edilmemiş girdilerin onaylanması saldırıların ana kaynağını oluşturmaktadır. Açık Web Uygulama Güvenliği Projesine (OWASP) göre 10 açıklıktan dördü geçersiz girdi denetimi ile ilgilidir. OWASP Top 10 raporunda 2008 yılından bu yana ilk üç saldırı hep form alanlarının istismar edilmesi sonucu oluşan saldırılardır. Geçersiz girdilerden web uygulamalarını koruma yetersizliği organizasyonlar için hem maddi hem prestij yönünden pahalıya mal olabilmektedir [2].

	Sıralama			
	2004	2007	2010	2013
Enjeksiyon Saldırıları	6	2	1	1
Siteler Arası Komut Dizisi (XSS)	4	1	2	3
Oturum Çalma ve Oturum Çalma	3	7	3	2
Güvensiz doğrudan nesne erişimi		4	4	4
Siteler Arası İstek Sahteciliği(CSRF)		5	5	5
Güvenlik Ayarları Hataları	10		6	
Güvensiz Kriptografik Depolama	8	8	7	
Sayfaların Erişimini Kısıtlama Hataları		10	8	8
Yetersiz Taşıma Katmanı Koruması		9	9	
Kontrol Edilmeyen Tekrar Yönlendirmeler ve İletmeler			10	10
Hassas Veriyi Açıkta Bırakma				6
İşlev Seviyesi Erişim Kontrolü Eksikliği				7
Bilinen Güvensiz Araçların İstimarı				9

Tablo I - Owasp raporuna göre web uygulamalarına yapılan saldırılar [1]

Bu çalışmada web uygulamalarına yapılan saldırılardan en çok karşılaşılan Siteler Arası Komut Dizisi (XSS) ve SQL enjeksiyonu saldırıları üzerinde durulmuştur. Bu iki saldırı hakkında ikinci bölümde literatürde yapılan çalışmalara değinilmiş, üçüncü ve dördüncü bölümde saldırılar hakkında bilgi verilmiş ve genel çözüm önerileri listelenmiştir. Beşinci bölümde DVWA uygulaması kullanılarak; ilk kısımda örnek SQL enjeksiyonu saldırısı, ikinci kısımda XSS saldırısı yapılarak; bu saldırılar analiz edilmiştir. Ayrıca bu kısımlarda yapılan çalışmaların çoğunda ele anılan Snort programı için oluşturulan kuralların web saldırılarına karşı ne kadar etkili olduğu denenerek önleyici ve güvenli yazılım ile tarayıcı seviyesinde kullanılan güvenlik çözüm önerilerinin etkisi incelenmiştir. Literatürdeki çalışmalarda genellikle farkındalık oluşturulması, web uygulamalarında güvenlik duvarı kullanılması, güçlü şifreleme ve protokol kullanılması, açıklıkların önceden tespiti, sunucu veya tarayıcı tarafından içerik ve uzantı kontrolü gibi çözüm önerileri sunulmuştur. Yaptığımız çalışmada tartışma ve sonuç kısmında, uygulama sonrasında bu saldırılara karşı güvenli yazılım ve tarayıcı seviyesinde geliştirilen çözüm önerilerinin daha faydalı olacağı vurgulanarak elde edilen kazanımlar neticesinde özellikle protokoller ve algoritmalar açısından güvenlik önerileri sunulmuştur.

II. LİTERATÜR TARAMASI

Web uygulama güvenliği alanında yapılan çalışmalar sunucu ve kullanıcı tarafındaki çözüm önerileri olmak üzere ikiye ayrılmaktadır. Ayrıca yapılan çalışmaların çoğu imza tabanlı

güvenlik sistemleri için kural oluşturmak yönündedir. Bu bölümde daha çok anormallik tespitine dayanan çalışmalar üzerinde durulacaktır.

Cova ve diğerleri yaptıkları çalışmada web uygulamalarına yapılan saldırılarda özgün bir yaklaşım sunan anormallik tespitine dayanan Swaddler'i önermişlerdir. Swaddler, web uygulamalarının iç durumunu analiz eder ve uygulamaların kritik durumu ile normal durumu öğrenerek, oluşan farklılıklara göre alarm üretir. Bu sayede Swaddler uygulamaları uygunsuz ve anormal duruma getiren uygulamaları tespit eder [3].

Kruegel ve Vigna, web sunucularına ve web uygulamalarına yapılan saldırıların tespitinde birçok anormallik tespit tekniği kullanan saldırı tespit sistemi önermişlerdir. Sistem sunucu taraflı, istemci sorguları tarafından kullanılan programlardan ve sorgularda yer alan parametrelerden yardım alır. Sistem otomatik olarak profillerle ilgili Web uygulama verilerini analiz eder [4].

Vigna ve arkadaşları, zararlı davranışları bulmak için web isteklerini analiz eden ve çok katmanlı saldırıların tespiti için karmaşık dil kullanan saldırı tespit sistemi webSTAT'ı önermiştir. WebSTAT çift taraflı olarak hem ağdaki hem işletim sistemindeki olayları sunucu loglarını kullanarak ve ilişkilendirerek analiz eder [5].

Auxilia ve Tamilselvan, web uygulamalarının kötüye kullanımına dayalı negatif bir güvenlik modeli önermişlerdir. Bu negatif güvenlik modeli kurallardan oluşan her web mimarisi arasında sorgulayıcı bir yöntemle web uygulama güvenlik duvarı sağlar. Bu güvenlik duvarı dış korumayı artırmak, web uygulamalarına ulaşmadan atakları tespit etmek ve önlemek maksadıyla konuşlandırılır [6].

Vigna ve diğerlerinin yaptığı bir başka çalışmada bir web tabanlı anormallik tespit sistemi, bir tersine HTTP proxysi ve bir anormallik veritabanı tespit sisteminin birleştirilmiş bir sistemi önerilmiştir. Kullanılan web tabanlı anormallik tespiti ve bir SQL sorgu anormallik detektörü kullanılması tespit oranını artırmıştır [7].

Krueger, Vigna ve Robertson web sunucularına ve uygulamalarına yapılan saldırıları tespit etmek için farklı saldırı tespit teknikleri kullanan bir saldırı tespit sistemi önermişlerdir. Sistem istemci sorgularını analiz ederek web uygulamaları ile arasındaki ilişkiyi inceler [8].

Sekar, sunucuya yapılan istekleri ve verilen cevapları yakalayarak çalışan yeni bir teknik önermiştir. Bu teknikte cookie, SQL sorguları, HTTP istekleri, cevapları gibi kaynaklardan elde edilen bilgilerin hem sözdizimsel açıdan (syntax) hem içerik yönünden belirlenen kurallara göre kontrolü yapılır. Bu kontrolde isteğin şeklinin yazılım kuralları açısıında uygunluğu kontrol edilir ve aynı zamanda uygulamanın sıradışı hareketleri gözlenir [9].

XSS açıklıklarının istismar edilmesi kolay fakat tespit edilmesi ve önlenmesi zordur. Selvamani, Duraisamy ve Kannan'ın çalışmasında istemci tarafındaki çözüm önerileri sunulur. Bir web proksisi kullanarak XSS saldırısını önleyici kurallar geliştirilmiştir. Bu çalışmada tarayıcı tarafında olabilecek özel bir XSS saldırısı üzerinde durulmuştur. Zararlı kod web uygulamasına enjekte edilir ve sitenin açıklığı istismar edilerek kullanıcının gizli bilgilerine ulaşılabilir. Bu çalışma kişisel web güvenlik duvarı sunarak istemci taraflı bir çözüm öneri sunmaktadır. Bu çözümün olumlu tarafı web uygulama sağlayıcılarına güven duymaya gerek kalmaz. Bu da oturum id'leri ve cookieler gibi gizli bilgilerin hedef olmasını engeller [10].

Saldırı tespit sistemleri genelde saldırıların tespitinde saldırı imzası ve genel saldırı özelliklerini kullanırlar, fakat bu yeterli olmayabilir. Açık kaynak web güvenlik duvarı uygulaması "ModSecurity"; SQL enjeksiyonu, XSS saldırılarında veya XSRF tespitinde "Core Rule" kural setlerini kullanır. Ağ sistemleri saldırılara maruz kaldığında, ağ davranışlarını görüntülemek ve saldırı tekniklerini analiz etmek için IDS gibi bazı cihazlara güvenir. Bu cihazlardan gelen raporlar doğrultusunda sistem yöneticileri zafiyetleri giderir ve sistem güvenliğini sağlarlar. Bununla birlikte IDS sistemlerinin kabiliyeti geniş ve esnek web saldırıların tespitinde yetersizdir. Bu yüzden bu çalışmada ModSecurity core rule mimarisi kullanılarak Snort IDS programının önişleme özelliği geliştirilmiştir [11].

Lebeau ve arkadaşlarının yaptığı çalışmada son kullanıcıyı hedef alan XSS saldırıları incelenmiştir. Bu tür saldırılar sunucu cevap üretmek için kullanıcıdan bir girdi (form, url parametresi, cookie değeri) alındığında oluşur. Saldırgan, bir web uygulamasında kullanıcı girdisi kullanarak zararlı veri (örneğin JavaScript dilinde yazılmış, son kullanıcının tarayıcısında çalıştırılacak bir script) enjekte eder. Kullanıcı girdisinin kontrolü yapılmazsa saldırgan başarıya ulaşır [12].

Mookhey ve Burghate tarafından yapılan çalışmada bilinen XSS ve SQL enjeksiyonu saldırılarına karşı savunma çözüm önerisi olarak Snort IDS programına uygun kurallar önerilmiştir. Bu sayede saldırıların önleneceği yorumu yapılmıştır [13]. Yapılan birçok benzer çalışmada aşağıdaki XSS saldırılarında sıklıkla kullanılan ifadelerin, karakterlerin ve hexadecimal karşılıklarının kullanımının engellenerek bu saldırıların önlenebileceği önerilmektedir

Saldırı Türü	Saldırılarda Kullanılan Metakarakter ve Hex Karşılıkları
XSS	(\%3C)|<) ((\%2F)|\/) [a-z0-9\%] (\%3E)|>)
SQL Enjeksiyonu	/(\%27)|(\') (\-\-) (\%23) (\%6F)|o|(\%4F)) ((\%72)|r|(\%52) (#)

Tablo II - Xss ve sql saldırılarında kullanılan karakterler[13]

Saldırı Tipi	Anahatar Kelimeler
SQL Enj.	Or, --', '--, and, exec, select, insert, update, delete, drop, where, dbo, cast(,char(, union
XSS Sald.	Javascript, <scrip, /scrit>, document.write, eval(, expression(, <object, onload, onmouseover, onerror, windoes.open, <iframe, function, .location, onclick

Tablo III - Xss ve sql saldırılarinda kullanilan anahtar kelimeler [11]

Örneğin Alnabulsi ve arkadaşları Snort IDS programına bu ifadeler kullanıldığında alarm veren SQL saldırıları için 15 adet kural oluşturulmuştur [14].

Literatürdeki çalışmalardan görülmektedir ki tüm web saldırıları için etkili olan bir yönteme rastlamak zordur. Çoğunlukla önerilen saldırı tespit sistemleri için kural eklemek her saldırı için etkili olamamaktadır. Her saldırı için kural oluşturulsa bile yeni saldırı çeşitlerine karşı

yeterli olacağı kuşkuludur. Bu yüzden karmaşık saldırıların yaşandığı günümüzde açıklıkların önceden tespit edilmesi ve giderilmesi, mevcut saldırılara karşı hem imza tabanlı, hem de yeni saldırı türlerine karşı anormallik tabanlı saldırı önleme sistemlerinin kullanıldığı çözüm önerileri tercih edilmelidir.

III. SİTELER ARASI KOMUT DİZİSİ (XSS) SALDIRISI

OWASP ve CWE/SANS kuruluşlarının yayınladıkları son raporlara göre Siteler Arası Komut Dizisi Saldırıları (XSS) en sık karşılaşılan ve en ciddi saldırılardan birisidir. Bu saldırı saldırganın istemciye zararlı kod göndermesine imkan sağlayan bir çeşit kod enjeksiyonu açıklığıdır. Bu durum, bir web uygulaması HTML sayfalarında kullanıcı girdilerini uygun olarak kontrol etmeden referans olarak alması ve o veriyi web tarayıcısına kontrol etmeden veya şifrelemeden göndermesi sonucunda oluşur. Saldırgan, zararlı kodları uygulamanın HTML sayfalarında bu girdilerin içine gömebilir. İstemci bu yolla sömürülmüş web sitesini ziyaret ettiğinde; istemci tarayıcısı farkında olmadan zararlı kodu çalıştırabilir ve başarılı bir XSS saldırısı oluşur. Zararlı kod XSS saldırısında HTML, Javascript, VBScript ve Flash gibi birçok formatta olabilir. Bunun sonucunda XSS saldırıları hesap hırsızlığı, veri çalınması, cookie hırsızlığı ve zehirlenmesi, web içerik manipulasyonu ve hizmet dışı bırakma gibi şiddetli güvenlik ihlallerine sebep olabilir.

XSS birçok uygulama açıklığında ortaya çıkar. Bunlardan bazıları;

1. Kullanıcılar bilmeden saldırgan tarafından yüklenen zararlı kodları çalıştırabilir.
2. Saldırgan kullanıcı oturum cookielerinin süresi dolmadan kullanıcı oturumunu ele geçirebilir.
3. Saldırgan kullanıcıyı kendi istediği zararlı sunucuya yönlendirebilir.
4. Form özellikleri belli durumlarda saldırgan tarafından değiştirilebilir.
5. Bazı web siteleri bazı karakterlerin kullanımını kısıtlamadığından alternatif karakter kullanımı kullanıcıyı riske atabilir.
6. Kullanıcı tarayıcılarında programlama dillerinin çalıştırılması kısıtlansa bile, sunucuya giden istekte zararlı kod yazılarak bu kısıtlamaya izin verilebilir.
7. Saldırgan, kulanıcı tarayıcısında istediği URL'yi tıklattırıp program dosyası veya HTML'yi çalıştırtabilirse ayrıcalıklı yetkiye sahip olabilir.
8. SSL şifreli bağlantısı verinin içine yerleşen zararlı kodla bozulabilir. SSL bağlantısı verinin içeriğini kontrol etmez. Kullanıcı tarafında zararlı kod SSL olmayan siteye yönlendirebilir [10].

XSS saldırılarını üç gruba ayırmak mümkündür.

A. Kalıcı (Stored) XSS Saldırısı

Bu saldırı açıklığı, bir web sunucu programı kısıtlanmayan kullanıcı girdisini veri tabanında sakladığında ve program sonradan farklı kullanıcılar tarafından görüntülenen web sayfasında bu saklanan veriye ulaştığında ve referans gösterdiğinde ortaya çıkmaktadır. Bu XSS çeşidine daha çok forumlarda, bloglarda ve sosyal ağ sitelerinde rastlanır [15].

B. Kalıcı Olmayan (Reflected) XSS Saldırısı

Bu saldırı açıklığı HTTP isteğiyle gelen kontrol edilmemiş verinin sunucu tarafından alınması ve tekrar kullanıcıya geri gönderilmesiyle oluşur. Bu XSS tipinde kullanıcının yapması gerekenlerin ilki linke tıklamak, arama yapmak vb. uygulamaya özel fonksiyonlardan birini çalıştırmaktır. İkincisi, kullanıcı zararlı linke tıkladıktan sonra güvensiz siteye girmelidir. Kalıcı XSS'te olduğu gibi bu saldırı da sunucu tarafında uygun olmayan girdilerin kabul edilmesiyle mümkün olur [15].

C. Yerel (DOM Based) XSS Saldırısı

İstemcinin zararlı kodu sunucudan değil bir kullanıcıdan almasıyla oluşur. Bu saldırılar web sitesinin kendi kodları içinde yer alan açıklıkları hedef alır. Bu JavaScriptte Belge Nesne Modelinin dikkatsiz kullanımı sonucu oluşur. Zararlı yazılımla açılan diğer web sayfası aynı zamanda yerel sistemde ilk sayfanın kodlarını da değiştirebilir [16].

XSS saldırıları, site güvenliğini geçmek için HTML etiketleri (tag) ve biraz Javascript bilgisi dışında başka bir desteğe ihtiyaç duymadan web sitelerine karşı yapılmaktadır. Tarayıcılar Javascript formatındaki kodları kullanıcının bilgisi ve izni olmadan her zaman çalıştırırlar ve antivirüs yazılımları ile diğer masaüstü savunma mekanizmalarının kapsamı dışında kalırlar [17].

XSS saldırılarına karşı savunma olarak önerilen çözümlerin en önemlilerinden biri, önleyici kodlama uygulamalarıdır. Bu teknikte özel anlamlar içeren özel karakterler istemciye gönderilirken özel anlamları silinir. Bunun dışında bireysel seviyede güvenli yazılım kurallarına dikkat edilmeli, özellikle sisteme kullanıcı kodlarının enjektesine izin veren meta karakterlerin kullanılması engellenmelidir. Bunun için sadece sunucu tarafında değil özellikle tarayıcı bazlı çözüm önerileri geliştirilmelidir.

IV. SQL ENJEKSİYONU

SQL (Structured Query Language), veri tabanlarından veri seçme, silme ve güncelleme gibi işlemleri yapabilmek için kullanılan yapısal bir sorgulama dilidir. SQL Enjeksiyon ise, web uygulamalarından alınan kullanıcı girdileri ile oluşturulan SQL sorgularının manipülasyonu olarak tanımlanabilir [18]. Diğer bir ifade ile SQL enjeksiyonu arka planda çalışan SQL sorgularının çalışmasını değiştirmek için SQL meta karakterlerinin ve komutlarının web tabanlı girdi alanlarına enjekte edilmesidir [13].

SQL enjeksiyonu saldırılarında öncelikle veri tabanı ile ilgili saldırı amacına yönelik bilgi toplanması gerekmektedir. Veri tabanında bulunan tablo ve tablo alan isimleri denenen yanlış sorgular sonucunda oluşan ekrandaki hata mesajlarından yararlanılarak öğrenilebilir. Öğrenilen tablo ve tablo alan isimlerine göre saldırıya yön verilmesi gerekmektedir. Bu amaçla, ilk önce saldırı yapılacak web sitesi için neler yapılacağının belirlenmesi gerekir. Eğer bu web sitesi tamamen ele geçirilmek isteniyorsa buna yöneticinin kullanıcı adı ve şifresinin tutulduğu tabloyu bulmakla başlanabilir [18].

Veri tabanı kullanan bir web uygulamasındaki SQL

enjeksiyonu açığı tespit edildikten sonra yapılabilecek işlemler, uygulanan SQL ifadeleri sonucu alınan hatalar dikkate alınarak belirlenebilir. Böylece web sitesinin yönetimi ele geçirilebilir, veri tabanına yeni tablolar eklenebilir, var olan tablolar silinebilir. Veri tabanı yönetimi ele geçirilerek, sistem üzerinde yönetici hakkı elde edilebilir. Erişilen bilgilere göre saldırının türü, boyutu ve yapılabilecek işlemler farklılık gösterebilir. Bu nedenle, geliştirilen bir web uygulamasında muhtemel sızma veya saldırılara karşı, birçok husus göz önünde bulundurulmalıdır [19].

Yapılan kapsamlı literatür çalışmaları ile elde edilen bilgiler ve gerçekleştirilen örnek uygulamalar ile sızmalara karşı kazanılan tecrübeler ışığında, bu konuda ilk akla gelen öneriler aşağıda belirtilmiştir.

- Veri tabanında bulunan tablo ve tablo alanı isimlerinin kolay tahmin edilebilir olmaması,
- Web formlarında parametrik sorguların kullanılması,
- Web formlarında kullanılan ve veri tabanında bir kayıt satırını temsil eden sayısal değerler için (QueryString değeri) formlar arası geçişlerde bu değerlerin sayısal değer olup olmadığının kontrol edilmesi,
- SQL tabanlı web uygulamalarında kullanıcı, veri girişi yaptıktan sonra veri tabanına gönderilen SQL sorgusu karakterlerinde arama yaptırılarak tehlikeli karakterleri, SQL Sunucuda hataya yol açmayacak şekilde zararsız karakterlere çevrilmesi,
- SQL sunucusunda oluşacak hataların kullanıcı tarafından görüntülenmesi engellenmesi,
- Uygulamada web formlarına sorgu yazmak yerine, bu sorguları veri tabanı kısmında saklı yordam (Stored Procedure) olarak yazılması sağlanması,
- SQL enjeksiyon yönteminde kullanılabilecek sözcüklerin (select, insert, update vb.) bir fonksiyon ile filtrelenmesi,
- Sistem nesneleri için genel erişim verilmemesi, gerekirse kullanıcı bazında yetki verilmesi gerekir [18].

V. DVWA UYGULAMASI İLE SENARYONUN İCRASI

Yaptığımız çalışmada Firefox tarayıcısında çalıştırılan DVWA uygulaması, Snort 2.9.7.2 IDS Sistemi, Kali Linux 2.0, Burpsuite aracı, Sqlmap Aracı, Firebug ve Cookie Manager eklentisi kullanılarak XSS, ve SQL Enjeksiyonu saldırıları tatbik edilmiş; literatürde sıkça kullanılan Snort IDS sisteminin, tarayıcı seviyesinde kullanılan gömülü kodları engelleyen NoScript uygulmasının ve en önemlisi güvenli yazılım kurallarının varlığının bu saldırılara karşı etkinliği gözlenmiştir.

A. SQL Enjeksiyonu Saldırı Senaryosu

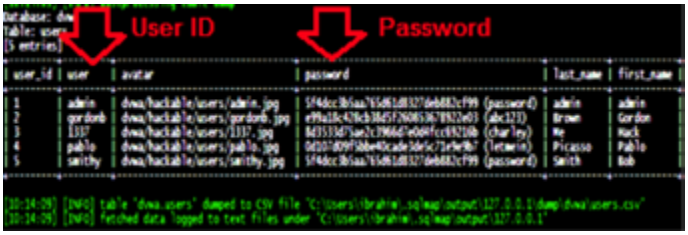
Yaptığımız SQL Enjeksiyonu saldırılarında aşağıdaki adımlar takip edilmiştir.

1. Bizim girdilerimizin uygulamanın davranışlarıyla aynı olduğunu ispatlamak için uygulamaya hata verdirme,
2. Kimlik doğrulamayı atlatmak için veritabanından kullanıcı isimlerini elde etme,
3. Veritabanından şifrelerin hashleri gibi yararlı bilgileri çekme,
4. Şifrelerin hashlerini kırarak kullanıcı isimlerini ve şifrelerini öğrenme [20].

Bunun için DVWA programının sol tarafındaki SQL Injection kısmı kullanılmış ve çeşitli sql sorguları kullanılarak parola ve şifre hashlerine(MD5) ulaşılmıştır.

Şekil 1. SQL Enjeksiyonu ile Kullanıcı Şifrelerine Ulaşılması

Yaptığımız bu saldırıda saldırgan mysql gibi programlama bilgisine sahip olmalıdır. Fakat saldırgan internetten rahatlıkla bulabileceği "havij" ve "SQLMap" gibi araçlarla proglamlama bilgisine sahip olmadan bu saldırıları gerçekleştirebilir. Aşağıda aynı senaryonun "SQLMap" aracılığı ile yapılması sonucu parola ve şifrelerin elde edildiği görülmektedir.



Şekil 2. SQLMap Aracı ile Kullanıcı ve Şifrelerin Elde Edilmesi

Log kayıtları incelendiğinde sqlmap aracının da arka planda bizim adımıza sql sorguları gerçekleştirdiği görülmektedir.

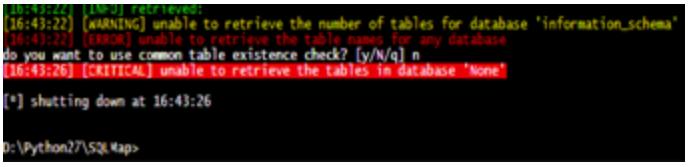
Örnek komut:

sqlmap.py -u "http://127.0.0.1/dvwa/vulnerabilities/sqli/?id=1&Submit=Submit#" --proxy="http://127.0.0.1:80" --cookie="security=low; PHPSESSID=ibfjqtd315k3q0iij4ibc74497" –D dvwa –T users --dump

Örnek Log Kaydı:

127.0.0.1 - - [17/May/2015:00:18:33 +0300] "GET http://127.0.0.1:80/DVWA/vulnerabilities/sqli/?id=1&Submit=Submit HTTP/1.1" 200 4775 "-" "sqlmap/1.0-dev-nongit-20150513 (http://sqlmap.org)"
127.0.0.1 - - [17/May/2015:00:18:33 +0300] "GET http://127.0.0.1:80/DVWA/vulnerabilities/sqli/?id=1&Submit=Submit&dzTr%3D4653%20AND%201%3D1%20UNION%20ALL%20SELECT%201%2C2%2C3%2Ctable_name%20FROM%20information_schema.tables%20WHERE%202%3E1--%20../%2F../%2F../%2Fetc%2Fpasswd HTTP/1.1" 200 4775 "-" "sqlmap/1.0-dev-nongit-20150513 (http://sqlmap.org)"

Güvenlik önerisi olarak örnek PHP uygulamamızın yazılım kısmında özel karakterler içeren sorguları düz metine çeviren "mysql_real_escape_string" fonksiyonunu kullandığımızda aynı saldırının başarısızlığa uğradığını görmekteyiz.



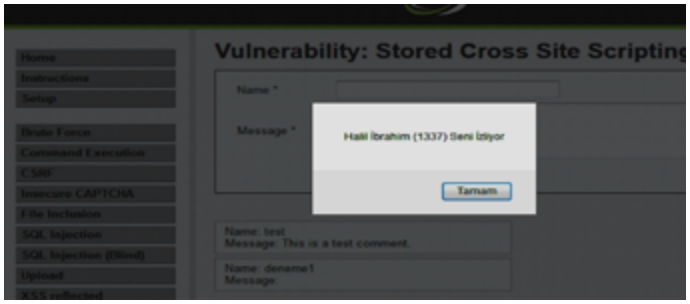
Şekil 3. Sqlmap Aracının Başarısız Olması

Sonrasında oluşan log kayıtlarında ise “302 Geçici Taşındı” hatası almaktayız.

127.0.0.1 - - [15/Aug/2015:16:42:43 +0300] “GET http://127.0.0.1:80/DVWA/vulnerabilities/sqli/?id=1&Submit=Submit HTTP/1.1” 302 - “-” “sqlmap/1.0-dev-nongit-20150513 (http://sqlmap.org)”

B. XSS Saldırı Senaryosu

XSS saldırılarında öncellikle web uygulamasında zafiyetin bulunması gerekir. Zafiyetin varlığını teyit etmek için <script>alert(“Halil İbrahim (1337) Seni İzliyor”) </script> komutu sitede çalıştırıldığında uygulamanın dışardan kod girişine izin verdiği görülmüştür.

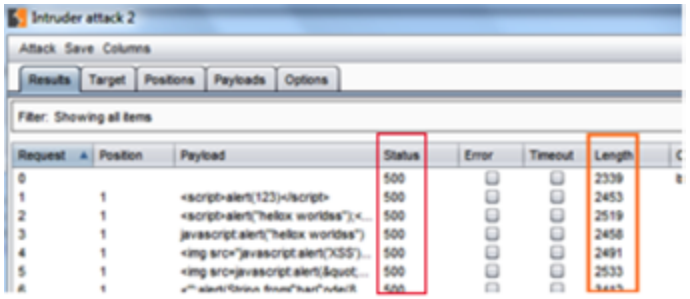


Şekil 4. Farklı Kullanıcı Giriş Yaptığında Zararlı Kodun Çalıştırılması

2. safhada sisteme kurulu olan Snort v2.9.7.2 IDS programına “scipt, alert, onclick..vb.” komutların çalıştırımasını engelleyici kural girilerek aynı saldırıda Snort’un alarm vermesi sağlanmıştır. Fakat saldırıların sürekli şekil değiştirmesi kural tabanlı güvenlik çözümlerini yetersiz kılmaktadır. Bu senaryoda engellenen ifadeler “ScriPt, alErt” büyük küçük harf değişiklikleri yapılarak aşılabilir.

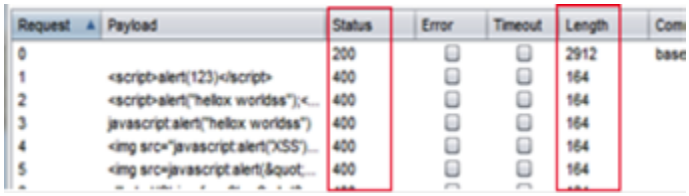
Şekil 5. Snort Programının XSS Saldırısına Karşı Alarm Üretmesi

3. aşamada hem mevcut güvenlik sistemlerini atlatmak hem de mevcut zafiyetin hangi komutlarla sömürülebileceğini bulmak için XSS saldırı komutları içeren bir sözlük [21] kullanılarak BurpSuite aracı ile uygulamaya bir nevi sözlük saldırısı gerçekleştirilmiştir.



Şekil 6. BurpSuite ile XSS Komutlarının Denenmesi

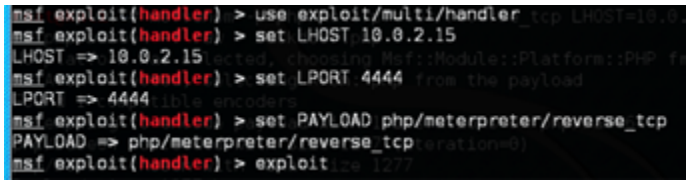
Burada Length bölümündeki değişiklikler o komutların enjekte edilebileceğini göstermektedir. Ayrıca yine Status bölümdeki 500 iç sunucu hatası XSS’in yapılabileceğini göstermektedir. XSS açıklığı bulunmayan başka bir uygulamaya aynı saldırı yapıldığında Length bölümündeki ifadelerin aynı kaldığını ve 400 hatasının oluştuğunu yani güvenlik nedeniyle ulaşılamadığını görmekteyiz. Bu durum, saldırı sonrası analizlerde 500 iç sunucu hatasının varlığının XSS saldırısına işaret edebileceğini göstermektedir.



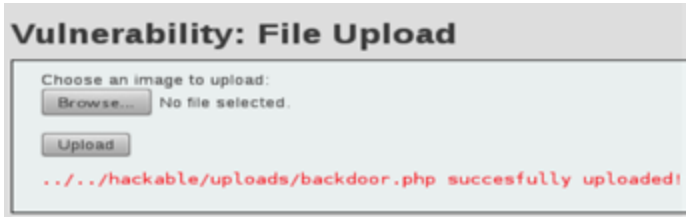
Şekil 7. XSS Açıklığı Bulunmayan Siteye Saldırı

4. aşamada zafiyeti bulunan DVWA uygulamasına zararlı PHP uygulaması enjekte etme yoluyla arka kapı oluşturulmuş ve uygulamaya uzaktan erişim sağlanmıştır. Bunun için Kali Linux 2.0 üzerinde kurulu gelen Metasploit aracı kullanılmıştır. Senaryoda Kalinin kurulu olduğu saldırgan IP adresi 10.0.2.15; DVWA uygulamasının kurulu olduğu kurban bilgisayar IP adresi 169.254.123.235’dir. İlk olarak hazırladığımız backdoor.php uygulamasını çalıştırdığımızda uzaktan komuta edecek IP adresi ve tünellemenin yapılacağı port ayarlanmıştır.

Şekil 8.Payload’un Şiferelenmesi ve Hazırlanması

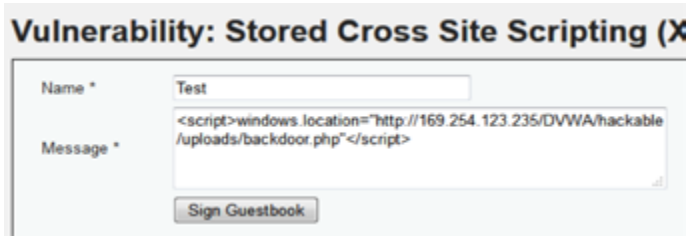


Şekil 9. Veri Aktarılacak IP ve Portun Belirlenmesi

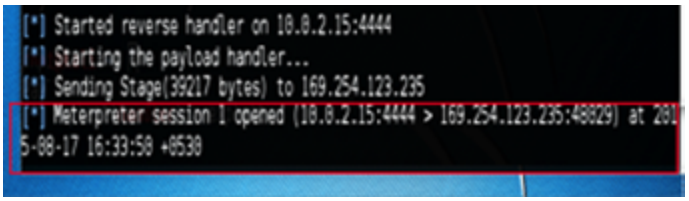


Şekil 10. Zararlı Yazılımın Sisteme Yüklenmesi

Ardından uygulamadaki stored XSS bölümündeki iletişim kutusuna zararlı kod enjekte edilmiş ve arka kapı sayesinde uygulamaya uzaktan bağlanılmıştır.



Şekil 11. Zararlı Yazılımı Çalıştıracak Kodun Enjekte Edilmesi



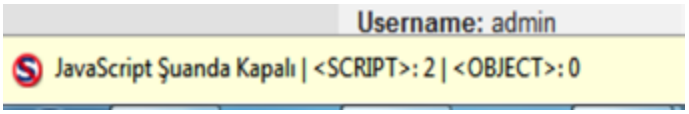
Şekil 12. Uygulamaya Uzaktan Bağlantının Başarılı Olması ve Uzaktan Erişimin Sağlanması

Saldırının bu kısmında meterpreter oturumun açılması uygulamanın kullanıcı adı ve şifrelere, cookie bilgileri bulunan dosyalara, uygulamadaki kayıtlı fotograf, veri gibi bilgilerin alınabilmesine ve sitedeki giriş sayfası gibi birçok alanda değişiklik yapabilmemize imkan tanımaktadır.

Yazılımsal olarak "mysql_real_escape_string, htmlspecialchars" gibi fonksiyonlarla XSS zafiyeti engellendiğinde uygulamanın zararlı kodları çalıştırmadığı sadece string ifadesi olarak kabul ettiği görülmüştür.

Şekil 13. Girilen Zararlı Kodun String İfadesine Dönüşmesi

Senaryo dahilinde yapılan saldırılar Firefox tarayıcısında eklenti olarak çalışan NoScript uygulaması kullanıldığında başarılı olunamamıştır.



Şekil 14. NoScript Uygulamasının Siteye Gömülü Zararlı Java Kodunu Engellemesi

C. Senaryo Dahilinde Yapılan Saldırıların Analizi

Yapılan saldırılar analiz edildiğinde;

- Yaptığımız çalışmada SQL Enjeksiyonu saldırısında; SQLMap aracı kullanılarak kullanıcı adları ve hashlerine ulaşılmıştır. Ardından hashler kaba kuvvet saldırısı yapılarak MD5 algoritması kullanıldığından kolayca çözümlenmiştir.
- Yaptığımız SQL Enjeksiyonu saldırısında açıklık bulunan uygulama, yazılımsal anlamda güvenli hale getirildiğinde saldırının başarısız olduğunu ve log kayıtlarından "302 Geçici Taşındı" hatası oluştuğu görülmüştür.
- Literatürde önerilen önleyici kurallar, Snort IDS programına eklendiğinde; kuralların kapsamına giren karakterler kullanarak yaptığımız XSS saldırılarının önlendiği görülmüştür. Fakat aynı saldırıların büyük, küçük karakter değişiklikleriyle veya kapsam dışındaki başka komutlar kullanılarak yapıldığında başarıya ulaştığı görülmüştür. Bu durum imza tabanlı güvenlik sistemlerin yeni saldırıları engellemede başarısız olduğunu göstermektedir.
- Ayrıca diğer bir XSS saldırısı olarak uygulamaya zararlı php yazılımı gömülerek uzaktan erişim imkanına sahip olunmuştur ve aynı kurallarla kullanılan Snort IDS programı bu saldırıyı önlemede başarısız olmuştur. Sistemde bulunan virüs programı da saldırılar esnasında alarm üretmemiştir.
- Yazılımsal anlamda güvenlik için çeşitli önleyici değişiklikler yapıldığında veya tarayıcı seviyesinde Firefox eklentisi olan NoScript gibi, istenmeyen Java kodlarının çalışmasını engelleyen programlar kullanıldığında yapılan tüm XSS ve SQL Enjeksiyonu saldırılarının başarısız olduğu görülmüştür.
- Bu sonuçlar göstermektedir ki XSS ve SQL Enjeksiyou saldırılarına karşı güvenlik, uygulama katmanında alındığında daha başarılı olunmaktadır.
- Ayrıca bu saldırı türlerinde sistemlerin, log kayıtlarından saldırının başarısız olması durumunda 300 ile başlayan hata kodları ürettiği gözlenmiştir.

VI. TARTIŞMA ve SONUÇ

Günümüzde web uygulamalarına yapılan saldırılar arasında XSS ve enjeksiyon saldırıları ciddi yer tutmaktadır. Yapılan çalışmaları incelediğimizde çözüm önerisi olarak genelde; filtreleme, güvenli kodlama, imza oluşturma, anormallik tespiti, hibrit sistemlerin kullanılması, bireysel farkındalık, model oluşturma, sunucu taraflı çözüm önerileri, kullanıcı taraflı çözüm önerileri gibi sonuçlara ulaşılmıştır [22].

Yaptığımız çalışmanın sonucunda özellikle güvenli yazılım eksikliklerinin bulunduğu, cookie bilgilerinin güvenliğinin yetersiz olduğu, kullanıcı adı ve şifreleri gibi değerli bilgilerin sistemlerde saklandığı, bu bilgilere sahip olunduğunda sistemlere girişte başka kimlik doğrulamanın yapılmadığı ve bunların şifrelenmesinde DES gibi zayıf algoritmaların veya MD5 gibi yetersiz özet fonksiyolarının kullanıldığı göze çarpmaktadır. Bu zafiyetleri kullanan saldırıları engellemek için yukarıda belirtilen çözüm önerilerine ek olarak;

- Öncelikle ulusal ve uluslararası alanda ağ güvenliği, uygulama güvenliği, yazılım güvenliği konularında geçerli güvenlik standartları sistemlere entegre edilerek uygulanmalıdır. Bu sayede insanlara yol haritası sunulmuş olacaktır.
- Hassas sistemler güvenliği zayıf sistemlerle korunmamalıdır. Diğer bir ifade ile eğer sahip olunan veri sistemden daha önemli ise o sisteme girmemelidir, işlenmemelidir ve en önemlisi o sistemde depolanmamalıdır. Örneğin etki alanı yönetici hesabı diğer kullanıcılarında olduğu ağa direk kullanıcı adı ve şifre özetini kullanarak giriş yapmamalıdır. Çözüm olarak geçici bir hesap yaratılabilir ve sonradan bu hesap silinebilir, veya özel olarak korunmuş ve internet bağlantısı olmayan sistemlere sadece giriş yapmasına izin verilebilir.
- Kullanıcılara asgari yetki verilmelidir. Kullanıcıların ihtiyaçlarından fazla yetkiye sahip olmaları zararlı yazılımlara ve sosyal mühendislik destekli saldırılara karşı sistemleri ve kurumları daha savunmasız hale getirir. Bu prensip uygulandığında açıklıkların %92 sinin engellendiği ortaya çıkmıştır [23].
- NTLM ve LM challenge-response Protokolleri kullanılmamalıdır. Bunun yerine NTLMv2 ve Kerberos kullanılmasının daha güvenli olduğu belirtilmektedir. Ayrıca NTLMv2 şifrelerin özetleri windowsta tutulmadığı için bu saldırılara karşı daha güvenlidir. Özetlemede daha güvenli olan HMAC-MD5 kullanılmalıdır [23].
- Gizli Bilgilerin önbelleklenmesi sınırlanmalıdır. Saklanan kimlik doğrulama için kullanılan gizli bilgilerin herzaman açıklık olduğu unutulmamalıdır. Bazı kurumlarda tek bir şifreyle birçok yere giriş yapılabilmektedir ve bir kere girildiğinde sistemde hemen saklanmaktadır. Buda depolanan (cookie) bilgilerin ele geçirilmesine sebep olabilir [24].
- Hem istemci (HIDS) kendi yerel hesabında ve gruplarında hem sistem yöneticisi (NIDS) günlük olarak ağda anormallik analizi yapmalıdır.

- Tüm bu çözüm önerilerin yanında İnternet tarayıcılarının kullanıcıları koruyucu önlemler alması gerekmektedir, çünkü XSS ve SQL saldırılarında tarayıcılar aktif olarak rol almaktadır. Firefox tarayıcısına eklenti olarak sunulan "NoScript" uygulaması örnek bir koruyucu uygulamadır.

Yukarıdaki önerilerin gereği yapıldığında sistemlerde belli bir seviyede güvenlik sağlanmış olacaktır. Fakat teknolojinin ve saldırganların hızla geliştiği günümüzde bu güvenlik çözümlerinin de yeterli olmadığı saldırılar karşımıza çıkabilir. Bu yüzden bireysel farkındalık kapsamında teknolojiye ayak uydurarak bireyler kendini eğitmeli, geliştirmeli ve esnek güvenlik çözümleri geliştirilmelidir. Unutulmamalıdır ki ağa bağlı olan en alt unsur dahi tüm sistemler için risk niteliği taşımaktadır. Bu yüzden güvenlik kapsamında baştan sona tüm unsurların güvenliği hesaba katılmalıdır.

KAYNAKLAR

[1] D. Wichers, "OWASP Top-10 2013 Report", https://www.owasp.org/images/1/17/OWASP_Top-10_2013--AppSec_EU_2013_-_Dave_Wichers.pdf, 2013.

[2] A. Razzaq, K. Latif, H.F. Ahmad, A. Hur, Z. Anwar, P.C. Bloodsworth, "Semantic security against web application attacks", Information Sciences, 2014, sayı 254, s.19–38.

[3] M. Cova, D. Balzarotti, V. Felmetsger, and G. Vigna, "Swaddler: An Approach for the Anomaly-Based Detection of State Violations in Web Applications", Proc. Int'l Symp. Recent Advances in Intrusion Detection (RAID '07), Lecture Notes in Computer Science, 2007, sayı 4637, s.63-86.

[4] C. Kruegel, and G. Vigna, "Anomaly Detection of Web-Based Attacks", Proc. 10th ACM Conf. Computer and Comm. Security (CCs'03), 2003, s.251-261.

[5] G. Vigna, W.K. Robertson, V. Kher, R.A. Kemmerer, "A Stateful Intrusion Detection System for World-Wide Web Servers", Proc. Ann. Computer Security Applications Conf. (ACSAC '03), 2003, s.34-43.

[6] M. Auxilia, D. Tamilselvan, "Anomaly Detection Using Negative Security Model in Web Application", Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference, 2010, s.481-486.

[7] G. Vigna, F. Valeur, D. Balzarotti, W.K. Robertson, C. Kruegel, and E. Kirda, "Reducing Errors in the Anomaly-Based Detection of Web-Based Attacks through the Combined Analysis of Web Requests and SQL Queries", J. Computer Security, 2009, 17(3), 305-329.

[8] C. Kruegel, G. Vigna, W. Robertson, "A multi model approach to the detection of web based attacks. Journal of Computer Networks,2005, 48(5), s.717-738.

[9] R. Sekar, "An Efficient Black box Technique for Defeating Web Appli-cation Attack", Proc. Network and Distributed system security sump.(NDSS), 2009.

[10] K. Selvamani, A. Duraisamy, A. Kannan, "Protection of Web Applications from Cross-Site Scripting Attacks in Browser Side". (IJCSIS) International Journal of Computer Science and Information Security, 2010, 7(3).

[11] C.H. Yang, and C.H. Shen, "Implement Web Attack Detectıon Engıne Wıth Snort By Usıng Modsecurıty Core Rules", The E-Learming And Information Technology Symposium Tainan, Taıwan, 2009.

[12] F. Lebeau, B. Legeard, F. Peureux, and A. Vernotte, "Model-Based Vulnerability Testing for Web Applications", 2013 IEEE Sixth International Conference on Software Testing, Verification and Validation Workshops, 2013, s.445-452.

[13] K.K. Mookhey, and N. Burghate, "Detection of SQL Injection and Cross-site Scripting Attacks. Security Focus", http://www.securityfocus.com/infocus/1768, 2004, Son Erişim Tarihi:05.05.2015.

[14] H. AlNabulsi, I. Alsmadi, M. Al-Jarrah, "Textual Manipulation for SQL Injection Attacks", I.J. Computer Network and Information Security 1, 2014, s.26-33.

[15] J. Pauli, "The Basics of Web Hacking- Tools and Techniques to Attack the Web", E-Book, Elsevier, 2013, s.105-123.

[16] L.K. Shar, H.B. Kuan Tan, "Automated removal of cross site scripting vulnerabilities in web applications", Information and Software Technology, 2012, sayı 54, s.467–478.

[17] M. Shema, "Hacking Web Apps, Detecting and Preventing Web Application Security Problems", http://dx.doi.org/10.1016/B978-1-59-749951-4.00002-3, 2012, Newnes e-kitap.

[18] D. Demirol, R. Daş, M. Baykara, "SQL Enjeksiyon Saldırı Uygulaması ve Güvenlik Önerileri", 1st International Symposium on Digital Forensics and Security (ISDFS'13), 20-21 May 2013, Elazığ, Turkey.

[19] N. Sakthipriya, K. Palanivel, "Intrusion Detection for Web Application: An Analysis", International Journal of Scientific & Engineering Research, 2013, 4(5), s.1824-1827.

[20] J. Pauli, "The Basics of Web Hacking- Tools and Techniques to Attack the Web". Syngress e-kitap, 2013, s.64-86.

[21] İnternet: Collection of Cross-Site Scripting (XSS) Payloads, http://www.smeegesec.com/2012/06/collection-of-cross-site-scripting-xss.html, Son erişim Tarihi:30.07.2015.

[22] I. Hydara, A.B. Sultan, H. Zulzalil, N. Admodisastro, "Current state of research on cross-site scripting (XSS) – A systematic literature review", Information and Software Technology, 2015, sayı 58, s.170–186.

[23] B. Ewaida, "Pass-the-hash attacks: Tools and Mitigation", SANS Institute InfoSec Reading Room, 2010.

[24] D. Stirnimann, "Windows Attack - Gain Enterprise Admin Privileges in 5 Minutes", www.csnc.ch, 2010, Son Erişim Tarihi: 06.06.2015.

Halil İbrahim ULUS Gazi Üniversitesi Fen Bilimleri Enstitüsü

Bilgi Güvenliği Mühendisliği Bölümünde yüksek lisans eğitimine devam etmektedir. Araştırma konuları arasında siber güvenlik, web uygulama güvenliği, SIEM sistemleri yer almaktadır.

Mehmet DEMİRCİ Gazi Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliğinde Yrd. Doç. Dr. olarak görev yapmaktadır. Araştırma konuları arasında ağ sanallaştırma, yazılım tanımlı ağlar, İnternet mimarisi, ağ ve bilgi güvenliği yer almaktadır.

ÇOKLU PARMAK İZİ TABANLI, YENİ BİR BİYOMETRİK KİMLİKLENDİRME TEKNİĞİ

Y. Sönmez, İ.L.Belenli ve E. Avcı

Özet — Biyometrik sistemler; klasik şifre kontrolü, kartlı geçiş vb. tekniklerden kavram olarak çok farklıdır. Çünkü biyometrik özellikler, kişinin değiştirmesi veya bir başkasına aktarması mümkün olmayan özelliklerdir. Bu noktadan hareketle biyometri, güvenlik seviyesini büyük ölçüde arttıran bir tekniktir. Biyometri tabanlı sistemler, kimlik belirleme uygulamalarında her ne kadar yüksek güvenlik ve başarı düzeyi vaad etse de, klasik parmak izi tanıma yöntem ve cihazlarının bilinen bazı dezavantajlarının var olması; parmak izi tanıma işlemi gerçekleştiren biyometrik yöntemlerin-sistemlerin güçlendirilmesi ve güvenlik düzeylerinin daha da arttırılmasının önemli olduğunu göstermektedir. Yapılan deneysel çalışmamız da parmak izine dayalı oluşturulan şifrelerin; parmak izinin kopyalanıp çalınmasına karşı yeni bir teknik ile güvenliğini arttırmaya yönelik bir yöntem geliştirilmiştir. Bu yöntemde biyometrik kimliklendirme kullanılan tek parmak izine dayalı şifrelemede parmak izinin kopyalanıp çalınması durumuna karşı tüm parmaklardan alınan izlerin statik veya dinamik kombinasyonlu bir şifreleme yapılarak parmak izine dayalı biyometrik kimliklendirmenin güçlendirilmesi önerilmiştir. Çalışmanın birinci kısmında klasik tek parmak izine dayalı şifrelemede parmak izinin nasıl birkaç basit adımda kopyalanıp çalındığı incelenmiştir. İkinci kısımda ise mevcut dezavantajlı durumun ortadan kaldırılmasına yönelik önerilen yöntem olan çoklu parmak izine dayalı statik veya dinamik kombinasyonlu şifrelemenin kavramsallaştırılmasına yer verilmiştir. Üçüncü ve son kısımda ise önerilen yöntemin gerçek zamanlı bir uygulaması yapılmış ve sonuçlar tartışılmıştır.

Anahtar Kelimeler — Biyometrik kimliklendirme sistemi; Parmak izi tanıma; Güvenliği arttırılmış parmak izi şifreleme.

Abstract — Biometric systems are very different in concept from the classic techniques such as password control and passing with card. Because, biometric features are such kind of features whose being changed and being transfered to somebody else isn't possible. For this reason, biometry is a technique which greatly improve the security level. Biometrics-based systems promises high security and achievements in the applications of identification. However, there are some known disadvantages of classical fingerprint identification methods and devices. Therefore empowering the biometric systems that carry out fingerprint identification and improving their security level is important. And in our experimental study, we tried to develope a method that improves the security level against the fingerprints being copied and stealed by someone else. At this point, against single-fingerprints being copied and stealed, by enciphering all fingerprints with their static and dynamic combinations we propose to empower biometric identification based on fingerprints. So, in the first part of the study, it is shown how in very simple steps, the fingerprints are copied and stealed in classical single-fingerprints encryption. And in the second part, we give place to the conceptualization about the static and dynamic encryption based on multiple fingerprints. And in the last part, we devoloped a real-time application of the proposed method and thereby discuss the results.

Keywords — Biometric identification system; Fingerprint recognition; Improved security fingerprint encryption.

I. GİRİŞ

Teknolojinin hızla gelişmesiyle kişiye ait sayısal verilerin artması; bu verilerin korunmasına yönelik yöntemlerin gelişmesi yönündeki artışı beraberinde getirmiştir[1]. Bu koruma veya güvenlik yöntemlerindeki temel amaç kişiye ait veriyi sadece o kişiye sunmak ve yabancı kişilere karşı veriyi kapatmaktır. Yani veri kişi eşlemesini doğru ve güvenli şekilde sağlamaktır [1,2]. İşte bu veri kişi eşlemesi doğru, güvenilir ve sağlam bir kimlik tespiti gerekliliğini de beraberinde getirmiştir. Biyometrik sistemler önceden belirlenen ve depolanan davranışsal ve fizyolojik özelliklerle karşılaştırma yaparak kişi kimlik eşleştirmesini doğrulamaktadır. Kişi kimlik belirleme uygulamalarında sıklıkla tercih edilen biyometrik güvenlik sistemleri her zaman daha kaliteli çözümler sunmaktadır. Çünkü kişiye ait biyometrik özellikler kişinin değiştirmesi veya bir başkasına aktarması mümkün olmayan özelliklerdir. Bu biyometrik güvenlik sistemlerinden en önemlisi ve pratikte en çok kullanılanı parmak izi tanıma yöntemidir [3,4]. Araştırmacılar biyometrik güvenlik sistemlerinden söz ederken "kişilerin fiziksel ya da davranışsal özelliğine dayanarak gerçekleştirildiği için başkasına devredilmesi, unutulması ya da kaybedilmesi mümkün değildir." [5] . Prensibine dayanarak parmak izinin tek yumurta ikizlerinde bile farklı olduğu bilgisine sahip olduklarından biyometrik güvenlik sistemlerinin kişiye özgü olduğu kuramını savunmaktadırlar.

Bu kuram ışığından bakıldığında kişi veri eşleştirmesi yapılırken bir parmak izinin kullanılması yeterli olduğu görüşü çıkarılabilir . Ayrıca hali hazırdaki uygulamalarda tek parmak izine dayalı şifreleme yöntemlerinin kullanıldığı gözlemlenebilir. Ancak aşağıda bir parmak izinin bir kaç basit hamle ile nasıl kopyalanıp-çalınabildiği yöntemi anlatılmıştır. Bu yöntemde anlatılan adımlar uygulandığında istenilen kişinin parmak izinin kopyası oluşturulmuş ve parmak izi okuyucu sisteminden okutularak parmak izine dayalı şifreleme yapan sisteme giriş yapılarak bir kopyalama-çalma işlemi gerçekleştirilmiştir.

A. Bir parmak izinin kopyalanmasının aşamaları

Bu kısımda şifre için kullanılan bir parmak izinin nasıl kopyalanıp şifre olarak kullanıldığı anlatılmıştır [6]. Bu adımları uygulayan kişi tarafından parmak izi kopyalanıp bir şifre olarak kullanılmasında başarılı sonuçlar elde edildiği açıkça ifade edilmektedir.

Ön hazırlık aşaması:

Parmak izinin kopyalanması ve yapay bir parmak izinin oluşturulması için çeşitli malzemelere ihtiyaç vardır. Bunlar fotokopi tozu, makyaj fırçası veya benzeri bir yumuşak fırça, iyi bir makro fonksiyonuna sahip dijital kamera, görüntü işleme programı, şeffaflık baskı özelliği olan yüksek çözünürlüklü yazıcı, ışığa duyarlı foto aşındırma spreyi, UV ışığı (isteğe bağlı, normal bir ampul), Bakır kaplama devre

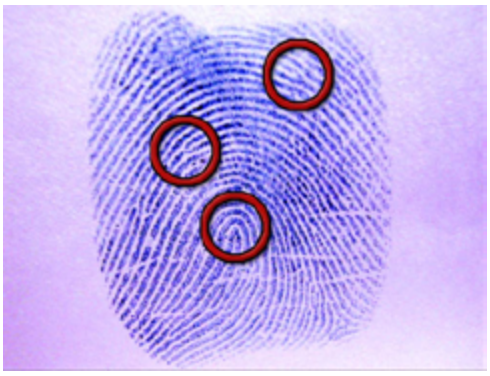
kartı, NaOH (kostik) karışımı (yaklaşık 3DL), FeCl3 (Demir klorür) karışımı (yaklaşık 3DL), Yumuşak suluboya fırçası jelâtin yaprakları (40g jelâtin + 1/2dl su) su ısıtıcısı ve buzdolabı gibi malzemelerdir.

Parmak izinin elde edilmesi için aşağı adımlar sırasıyla uygulanır.

- Yapay izi oluşturulacak olan parmağı gizlice izlenir (Bardak, kapı kolu vs).
- Fotokopi tozunu yavaşça parmak izine serpilir.
- Aşırı tozu atmak ve baskı net yapmak için makyaj fırçası kullanılır.
- Kamera ile yüksek çözünürlüklü bir fotoğraf çekilir.
- Ölçek tanımlayarak iki karakteristik nokta arasındaki mesafeyi ölçülür.
- Toz silinir ve parçayı temizlenir.

Son adıma ulaşıldığında Şekil 1'deki gibi bir görüntü elde edilir. Şekil 1'de görülen görüntü üzerinde gerçek parmak izi üzerinde de olduğu gibi hatlar vadiler ve sırtlar denilen hatlar açıkça gözlenmelidir.

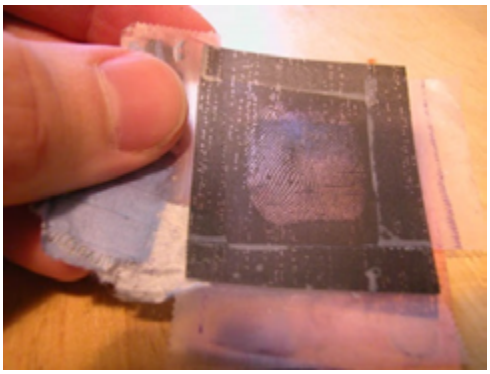
Şekil 1. Bardak üzerindeki gizli parmak izinin tozla görünür hale getirilmesi.



Şekil 2. Parmak izi üzerindeki hatlar vadiler ve sırtlar.

Parmak izinin oluşturulması için aşağıdaki adımlar sırasıyla uygulanır.

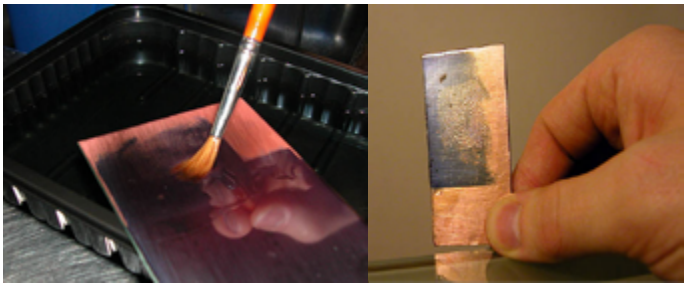
- Resim işleme programı kullanılarak parmak izi görüntüsünü düzenlemek için hazırlanır.
- Negatif (invers) görüntüyü elde edilip yazdırılır.



Şekil 3. Görüntü negatifinin şeffaf kâğıda yazdırılması.

Elde edilen parmak izi görüntüsünden kalıp çıkarılması.

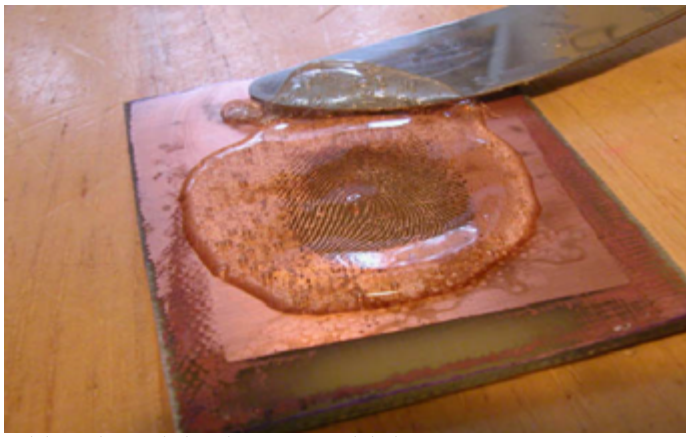
- Fotoğrafı devre kartı üzerine sprey ile cilalayarak bir süre kurumaya bırakın.
- Şeffaf bandı yaklaşık 15 dakika UV ışığı altında bırakın.
- Şeffaf bandı çıkarın.
- NaOH çözeltisi kullanarak cilalamayı geliştirin sulu boya fırçası kulanın.
- Parmak izi bulunan şeffaf bandı su ile yıkayın.
- FeCl3-çözeltisi kullanarak bakır levhayı temizleyin.
- Bakır levhayı bol su ile yıkayın
- Kalan cila kalıntılarını durulayın. (sabun veya alkol kullanmayın.)



Şekil 4. Kalıp çıkarma aşamaları.

Kalıbı çıkarılan parmak izinden artık bir sahte parmak izi oluşturabiliriz.

- Jelatin yaprağı 5 dakika soğuk suda bekleterek yumuşatın.
- 1/2dl suyu kaynayana kadar ısıtın.
- Sıcak su içine yumuşatılmış jelatin yaprağı bırakın.
- 10 dakika boyunca karıştırın.
- Karışımın biraz soğumasını bekleyin.
- Jelatini baskı kalıbını kapsayacak şekilde kalıbın üzerine dökün.(Parmak çok kalın olmayacak şekilde uygulayın)
- Baskı kalıbını en az 15 dakika buzdolabında bekletin böylece jelatin kurumuş olacaktır.
- Jelatin pıhtılaşmış hale gelince kalıptan ayırabilirsiniz. (Köşesinden hafifçe bir bıçak yardımıyla yavaş yavaş kaldırın)
- Jelatin artık gerçek bir parmak izine dönüşmüş olmalıdır.
- Elde edilen jelatindeki parmak izini oda sıcaklığında tutmaya özen gösteriniz.



Şekil 5. Jelatinin kalıptaki iz üzerine dökülmesi.

Son aşama oluşturulan parmak izinin gerçek bir sistem üzerinde denenmesi.

Bu aşamalardan sonra artık elinizde gerçek bir parmak gibi hissedilen jelâtinden bir parmak olmalıdır. Sahte parmak izi ile yani jelâtin parmağı okuyucuya yerleştirerek sisteme giriş yapabiliriz. Buradaki hassas nokta oluşturulan jelâtinini okuyucuya çok sert veya çok hafif olarak bastırılmamasıdır. Yukarıda anlatılan yöntemde şifre olarak kullandığımız parmak izinin kopyalanması ve yeniden oluşturulması anlatılmıştır. Bu kopyalayıp çalma yöntemine karşı yani parmak izimizin kopyalanmaması için bilinen veya literatüre geçmiş herhangi bir yöntem tespit edilmemiştir.

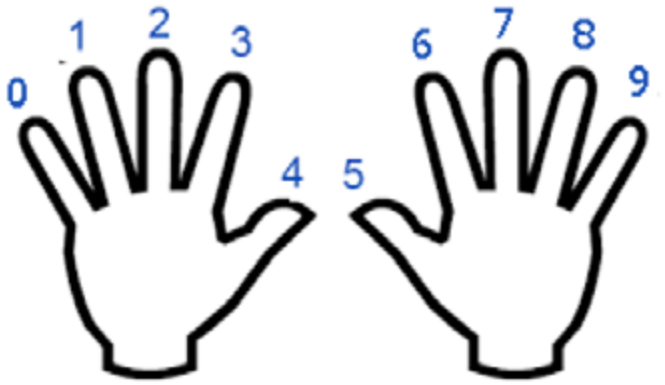
II. YÖNTEM

Sayısal şifreler 0 dan 9 kadar olan rakamların kombinasyonuna dayanarak genellikle dört haneden oluşmaktadır. Sayısal şifrelemede rakamlar kullanılarak oluşturulan kombinasyona göre yani belirli sırada girilmesi beklenir.

Dört haneli bir şifrenin tahmin edilme olasılığı 10^4 =10000 dir. Sayısal veriye dayalı şifreler genellikle üç kez denemeden sonra bloke edilmektedir. Parmak izi tabanlı şifrede ise tek bir parmak izi ve genellikle sağ veya sol işaret parmağın izi ile kullanılmaktadır. Parmak izi şifrelemede genellikle bloke işlemi yapılmamaktadır; bloke yerine okuyucu parmağı okuyup veritabanı karşılaştırması yaparak böyle bir parmak izinin olmadığını iletmektedir. Bu işlem bir şifre mantığına dayanmadığından dolayı herhangi bir bloke işlemi oluşmamaktadır. Bu durumun dezavantajı kopyalanıp çalınan bir parmak izinin defalarca denenmesine karşı herhangi bir kontrolün olmamasıdır.

Geliştirilen yöntemde ise sayısal şifreler dinamik veya statik olarak çoklu parmak izi tabanlı kombinasyona dayanmaktadır.

Statik kombinasyonda prensip her parmağımıza hayali olarak belirlediğimiz sıraya göre bir rakam verilir örneğin sol serçe parmak sıfırdan veya sağ serçe parmak sıfırdan başlayacak şekilde numaralandırma yapılır. Böylece her parmak 0 dan 9'a kadar olan bir rakamı temsil eder. Statik kombinasyona dayalı çoklu parmak izine dayalı şifre oluştururken parmağınızdaki hayali rakamlar olduğunu varsayarak bu kombinasyona göre sırasıyla parmağınızı okutup şifrenizi güvenli bir şekilde kullanmış olursunuz. Örneğin şifreniz 1453 gibi dört haneden oluşan bir kombinasyona dayanmaktadır. Okuyucuya parmaklar sırasıyla şu şekilde okutulur. 1- Sol yüzük parmak, 4- sol baş parmak 5- Sağ baş parmak 3- sol işaret parmak olacak şekilde okuyucuya girip başarılı bir şekilde sisteme giriş yapılır. Statik kombinasyonda şifrenin kırılma olasılığı 4 haneden oluşan bir şifre için 10^4=10000 olasılığın yanında kişinin on parmak izinin de kopyalanması gerekmektedir. Ayrıca parmak izi okuma işlemine bloke işlemi de getirilerek on parmağın izinin de kopyalanmasına karşıda ekstra bir güvenlik adımı da oluşturulmuştur.



Şekil 6. Sol baş parmak 0 dan başlayan statik kombinasyon için hayali rakamlar.

Dinamik kombinasyonda ise prensip statik kombinasyondan farklı olarak her parmağımıza hayali olarak belirlediğimiz sıraya göre bir rakam vermek yerine sistem sizden rastgele oluşturduğu kombinasyona göre parmak izlerinizi okutmanızı istemektedir. Kişi ekranda görsel olarak beliren parmakları okutarak sisteme başarılı şekilde giriş yapabilir. Dinamik kombinasyonlu çoklu parmak izi sistemi kullanımı oldukça basit ve hızlıdır kişi herhangi bir rakam kombinasyon ve sırayı aklında tutmak zorunda değildir. Bu kadar basit ve hızlı bir yönteme karşılık güvenlik en üst seviyededir. Çünkü sistem parmak izlerini rastgele kendi üretip isteyebildiği gibi SMS OTP yada mobil OTP olarak üretip isteyebilir. Dinamik çoklu parmak izi tabanlı şifrelemenin henüz bilinen bir dezavantajı veya kopyalanıp çalınmasına karşı bir yöntem tespit edilmemiştir.

III. UYGULAMA

Kimlik doğrulama sisteminin uygulaması 2 aşamadan oluşmaktadır. İlk aşamada kullanıcı kendi belirleyeceği bir eşik değeri baz alarak sisteme kaydını gerçekleştirecektir. Seçmiş olduğu eşik değer kaç parmağını sisteme kaydedeceğini ve sisteme giriş esnasında kaç parmağını kullanacağını belirleyecektir. Örneğin eşik değer olarak (7,4) değerini belirleyen bir kullanıcı, sisteme 7 parmağını tanıtacaktır. Doğrulama yapmak istediğinde ise sistem kendisinden rastgele 4 parmağını tanıtmasını isteyecektir. Her bir parmağında rakamsal olarak ifade edilişi Şekil 7'de gösterilmiştir.

Şekil 7. Parmakların rakamsal olarak ifade edilişi

Kullanıcı sisteme kayıt işlemini gerçekleştirdikten sonra ilk aşama tamamlanacaktır. Sisteme giriş yapmak istediğinde ise giriş ekranı ile karşılaşacaktır (Şekil 8).

Şekil 8. Kimlik doğrulama sisteminin uygulaması

Giriş ekranında sistem rasgele ürettiği rakamsal değeri baz alarak kullanıcıya hangi parmağını okutması gerektiğini görsel olarak belirtecektir. (7,4) eşik değerini kullanarak sisteme kaydolmuş bir kullanıcı için sistem şunu yapar. Rastgele 4 parmak seçer ve ekranda bu dört parmağı sırasıyla göstererek, kullanıcıdan bu parmakları girmesini bekler. Kullanıcı sırasıyla birinci, ikinci, üçüncü ve dördüncü parmağını doğru ve sırasıyla okutursa sisteme başarılı bir şekilde giriş yapmış olacaktır.

Örneğin sistem rastgele 1,4,9 ve 6 rakamlarını (parmaklarını) seçmiş olsun. Seçilen bu parmaklar ekranda kullanıcıya sırasıyla gösterilecektir. Kullanıcı, Şekil 9'da ki gibi parmaklarını sırasıyla sisteme okutarak, doğrulama işlemini gerçekleştirmiş olacaktır.

Şekil 9. Rasgele oluşan şifreye gore okutulan parmaklar

IV. SONUÇ

Günümüzde bilgi güvenliğini sağlayabilmek çok büyük bir problemdir. Bilgilerimizin, üçüncü kişilerin eline geçtiği takdirde, hukuki olarak çok durumlar ile karşı karşıya kalabiliriz. Bu nedenle biyometrik güvenlik sistemleri, bilgilerin korunması anlamda büyük önem arz etmektedir [7].

Biyometrik güvenlik sistemlerinin, ek bir bilgi, donanım, yazılım, şifre, araç kullanmak zorunluluğunun olmaması, çalınma, unutulma, kaybolma gibi tehlikelerin yok denebilecek kadar az olması gibi avantajları nedeniyle ilerleyen zamanlarda, kimliklendirme\doğrulama sistemlerinde çok daha aktif rol oynayacağı öngörülmektedir.

Bu çalışmada öncelikle tek bir parmak izinin bir kaç basit hamlede nasıl kopyalanabileceği detaylı bir şekilde anlatılmıştır. Böylesine basit bir kopyalama tekniği kullanılarak bile sahteciliği yapılabilen bir biyometrik güvenlik sisteminin, sanıldığının aksine aslında çokta güvenli olmadığı açıktır. Kolay kopyalanabilmesinin neticesinde, biyometrik doğrulama sistemlerinde tek bir parmak izinin kullanılmasının ne kadar doğru olduğu görüleceği üzere tartışmaya açıktır.

Geliştirilen yöntemin amacı, tek bir parmak izine dayalı biyometrik kimliklendirme sistemlerinde karşılaşılan kopyalanma ve çalınma olaylarına karşı var olan zafiyetlerini gidermektir. Yöntemin sunduğu çoklu parmak izi hipotezinin, belirli bir kombinasyona dayanan ve kişi faktörünü de ele alan bir yapıda olması ise ayrıca güvenliği arttırıcı bir unsurdur. Yapılan test çalışmalarının sonucunda, tek parmak izi tabanlı şifrelemeye karşı çoklu parmak izi tabanlı statik veya dinamik şifrelemenin başarım oranının bariz şekilde daha yüksek olduğu gözlenmiştir.

Bu makalede sunulan "Çoklu parmak izi tabanlı biyometrik şifreleme yöntemi" adlı yöntem literatürde ilk çalışmadır ve biyometrik kimliklendirme çeşitlerinden olan parmak izine dayalı şifreleme alanında bir hipotezdir. Tüm hakları yazarlara aittir.

KAYNAKLAR

[1] Y., Sönmez, M.Karabatak, E. Avcı, "Uygulamalarda Şifre Güvenliği İçin Yeni Bir Yaklaşım", Uluslararası Bilgi Güvenliği ve kriptoloji Sempozyumu (2013): 311-315.

[2] G., Canbek, Ş. Sağıroğlu, "Bilgi, bilgi güvenliği ve süreçleri üzerine bir inceleme", Gazi Üniversitesi Politeknik Dergisi, 9.3 (2006).

[3] Ü. A., Aydın, C., Acartürk. "Kullanılabilir Güvenlik ve Grafik Şifreler", Türkiyede İnternet Konferansı. 2012

[4] N. Özkaya, N., Sağıroğlu, Ş., "Açık Anahtar Altyapısı ve Biyometrik Sistemler", I. Ulusal Elektronik İmza Sempozyumu, 7–8 Aralık 2006, s.283-290, Ankara, Türkiye.

[5] R., Şamlı, M.E., Yüksel, "Biyometrik güvenlik sistemleri", Akademik Bilişim'09 (2009): 11-13.

[6] "Hacking Fingerprint Recognition Systems", https://pacsec.jp/psj06/psj06krissler-e.pdf- son erişim "20/09/2015"

[7] Belenli I.L., Tuncer T., Demir F.B., Avci E., Ulas M., "A Secure Web Application Based Visual Cryptography and Secret Sharing", Journal of Multidisciplinary Engineering Science and Technology, Mart 2015, s. 443,446

Not: Bu makalede yapılan deneysel çalışmalarda kullanılan

parmak izleri “ http://biometrics.idealtest.org/ “ adresinde araştırmacılar için sunulan test verilerinden elde edilip gerekli izinler alınmıştır.

Yasin Sönmez - yasin.sonmez@dicle.edu.tr -1986 yılında Diyarbakır’da doğdu. 2010 yılında F.Ü. Bilgisayar Öğretmenliği Bölümünde lisans, 2012 yılında Elektronik ve Bilgisayar Eğitimi ABD’da yüksek lisans programını tamamladı. Halen F.Ü. Yazılım Mühendisliği ABD’da doktora öğrenimine devam etmekte olup Dicle Üniversitesi Teknik Bilimler M.Y.O. Öğretim Görevlisi olarak çalışmaktadır.

İbrahim Levent BELENLİ - ibrahimbelenli@hotmail.com.tr- 1989 yılında Mersin Silifke’de doğdu. 2012 yılında F.Ü. Bilgisayar Öğretmenliği Bölümünde lisans, 2015 yılında Yazılım müh. ABD’da yüksek lisans tamamladı. Halen İnönü.Ü. Bilgi İşlem Daire Bşk.’da görev yapmaktadır.

Engin Avcı - enginavci23@hotmail.com- 1978 yılında Elazığ’da doğdu. 2000 yılında F.Ü. Elektronik Öğretmenliği Bölümünde lisans, 2002 yılında Elektronik ve Bilgisayar Eğitimi ABD’da yüksek lisans ve 2005 yılında Elektrik-Elektronik Mühendisliği ABD’da doktorasını tamamladı. Halen F.Ü. Yazılım Mühendisliği Bölümü’nde öğretim üyesi olarak görev yapmaktadır.

BİYOMETRİK SİSTEMLERDE GÜVENLİK ÜZERİNE BİR İNCELEME

Ceren GÜZEL TURHAN, Eyüp Burak CEYHAN, Şeref SAĞIROĞLU

Özet — Biyometrik teknolojilerdeki gelişmeler bu alanda yeni bir endüstrinin ortaya çıkmasına neden olmuştur. Günümüzde biyometrik teknolojiler geleneksel doğrulama mekanizmaları yerine kabul görmüştür. Bu teknolojilerin geleneksel yaklaşımlara göre çok daha güvenli oldukları kabul edilmiştir; fakat yapılan araştırmalarda biyometrik sistemlere de kolaylıkla sızılabildiği görülmüştür. Bu çalışmada bu ihtiyaçtan yola çıkılarak biyometrik teknolojilerde güvenlik kavramına odaklanılmıştır. Bu amaçla, biyometrik sistemlerde güvenlik üzerine literatürde yer alan çalışmalar kapsamlı şekilde incelenmiştir. Güvenlik üzerine sunulan modeller de tanımlanan saldırıya maruz kalınabilecek noktalar ele alınarak biyometrik güvenlik konusunda bir farkındalık oluşturmak amaçlanmıştır.

Anahtar Kelimeler — Biyometri, biyometrik sistemler, güvenlik, biyometrik saldırılar

Abstract — Advances on biometric technologies occur a new industry on this field. Nowadays, biometric technologies are considered as a verification mechanism against traditional mechanisms. Although these technologies are supposed to be much more secure, it is seen that it is also possible to penetrate these systems. Therefore, this study focuses on security requirement on biometric systems. In this study, a comprehensive review is presented to answer this need. It is aimed to raise awareness about biometric security with analyzed attack points models.

Index Terms — Biometrics, biometric systems, security, biometric attacks

I. GİRİŞ

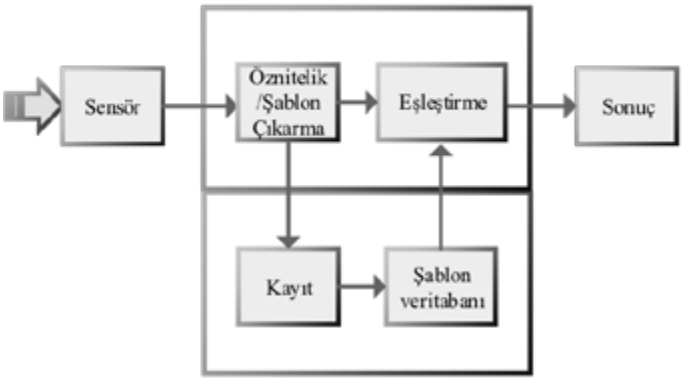
Teknolojik gelişmeler biyometrik tabanlı doğrulama sistemlerinin ortaya çıkmasına ve birçok alanda yaygın kullanımına olanak tanımıştır. Parmak izi gibi biyometri adı verilen karakterleri tanımaya odaklı çalışan doğrulama mekanizmaları ile yetkili kişileri sisteme sızmaya çalışan kişilerden ayırt edebilmek söz konusu olmaktadır. Bu durum, farklı biyometrik özelliklere dayalı sistemlerin geliştirilmesine neden olmuştur.

Geleneksel olarak şifre, anne kızlık soyadımız gibi bildiğimiz bilgiler doğrulanarak güvenlik sağlanmaya çalışılmaktadır. Şifre, kullanıcı adı gibi bilgiler bir veritabanında tutularak sisteme erişmeye çalışan bir kişiden alınan bilgiler ile eşleştirme yapılmaktadır. Eşleştirme yani doğrulama işlemi gerçekleştirilirse ise sisteme erişim yetkisi elde edilmektedir. Bu geleneksel yaklaşım performans açısından etkin olmakla birlikte bilgilerin saldırganlar tarafından elde edilmesi ile saldırganların sisteme yetkisiz erişimi gibi tehlikelere maruz kalabilmektedir. Doğrulama işlemi bildiklerimizin sorgulanması yerine sahip olunan bir akıllı kart ya da anahtar gibi nesneler ile de gerçekleştirebilmektedir. Bu doğrulama yaklaşımında ise kaybetme, çalınma gibi durumlarda sistem yetkisiz kişilerin hedefi haline gelebilecektir. Bahsedilen geleneksel yöntemlerin yerini günümüzde biyometrik denilerek nitelendirilen kendimize ait olan özellikler ile yapılan doğrulama işlemleri almaktadır. Biyometrik doğrulama mekanizmaları çalınamaz ve kaybedilemez olmaları nedeniyle diğer doğrulama mekanizmaları arasından en güvenilir olanı olarak nitelendirilmektedir [1]. Biyometrinin daha güvenli bir doğrulama mekanizması olarak değerlendirilmesi bu konu üzerine çalışmalara neden olmuştur. Bu çalışmalar farklı biyometrik karakterlerin doğrulama amacıyla kullanılabildiğini ortaya koydukları gibi farklı alanlarda uygulamalar geliştirilmesine imkan vermiştir. Güvenli olarak algılanan biyometri tabanlı doğrulama mekanizmalarının düşünüldüğü kadar güvenli olmamaları çeşitli çalışmalarda ortaya koyulmuştur. Bu nedenle, makale çalışması kapsamında günümüzde hemen her alanda hayatımızda yer edinen biyometri tabanlı sistemlerdeki güvenlik kavramı üzerine odaklanılarak bu alanda yapılan çalışmalar üzerine bir inceleme sunulmuştur. Daha önce yapılan çalışmalarda tanımlanan farklı güvenlik problemleri ele alınmıştır.

Bu makale 4 bölümden oluşmaktadır. Makalenin devam eden kısmı şu şekilde organize edilmiştir. 2. bölümde biyometrik sistemler ele alınarak basit bir biyometrik sistemde olması gereken bileşenler tanımlanmıştır. Biyometrik sistemlerde biyometrik karakter olarak kullanılabilecek karakterler kategorik olarak ele alınmıştır. Biyometrik teknolojilerin kullandıkları karakterlere göre uygulama alanları açıklanmıştır. 3. bölümde biyometrik sistemlerde güvenlik sağlamak üzere tanımlanan modellere ilişkin çalışmalar incelenmiştir. Son olarak sonuç bölümünde ise biyometrik sistemlerde güvenlik gereksinimi vurgulanarak bu amaçla yapılan çalışmalardan elde edilen kazanımlar değerlendirilmiştir.

II. BİYOMETRİK SİSTEMLER

Basit bir biyometrik sistem temel olarak 4 adımdan oluşacak şekilde incelenebilmektedir. Bu adımlar biyometrik verilerin bir algılayıcı aracılığıyla sisteme alınması, alınan veriden öznitelik vektörlerinin elde edilmesi, elde edilen öznitelik vektörlerinin daha önce yapılan bir kayıt işlemi ile veritabanlarına kaydedilen şablonlar ile eşleştirmesi ve yapılan eşleştirme sonucunda alınan skora göre sisteme erişim kararının oluşturulmasıdır [2]. Şekil 1'de temel bir biyometrik sistem modeli gösterilmektedir.



Şekil 1. Biyometrik sistem modeli

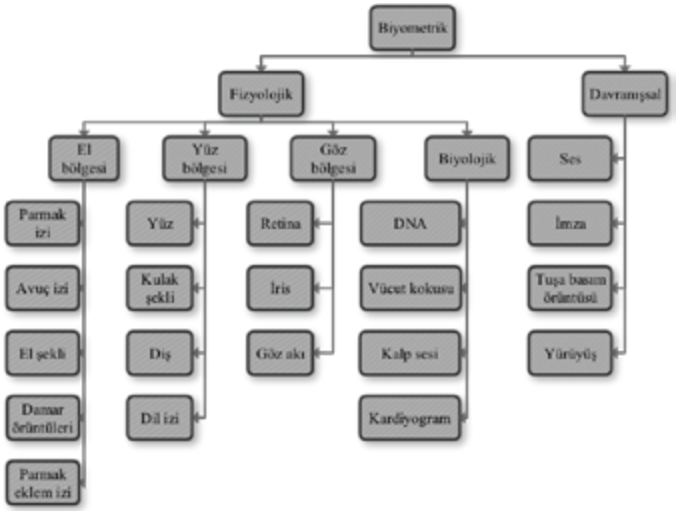
Biyometrik sistemler doğrulama amacıyla kullanılmalarının yanı sıra tanıma işlemlerini gerçekleştirebilmek üzere kullanılabilmektedir. Doğrulama sistemleri, veritabanında daha önceden kaydedilen bir şablon ile karşılaştırma yapılarak kişinin doğru kişi olup olmadığının doğrulanmasını

hedeflenmektedir. Bu sistemler erişim kontrolü, güvenlik, takip gibi amaçlara hizmet etmek üzere farklı uygulama alanlarına sahiptir. Bu uygulamalarda şüphesiz güvenlik önemli bir konuya sahiptir. Tanıma sistemlerinde ise veritabanında yer alan tüm şablonlar ile örnek veri karşılaştırılarak biyometrik özellikten kimlik tespiti yapılmaya çalışılmaktadır [3].

Biyometrik sistemlerde rol alan 3 farklı kullanıcı tanımlanabilmektedir. Kullanıcılar taklitçi, saldırgan veya yetkili kullanıcı olarak gruplandırılabilmektedir. Yetkili bir kişi gibi görünerek biyometrik sistemlere erişebilen ya da erişmeye çalışan kişiye taklitçi adı verilmektedir. Biyometrik sistemlere düzenlediği saldırılarla erişmeye çalışan ya da hizmet aksattırma işlemini hedefleyen kişi saldırgan olarak ifade edilmektedir. Yetkili kullanıcı ise biyometrik sistemlere erişim yetkisi olan kişidir [4].

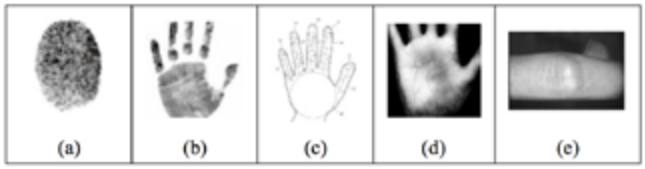
A. Biyometrik Sistem Karakterleri

Biyometrik sistemler bir algılayıcı ile dış ortamdan alınan biyometrik verinin tipine göre adlandırılmaktadır. Biyometrik sistemlere konu olan biyometrik karakterler fizyolojik ve davranışsal karakterler olmak üzere iki grupta ele alınmıştır. Fizyolojik biyometrik karakterler kategorisi altında Şekil 2'de görüldüğü gibi biyometrik karakterlerin yer aldığı bölgelere göre bir alt kategorizasyon yapılabilmektedir. Fizyolojik karakterler el, yüz ve göz bölgesinde bulunan fiziksel özelliklerimiz olabildikleri gibi DNA, vücut kokusu, kalp sesi ve kardiyogram gibi biyolojik özellikler de olabilmektedir.



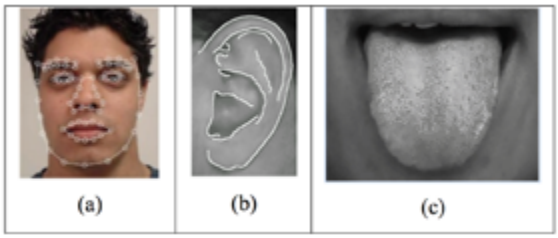
Şekil 2. Biyometrik sistemlerde kullanılan biyometrik karakterler [5]

Özellikle parmak izi olmak üzere el bölgesinin önemli doku bilgileri içermesi bu karakterler tabanında çalışan tanıma sistemlerinin yaygın şekilde kullanılmasına neden olmuştur. El bölgesinde ele alınabilecek karakterler Şekil 3'de gösterildiği gibi, bilinen en eski biyometrik karakter olarak ele alınan parmak izinin yanı sıra avuç izi, el geometrisi, damar örüntüleri ve parmak eklem izi gibi karakterlerdir. Parmak izi sistemlerinde sırt iskeleti, sırt örüntüleri, vadi ve geçit gibi öznitelikler dikkate alınarak eşleştirici bir skor elde ederek tanıma işlemi gerçekleştirilmeye çalışılmaktadır. Avuç içi karakterine dayalı çalışan sistemlerde temel çizgiler, kırışıklıklar, yoğunluk, avuç dokusu, ortalama ve varyans gibi öznitelikler kullanılmaktadır. Eller üzerinde yer alan damar izini tanımak üzere geliştirilen sistemlerde ise damar çatallanmaları gibi öznitelikler ayırt edici olmaktadır [5].



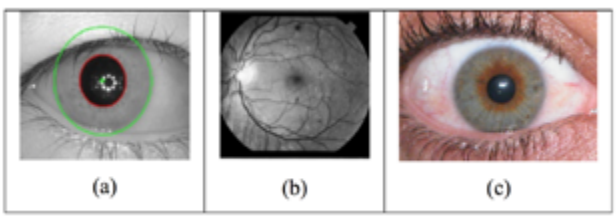
Şekil 3. El bölgesi biyometrik karakterleri (a) parmak izi (b) avuç izi (c) el geometrisi (d) damar izi (e) parmak eklem izi

Yüz bölgesi parmak izi gibi ayırt ediciliği yüksek bir biyometrik karakterlerden oluştuğu için en çok çalışılan bölgelerden biri olmuştur. Şekil 4'de biyometrik sistemlerde kullanılan yüz bölgesi karakterleri gösterilmiştir. Yüz bölgesinde en fazla ön plana çıkan biyometrik karakter yüz olmuştur. Yüz üzerinde gözler, ağız ve burun arası uzaklıklar gibi öznitelikler tabanında tanıma problemi çözümlenmeye çalışıldığı gibi tüm görüntü, yüz sınırları gibi farklı öznitelikler belirlenerek yüzden kimlik tespiti yapılmaya da çalışılmıştır. Kulak, şekli ve kepçe olarak ifade edilen doku yapısı itibariyle ayırt edici bir özellik olarak önerilmiştir [6]. Kulak için tanımlanan öznitelikler kulak boyu, genişliği, yüksekliği, rengi ve sınırları gibi öznitelikler olmuştur [5]. Dil izinin eşi olmayan bir biyometrik karakter olması bu karaktere dayalı olarak yapılan çalışmalar da vardır [7]. Bu çalışmalarda dil genişliği, kalınlığı, dokusu ve şekli gibi öznitelikler tanımlanmıştır [5].



Şekil 4. Yüz bölgesi biyometrik karakterleri (a) yüz (b) kulak (c) dil

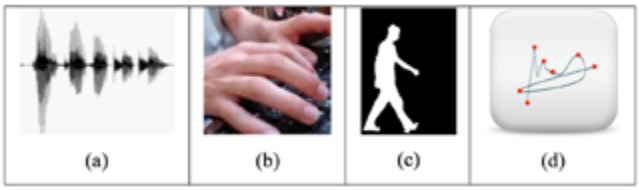
Göz bölgesinde yer alan biyometrik karakterlerin daha doğru, güvenilir ve sabit yapıda olmaları göz bölgesini tanımak üzere geliştirilen çok sayıda yaklaşıma neden olmuştur [5]. Göz bölgesinde ele alınan karakterler Şekil 5'de gösterilmektedir. İris tabanlı biyometrik sistemlerde renk, şekil ve iris dokusu özniteliklerinden yararlanılarak tanıma problemi çözümlenmeye çalışılmaktadır. İrisin ayırt edici özelliğinin aksine bu karaktere dayalı geliştirilen sistemlerin pahalı ve saldırıya açık olması sistemlerin kullanılabilmesine engel teşkil etmektedir [8]. Retina için damarlar ve optik alanı gibi öznitelikler ayırt edici olmaktadır. Gözün geri kalan kısmını oluşturan göz akı bölgesi ise göz damarlarından oluştuğu için bu damarlar yardımıyla eşleştirme yapılmaya çalışılmaktadır [5].



Şekil 5. Göz bölgesi biyometrik karakterleri (a) iris, (b) retina (c) göz akı

Biyolojik karakterler, medikal sensörler ile alınan DNA, vücut kokusu ve kalp sesi gibi kişiye özgü olarak nitelendirilen biyolojik bulgulardır. DNA, kişiye özgü olarak kullanılabilecek

en iyi ayırt edici biyometrik karakterlerden biridir. Kişiden alınan saç, kan ve tırnak gibi örneklerden DNA kodu kolaylıkla elde edilebilmektedir [5]. Bu sebeple başkasına ait DNA kodu, bir kişi tarafından kolaylıkla elde edilebilecek bir bilgidir. DNA kodunu değerlendirebilecek bir uzmana olan gereksinim DNA tabanında çalışabilecek gerçek zamanlı uygulamaları mümkün kılmamaktadır [6]. Vücut kokusu, bir sensör aracılığıyla alınarak sınıflandırılabilecek olan bir biyometrik karakterdir. Alınan kokuyu sınıflandırabilen e-nose olarak adlandırılan sistemler mevcuttur [5].



Şekil 6. Davranışsal biyometrik karakterler (a) ses (b) tuşa basma örüntüsü (c) yürüyüş tarzı örüntüsü (d) imza

Davranışsal karakterler insanların kendilerini ifade edebildikleri, ayırt edici olan ses, imza, tuşa basım gibi davranışlarına ait özellikleridir. Davranışsal karakterler olarak nitelendirilen biyometriklere Şekil 6'da yer verilmiştir. Davranışsal karakterler tabanında en yaygın şekilde kullanılan sistemler ses, tuşa basım ve imza tanıma sistemleridir. Bu sistemler fiziksel ve biyolojik karakterlere dayalı sistemlere göre herhangi bir harici donanıma gerek duymadıkları için daha avantajlıdır. Davranışsal karakterlerden biri olan sesi tanımak üzere spektrum, ritim, durak ve enerji gibi öznitelikler tanımlanmıştır. Tuşa basım örüntüsü, davranışın analiz edilerek kimlik tespitini sağlayan karakterlerden biridir. Bu sistemler, tuşa basım işleminin kişiye özgü bir ritme sahip olduğu varsayımı üzerine ortaya çıkmıştır. Tuşa basım örüntüsüne dayalı sistemlerde tuşa basım ve tuş bırakma anına bağlı olarak tanıma işlemi gerçekleştirilmeye çalışılmaktadır [9]. Bu nedenle tuşa basım, bekleme, gecikme süreleri ile hız gibi öznitelikler bu sistemlerde ayırt edici özellikler olarak kişileri birbirinden ayırt etmek üzere kullanılmaktadır. Yürüyüş tarzı tabanlı çalışan sistemler ise hız, bir adım mesafesi ve siluet şekli gibi öznitelikler ile değerlendirme yapabilmektedir. İmza, üzerinde çalışılan farklı bir davranışsal karakterdir. Bu karakteri kullanan sistemlerde imza şekli, kalem yönü, ivme ve imza uzunluğu gibi hususlara dikkat edilmektedir [5].

Biyometrik Özellik	Kullanım kolaylığı	Sorunlar	Doğruluk	Güvenlik gereksinimi
Parmak izi	Yüksek	Kuruluk, kir ve yaş	Yüksek	Yüksek
El geometrisi	Yüksek	Elde hasar, yaş	Yüksek	Orta
Retina	Düşük	Gözlük	Çok Yüksek	Yüksek
İris	Orta	Işık	Çok Yüksek	Çok yüksek
Yüz	Orta	Işık, yaş, gözlük, saç	Yüksek	Orta
İmza	Yüksek	İmza değişikliği	Yüksek	Orta
Ses	Yüksek	Gürültü, ses kısıklığı	Yüksek	Orta

Tablo I - Biyometrik karakter karşılaştırmaları [1]

Biyometrik sistemlerde yaygın olarak kullanılan özellikler parmak izi, avuç izi, retina, iris, yüz, imza ve ses olmuştur. Biyometrik sistemlerde kullanılan biyometrik karakterlerin her birinin kendine özgü güçlü ve zayıf yönleri bulunmaktadır. Parmak izi, düşük maliyetli ve diğer karakterlere göre daha ayırt edici bir özellik olması nedeniyle çok sayıda sistemde kullanılmıştır. Parmak izi, kontrollü bir çevrede yeterli sayıda eğitim adımın sonunda tanıma yapabilecek akıllı ev sistemleri için iyi bir özellik olarak nitelendirilmiştir. Retina tanıma sistemleri performansı yüksek olan sistemler olmalarına rağmen belirli bir noktaya odaklanma gereksinimleri nedeniyle yaygın olarak kullanılan biyometrik sistemler olamamıştır. Yüz tanıma sistemleri yüz karakterlerine dayalı olarak analiz yapan sistemlerdir. Bir algılayıcı ile çalışan bu sistemlerde çok sayıda uygulamada kullanılmıştır. Ses tanıma sistemleri ise herhangi bir ek donanıma gerek duymaksızın sesi metin dosyalarına dönüştürerek işlem yapan sistemler oldukları için ilgi görmüşlerdir. Bu sistemlerin de gürültü gibi dış etkenlerden etkilenmesi sistemlerin pratikte kullanılabilirliğini olumsuz yönde etkilemektedir [10]. Biyometrik karakterlerin özelliklerine göre yapılan bir karşılaştırmaya Tablo I'de yer verilmiştir. Tabloda karakterler tabanında çalışan sistemlerin kullanım kolaylığı, karşılaşılabilecekleri sorunlar, performansları ve güvenlik gereksinimleri kıyaslanmıştır. Karakterlerin bu özellikleri dikkate alınarak uygulamaya yönelik biyometrik sistemler ortaya çıkmıştır. Parmak izine dayalı biyometrik sistemler, kullanım kolaylığı ve performanslarına rağmen parmak izi yüzeyinde olabilecek kuruluk, yaş ve kir gibi durumlar karşısında başarısız olmaktadır. Yüz karakteri için tablo incelendiğinde ise yüz tanıma sistemlerinin kullanım kolaylığı orta, performansı yüksek, güvenlik gereksinimi ise orta olarak nitelendirildiği görülmektedir. Yüz tanıma sistemlerinin ışık varyasyonları, saç, gözlük gibi aksesuarlar ile yaşa karşı hassas olmaları dikkat çekmektedir.

Biyometrik olarak kullanılabilecek karakterleri belirleyebilmek üzere ayırt edicilik, evrensellik, kalıcılık ve ölçülebilirlik gereksinimleri sağlanmalıdır. Dikkat edilmesi gereken diğer unsurlar ise performans, kabul edilebilirlik ve aldatılabilirliktir. Ayırt edicilik, farklı iki kişide aynı karakterin aynı olma olasılığının neredeyse sıfır olmasıdır. Evrensellik, tüm insanların biyometrik veri olarak kullanılacak karaktere sahip olmasıdır. Kalıcılık, biyometrik karakterin zaman içinde değişiminin söz konusu olmamasıdır. Performans, biyometrik sistemlerde kullanılan biyometrik karakterin yüksek tanımlama başarısına sahip olmasıdır. Ölçülebilirlik, biyometrik karakterin niceliksel olarak ölçülebilir olması anlamına gelmektedir. Kabul edilebilirlik, söz konusu biyometrik karakterin insanlar tarafından biyometrik veri olarak görülebilmesidir. Aldatılabilirlik ise biyometrik sistemin kolaylıkla kandırılarak yetkisiz kişiler tarafından erişilebilmesidir [11].

	Ayırt edicilik	Evrensellik	Kalıcılık	Ölçülebilirlik	Performans	Kabul edilebilirlik	Aldatılabilirlik
Yüz	D	Y	O	Y	D	Y	Y
İris	Y	Y	Y	O	Y	D	D
Tuşa basım	D	D	D	O	D	O	O
Avuç izi	O	O	Y	O	Y	O	O
Parmak izi	Y	O	Y	O	Y	O	O
Kulak	O	O	Y	O	O	Y	O
DNA	Y	Y	Y	D	Y	D	D
Retina	Y	Y	O	D	Y	D	D
İmza	D	D	D	Y	D	Y	Y
Ses	D	O	D	O	D	Y	Y
Yürüyüş	D	O	D	G	D	Y	O
Vücut kokusu	D	Y	Y	D	D	O	O

Tablo II - Biyometrik karakter karşılaştırmaları [11]

Biyometrik sistemler için ayırt edici olan bir karakterin performansı hız ve doğruluk bakımından yeterli değil ise uygulanabilir olamayacaktır. Pratikte uygulanacak biyometrik bir karakterin gereksinimleri sağlamasına ek olarak tüm bu unsurları taşıması gerekmektedir [6]. Bu sebeple biyometrik sistemlerde doğrulanacak karakterler sistem biyometrik karakter ölçütlerine ek olarak uygulama alanları da dikkate alınarak belirlenmelidir. Tablo II'de farklı biyometrik karakterlerin biyometrik karakter ölçütlerine göre karşılaştırmalarına yer verilmiştir. Tablo üzerinde Y ile ifade edilen ölçütler Yüksek, O ile ifade edilenler Orta ve D ile ifade edilenler ise Düşük anlamına gelmektedir. Tabloda görülebildiği gibi yüz ayırt ediciliği düşük, tüm insanlarda aynı olan bir karakter olması sebebiyle evrenselliği yüksek, yaşa bağlı değişken bir karakter olması sebebiyle kalıcılığı orta, niceliksel olarak değerlendirilebileceği için ölçülebilirliği yüksek olarak değerlendirilmiştir. Ayrıca bir biyometrik veri olarak kabul edilebilirliği yüksek ve başka bir kişiye ait yüz ile taklit edilebilir olabileceğinden aldatılabilirliği yüksek olarak ifade edilmiştir.

B. Biyometrik Teknoloji Uygulama Alanları

Biyometrik sistemlerin uygulama alanları sınır kontrol, suçlu tanımlama, erişim kontrolü, e-ticaret, bilgisayar oturum açma işlemleri, kimlik kartları, pasaportlar, görüntüleme sistemleri, akıllı telefonlarda kullanıcı doğrulama, kalabalık görüntüleme, elektronik bankacılık, video izleme ve adli bilişim gibi alanlardır. Biyometrik sistemlerin uygulama alanları Tablo III'de özetlenmiştir.

	Parmak izi	Avuç izi	El geometrisi	El damar örüntüleri	Parmak eklem izi	Yüz	Kulak	Dil izi	İris	Retina	Göz akı	Ses	Tuşa basma örüntüsü	Yürüyüş tarzı	İmza
Sınır kontrol	×					×			×	×					
Adli bilişim	×					×								×	
Suçlu tanıma	×					×			×	×				×	
Kimlik kartı	×					×			×						
Pasaport	×					×			×						
Bilgisayar oturum açma işlemleri	×	×	×	×	×	×	×	×	×		×	×	×		×
Erişim kontrolü	×	×	×	×	×	×	×	×	×	×	×	×			×
E-ticaret	×	×	×	×	×	×	×	×	×		×	×			×
Akıllı telefon	×	×	×	×								×	×		
Görüntüleme sistemleri	×	×	×		×		×		×					×	
Video izleme						×								×	
Kayıp çocuk tanıma	×					×			×						
Kalabalık görüntüleme						×								×	
E-banka															×

Tablo III - Biyometrik sistemlerin uygulama alanları [5]

Tablo III'de görüldüğü üzere erişim kontrolü uygulamalarında hemen her karakter tabanında geliştirilen sistemler kullanılabilmektedir. Sınır kontrol uygulamalarında ise en yaygın kullanılan biyometrik karakterler olan parmak izi, yüz, iris ve retina kullanılmaktadır. Yürüyüş tarzı karakterine dayalı sistemlerin ise belirli bir mesafeden tanımaya imkan verdiği için video görüntüleri üzerinde adli bilişim, suçlu tanımlama, görüntüleme sistemleri, kalabalık görüntüleme ve video izleme uygulamalarında kullanıldığı görülmüştür. Kimlik kartı ve pasaport kartlarında ise yüz, parmak izi ve iris gibi karakterler kullanılmıştır.

Biyometrik teknolojilere duyulan gereksinim bu alanda bir endüstrinin ortaya çıkmasına neden olmuştur. Bu endüstride farklı teknolojiler ile 2013 ve 2019 yılları arasında elde edilen ve beklenen gelirler BBC raporlarında [12] ifade edilmiştir. 2013 yılı verileri ile oluşturulan raporda piyasada en çok kullanılan biyometrik teknoloji parmak izi teknolojileri olmuştur. Tüm teknolojilerin özellikle de parmak izi teknolojilerine ait pazar payının önemli derecede arttığı görülmektedir. Bu raporda 2013 yılında 8,7 milyar dolara varacağı öngörülen pazar payının 2014 yılında yıllık % 19,8 artışla 11,2 milyar dolar, 2019 yılında ise 27,5 milyar dolar olacağı tahmin edilmiştir.

III. BİYOMETRİK SİSTEMLERDE GÜVENLİK

Geleneksel doğrulama mekanizmalarındaki verilerin çalınabilme ve kaybedilebilme gibi sorunlarının aksine biyometrik sistemlerin taklit edilemez ve kopyalanamaz olarak görülmesi bu sistemlerin en güvenli doğrulama mekanizmaları olarak vurgulanmasına neden olmuştur. Biyometrik sistemlerin geleneksel doğrulama mekanizmalarına göre üstünlüklerine karşın bazı problemleri mevcuttur. İncelenen pek çok çalışmada

biyometrik sistemlerin düşünüldüğü kadar güvenli sistemler olmadıkları görülmüştür. 2009 yılında gerçekleştirilen Black Hat konferansında bir araştırma grubu tarafından Asus, Toshiba ve Lenova dizüstü bilgisayarların bazı modellerinde bulunan gömülü biyometrik sistemlerde (Asus SmartLogon V1.0005, Toshiba Face Recognition 2.0.2.32 ve Lenova Veriface III) taklit edilen biyometrik veriler ile sistemlere kolaylıkla erişilebildiği gösterilmiştir [13]. Apple'ın parmak izi sensörüne sahip ilk akıllı telefonu olan iPhone 5s'e yapıştırıcı bir ürünle kopyalanan bir parmakla bir başkası tarafından erişilebileceği gösterilmiştir [14].

Birçok çalışmada saldırı vektörleri olarak tanımlanan saldırı noktalarından biyometrik sistemlere erişilerek, bu sistemlerde güvenlik açısından tehditler oluşturulabilmektedir. Olası güvenlik açıklarının önceden bilinmesi bu saldırı noktalarına karşı önlem alınarak sistemlerin güvenli kılınması açısından önem teşkil etmektedir. Bu gereksinim üzerine literatür incelendiğinde biyometrik sistemlerde saldırı vektörlerini tanımlamak üzere yapılan çeşitli çalışmalara rastlanmıştır. Bu alandaki çalışmalardan ilkini Ratha ve arkadaşları [15] tarafından yapılan çalışma oluşturmaktadır. Ratha ve arkadaşları yaptıkları çalışmada biyometrik sistemlerde güvenlik unsuru üzerine odaklanmışlardır. Bu çalışmalarında biyometrik sistemlerin saldırılara maruz kalabileceği güvenlik açıklarına değinmişlerdir. Bu amaçla bir saldırı noktaları modeli önermişlerdir. Önerilen model 8 farklı saldırı vektöründen oluşmaktadır. Modelde tanımlanan saldırı vektörleri;

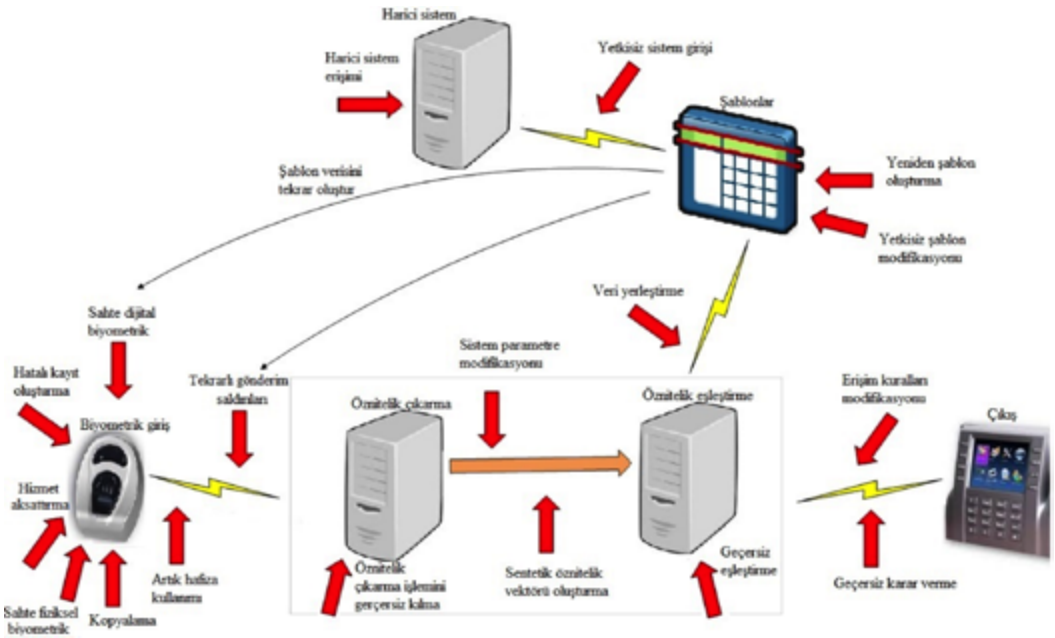
1. Sahte biyometrik,
2. Tekrarlı gönderim,
3. Öznitelik çıkarımının geçersiz kılınması,
4. Öznitelik vektörünün değiştirilmesi,
5. Eşleştiricinin etkisiz kılınması,
6. Veritabanına yetkisiz erişim,
7. Şablon verisinin değiştirilmesi,
8. Eşleştirici sonucunun değiştirilmesidir.

Bu modele ilişkin görsele Şekil 7'de yer verilmiştir. Şekilde kırmızı oklarla ifade edilen adımlardan sisteme yetkisiz kişiler tarafından erişilebileceği vurgulanmıştır.

Şekil 7. Ratha ve arkadaşlarının saldırı vektörleri modeli [15]

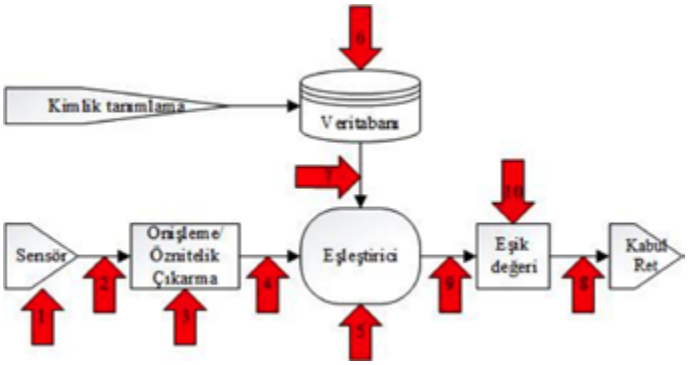
Cukic ve Bartlow [16] yaptıkları çalışma ile Ratha ve arkadaşlarının [15] yaptıkları çalışmaya kıyasla daha kapsamlı bir model sunmuşlardır. Çalışmalarında biyometrik sistemler için bir saldırı ağacı modellemişlerdir. Bu model ile 20 olası saldırı vektörü ile 22 güvenlik açığı tanımlanmıştır. Jain ve arkadaşları [17] kılçık modeli diye tanımladıkları modelleriyle biyometrik sistemlerin maruz kalabileceği problemleri kategorize ederek ele almışlardır. Modellerinde saldırıları, düşman ve sıfır çaba saldırıları olarak iki sınıfa ayırmışlardır. Roberts [4] yaptığı çalışmada tanımladığı 18 olası saldırı vektöründen oluşan model ile bu saldırı vektörlerine karşı savunma yollarını ele almıştır. Roberts'ın modeline Şekil 8'de yer verilmiştir. Modelde görülebildiği gibi;

- Sahte dijital biyometrik
- Hatalı kayıt oluşturma
- Hizmet aksattırma
- Sahte fiziksel biyometrik
- Kopyalama
- Artık hafıza kullanımı
- Tekrarlı gönderim
- Öznitelik çıkarma işleminin geçersiz kılınması
- Sistem parametre modifikasyonu
- Sentetik öznitelik vektörü oluşturma
- Harici sistem erişimi
- Yetkisiz şablon modifikasyonu
- Veri yerleştirme
- Geçersiz eşleştirme
- Erişim kuralları modifikasyonu
- Geçersiz karar verme



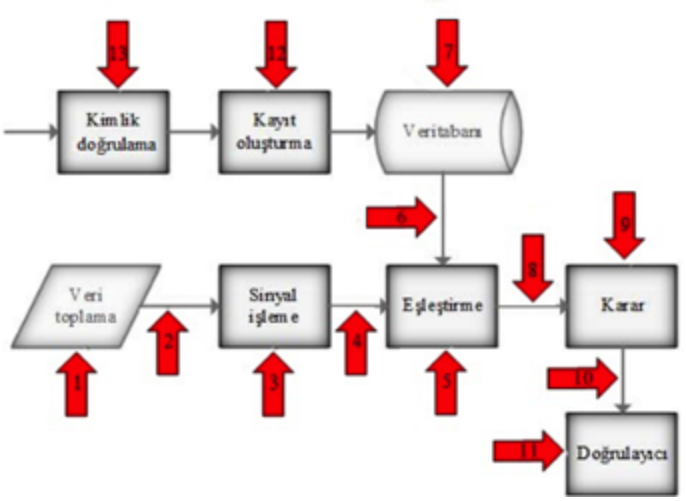
Şekil 8. Roberts'ın saldırı vektörleri modeli [4]

saldırı vektörleri tanımlanmıştır. Bu çalışmada yukarıda da sıralanan güvenlik sorunlarına karşı biyometrik sistemleri güvende tutabilmek üzere bazı savunma yolları önerilmiştir. Bunlar güvenlik yanıtı, rastsal biyometrik veri, hafızada tutma, gerçeklik tespiti, çoklu biyometrik, çok karakteristikli biyometrik, faktörlü doğrulama, ayırt edici biyometrik, veri bütünlüğü, şifreleme, dijital imza, şablon bütünlüğü, iptal edilebilir biyometrik, donanım bütünlüğü, ağ güvenliği, fiziksel güvenlik, aktivite günlüğü ve politikalar gibi yaklaşımlar olmuştur.



Şekil 9. Galbally saldırı vektörleri modeli [18]

Galbally [18], daha önce tanımlanan kılçık modelinde [17] yapılan kategorizasyonu dikkate alarak düşman saldırılarını doğrudan ve dolaylı saldırılar olarak kategorize ederek ele almıştır. Bu kategorizasyon 10 saldırı noktasından oluşan model üzerinde ifade edilmiştir. İlgili modele Şekil 9'da yer verilmiştir. Model incelendiğinde Ratha ve arkadaşlarının [15] geliştirdiği modele benzerlik dikkat çekmektedir. Bu çalışmada 4 ve 5 ile gösterilen saldırı noktaları 9 ve 10 ile tekrar ifade edilerek model detaylandırılmıştır.



Şekil 10. Alaswad saldırı vektörleri modeli [19]

Alaswad ve arkadaşları [19], 13 farklı saldırı noktası ile modellerini ifade etmişlerdir. Bu saldırı modeline Şekil 10'da gösterilmektedir. Şekilde görüldüğü gibi model bir biyometrik sistemde yer alan her adım için geliştirilmiştir. Bu çalışmada saldırı noktaları her bir saldırı seviyesi için ele alınmıştır. Her bir adımda ayrı olarak ele alınan saldırılara karşı alınabilecek savunma yollarına da yer verilmiştir.

IV. SONUÇ VE DEĞERLENDİRMELER

Bilgi teknolojilerindeki gelişmeler geleneksel doğrulama mekanizmalarının yerini biyometrik veri tabanlı doğrulama mekanizmalarının almasına neden olmuştur. Biyometrik verilerin kaybedilemez, kopyalanamaz ve unutulamaz gibi özellikleri nedeniyle biyometrik teknolojiler üzerine bir endüstri ortaya çıkmıştır. Makalede ele alındığı gibi farklı biyometrik karakterlere dayalı biyometrik teknolojiler ortaya çıkmıştır. Biyometrik teknolojilerin özellikleri dikkate alınarak uygulama alanları da farklılık gösterebilmektedir. Bu teknolojilerin hayatımızın bir parçası haline gelmesi, bu sistemlerin varsayıldığı kadar güvenli sistemler olup olmadıkları sorusunu gündeme getirmiştir. Literatürde biyometrik sistemlerin düşünüldüğü gibi güvenli sistemler olmadıkları örnekleri ile gösterilmiştir. Bu nedenle sunulan çalışmada biyometrik teknolojilerde güvenlik kavramına odaklanılmıştır. Biyometrik sistemlerde güvenlik konusunu ele alan çalışmalar incelendiğinde sistemlerin maruz kalabileceği saldırılar üzerine çeşitli çalışmalar yapıldığı görülmüştür. Bu çalışmalarda sunulan saldırı noktaları modellerindeki olası saldırılar değerlendirilmiştir.

Sunulan çalışma ile biyometrik teknolojilerin saldırılara ne kadar açık olduğu gözler önüne serilerek, ülkemizde de giderek kullanımı artan biyometrik sistemlerin güvenliği konusunda bir farkındalık oluşturulmaya çalışılmıştır. Biyometrik teknolojilerin, geleneksel doğrulama mekanizmalarına göre çok daha güvenli teknolojiler olarak nitelendirilmelerine rağmen beklenildiği kadar güvenli olmadıkları literatürdeki çalışmalar incelenerek ortaya konmuştur. Ayrıca, biyometrik karakterlerin kopyalanmasının, biyometriklerin değiştirilemez yapıları nedeniyle, şifre ve kullanıcı adı gibi bilgilerin bir başkası tarafından elde edilmesinden çok daha büyük tehdit oluşturduğu tespit edilmiştir. Bu nedenle, biyometrik sistemlerde güvenlik önlemleri alınarak gerekli politikalar sağlanması gerektiği sonucuna ulaşılmaktadır. Gelecek çalışmalarda, günümüzde kullanımı hızla artan biyometrik sistemlerin daha güvenli ve performanslı olmasını sağlayacak yaklaşımların üzerinde daha fazla durulması gerektiği değerlendirilmektedir.

KAYNAKÇA

[1] S. Liu, M. Silverman, "A practical guide to biometric security technology", IT Professional, vol. 3, no. 1, pp. 27–32, 2001.

[2] K. Delac, M. Grgic, "A survey of biometric recognition methods", in 46th International Symposium Electronics in Marine, Crotia, pp.184-193, 2004.

[3] S. Prabhakar, S. Pankanti, A. K. Jain, "Biometric recognition: Security and privacy concerns", IEEE Security & Privacy, no. 2, pp. 33–42, 2003.

[4] C. Roberts, "Biometric attack vectors and defences," Computers & Security, vol. 26, no. 1, pp. 14–25, 2007.

[5] J. Unar, W. Senga, A. Abbasia, " A review of biometric technology along with trends and prospects", Pattern recognition, vol. 47, pp. 2673–2688, 2014.

[6] A. K. Jain, A. Ross, S. Prabhakar, "An introduction to biometric recognition", IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 4–20, 2004.

[7] D. Zhang, Z. Liu, J. Q. Yan, P. F. Shi, "Tongue-print: A novel biometrics pattern", In Advances in Biometrics, vol.

4642, pp. 1174–1183, 2007.

[8] F. Monrose, A. Rubin, "Authentication via keystroke Dynamics", In Proceedings of the 4th ACM conference on Computer and communications security, Switzerland, pp. 48–56, 1997.

[9] P. R. Dholi, K. P. Chaudhari, "Typing pattern recognition using keystroke Dynamics", Mobile Communication and Power Engineering, vol. 296, pp. 275-280, 2013.

[10] S. Liu, M. Silverman, "A practical guide to biometric security technology", IT Professional, vol. 3, no. 1, pp. 27-32, 2001.

[11] A. K. Jain, R. Bolle, S. Pankanti, "Biometrics: personal identification in networked society", Springer Science & Business Media, 1999.

[12] BBC, "Biometrics: Technologies and Global Markets", [Çevrimiçi]: http://www.bccresearch.com/market-research/information-technology/ biometrics-technologies-ift042d.html. [Erişim Tarihi: 14 Ağustos 2015]

[13] Z. Zhang, D. Yi, Z. Lei, S. Z. Li, "Face liveness detection by learning multispectral reflectance distributions", IEEE International Conference on. Automatic Face & Gesture Recognition and Workshops (FG 2011), USA, pp. 436–441, 2011.

[14] A. Hadid, "Face biometrics under spoofing attacks: vulnerabilities, countermeasures, open issues, and research directions," IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), USA, pp. 113–118, 2014.

[15] N. K. Ratha, J. H. Connell, R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," IBM systems Journal, vol. 40, no. 3, pp. 614–634, 2001.

[16] B. Cukic, N. Bartlow, "Biometric system threats and countermeasures: a risk based approach," in Proceedings of the Biometric Consortium Conference (BCC05), USA, 2005.

[17] A. K. Jain, A. Ross, S. Pankanti, "Biometrics: a tool for information security," IEEE Transactions on Information Forensics and Security, vol. 1, no. 2, pp. 125–143, 2006.

[18] J. Galbally, "Vulnerabilities and attack protection in security systems based on biometric recognition," IEscuela Politecnica Superior, Universidad Autónoma de Madrid, PhD Tezi, 2009.

[19] A. O. Alaswad, A. H. Montaser, F. E. Mohamad, "Vulnerabilities of biometric authentication 'threats and countermeasures'," International Journal of Information & Computation Technology, vol. 4, no. 10, pp. 947–958, 2014.

SECURING BIOMETRIC FACE IMAGES VIA STEGANOGRAPHY FOR QR CODE

Sercan Aygün, Muammer Akçay

Abstract — Recent evaluations in technology require deep awareness in terms of security. Both the users and developers need to bear in mind that there is no perfect way to be in secure. At least by using the brute force approach, every password has its own way to be deciphered, even though it takes years with supercomputers. Passwords can be changed with the ease of use. Whereas, there are some other authentication methods that serve quite convenient advantages like portability. Biometric features of our own are the part of biological human body which have admirable structure to be distinguished perfectly. On the other hand, these unique features are to be secured carefully in the application based usage the reason why cannot be changed physically like passwords. Therefore, in this paper it is proposed to use quick response code-QR for biometric face features to be ciphered by steganography and Relational Bit Operator-RBO. Recent access control systems require aforementioned biometric structures especially for e-government and e-passport applications. This paper presents a new approach to be used in application related biometric face image features in tiles to be safely transmitted. Consequently, by using an operator, first a pattern is obtained in the light of image processing and cryptography. After, the pattern is mixed up randomly by saving the actual positions of each element. Finally, the new pattern, in other words new image, is embedded into QR code by inserting the actual parameters of each element via steganography, too. Following sections present relatively the introduction, literature review, face biometry related operator, proposed method and conclusion at the end.

Index Terms — Biometry, cryptography, face biometry, QR code, relational bit operator, steganography.

I. INTRODUCTION

Security plays an important role in the human history. From the ancient times to the World Wars there were many stories about the cryptographically constructed approaches and systems. Also with recent technological developments, cyber security has become very crucial area that appeals researchers. In this work, it is aimed to propose a model that hides biometric face image features into QR code that is hard to be obtained by attackers. In daily technologies, biometric authentication provides more suitable ways for users to be supplied by easy to carry personal identification indicators. These indicators are every time with the human, the reason why they are the part of human body. There is no need to carry extra cards or remember things like passwords. Hardware engineers add some biometric sensors like fingerprint sensors into the personal computers and mobile phones to make access controls more robust, too. Besides, banking also uses bio-signatures during ATM machine operations. Thus, public is getting used to employ these access methods in their daily authentication systems. Furthermore, some other governmental areas are requiring biometric face and finger data for verification purposes. For instance in Turkey, the new biometric data inserted national ID cards are going to be in use near future. Therefore, handling these significant biometric data must be securely implemented.

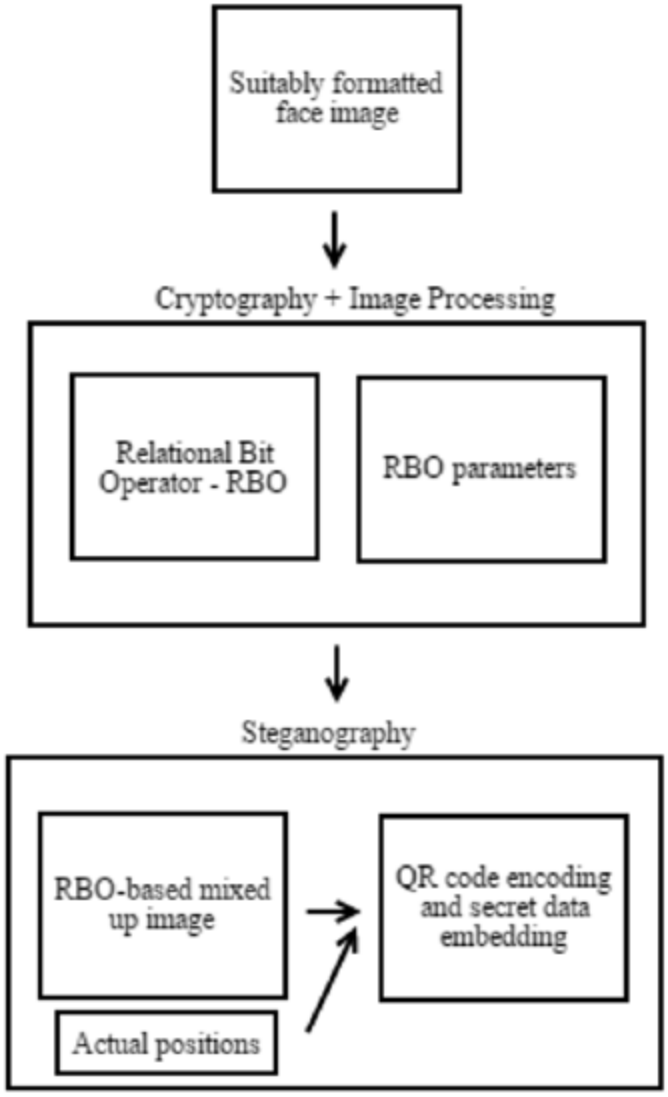


Fig. 1. Summary of the general approaches that are used in this paper.

In Fig. 1, it is illustrated the basic points of this work that are gathered together. Input image that obeys the biometric face picture format is inputted, which is cropped and its color conversion managed. After, the feature extraction process starts. In that step, getting features both must be meaningful for image analysis knowledge and must be cryptographically manageable. Therefore, a convenient approach is needed which is going to be presented in Section III. Then, the operator related new image –the image feature data– is obtained in matrix format. This is then randomly mixed up by caching the first positions of each row-column based pixel element. Up to that point, image processing is deployed in the frame of ciphering data where all mixing data will be placed in QR codes by hiding actual row-column based location keys into QRs via steganography.

II. RELATED WORKS

In the literature, there are some examples for the usage of QR code to be an element as a security level increaser. Chen and Wang propose to hide some data in QR code by

considering some of its useless parts [1]. In that work, the lossy and lossless data are categorized. One can be partially lost, there is no matter, because during reconstruction, error correction feature of QR code works well. They add that the proposed scheme is resistant to JPEG attacks. 25% is the well enough rate for error correction on their work. On the other hand, Zigomitros and Patsakis make vice versa proposal that embedding QR coded data into one another image [2]. This is good for compression applications according to their final draw. Besides, during web searches, embedding some data into several images should speed up the search time by checking the text data inside of QR. Chung et al. proposes a similar approach to [1] by embedding lossless text data into QR code and using the regular areas in the QR code even like cropping it for reduction of dimension [3]. Therefore, it can be understood that there are some parts which are omittable and can be used for saving the key of the ciphered data.

There are also some other applications of QR code like to be used in medical applications. Chang et al. uses QR code in a hospital environment to embed secret data of patients etc. [4]. Maheswari and Hemanth give Fresnelet Transform and Least Significant Bit (LSB) related to steganography knowledge where LSB is the one also considered in this study [5]. Ramesh et al. uses the general trend as embedding text data into QR code, but this time by using Discrete Wavelet Transform (DWT) to make encoding and decoding operation in frequency domain. It is proposed that the highly secure output is obtained [6]. In the literature, the general trend is to embed text data into QR and to handle this data. From the inspiration of that, in this work it is going to be added image data bytes up into the QR which is different from the standard text data.

For the visual cryptography, one of the state-of-art study [7] exhibits how to secure the biometric data in visual cryptography by Ross and Othman. In Visual Cryptography Scheme (VCS), the original binary image T is encrypted into n images where n is the number of noisy images. The scheme, namely k-out-of-n uses the Boolean operation as follows:

$$T = S_{h1} \oplus S_{h2} \oplus S_{h3} \oplus \ldots \oplus S_{hk} \qquad (1)$$

Reconstruction of original image T is only possible under the condition that k or more out of n images will be used. Encryption for each pixel is done by using subpixels, namely shares, and the independent random choices. Ross and Othman uses one step higher approach of VCS named as Gray-Level Extended Visual Cryptography Scheme (GEVCS) for face images. There are one private and two host face images. Host images look alike private image in terms of geometry and appearance. Transparency, the number of white pixels, is obtained by the control of subpixels in shares of host images during encryption which is done via Active Appearance Model (AAM). This model contains training set annotation, texture model and combined AAM building. After, selection of the hosts, image registration & cropping, secret encryption & reconstruction are handled. In the final phase, GEVCS does secure private image O in the two host images, Hs1 and Hs1, by resultants S1 and S2.

III. FACE BIOMETRY AND RELATIONAL BIT OPERATOR

Authentication is the one process that the face recognition is employed. Therefore, biometry plays an important role. In this section, biometry related knowledge and a new operator will be presented.

A. Biometry

It is better to understand what the biometry is and its subdivisions to make it clear before security related issues. Every human being has its own genetic structure which is perfectly different from one other. This phenomenon makes every person unique in the way of behavior, appearance, character, even illnesses, habits, and so on. Biometric data from biological being also the one makes it peerless. Biometry can be thought of three ways: physical, behavioral and the chemical. All the features used for authentication purposes are in the one of these subdivisions like fingerprint, face, signature, DNA, ear etc. Face biometry is relatively easy to be captured via image acquisition. Therefore, this paper uses the face biometry to be secured.

B. Relational Bit Operator

Extracting feature from biometric images is an important step in image processing related approaches. There are some methods to obtain a pattern from face images. Local Binary Pattern (LBP) is the one proposed by Ojala used for both in face detection algorithms and for texture analysis. Referencing the center pixel by considering the neighbors a pattern is captured. The size of the operator that effects neighborhood and the direction of the operator for processing each neighboring pixel are crucial parameters [8].

From the inspiration of LBP, a relatively new operator is proposed to search for relations of each neighbors. Fig. 2 shows the rectangle shape operator which is the 1 step size and 8-neighbouring based.

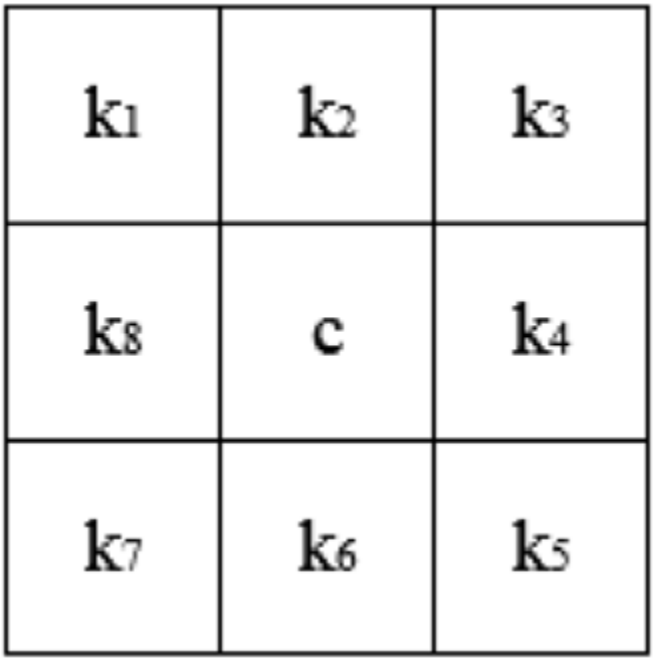
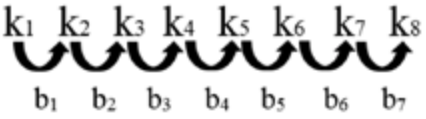


Fig. 2. RBO-Relational Bit Operator illustration with center pixel c, and all kx neighbors.

In Fig. 2, k's are all the neighbors in that rectangle shape and it is 1–pixel long distanced operator. The relations based on the neighbors can be calculated in some ways like the one in the following pattern as b1b2b3b4b5b6b7 format. The starting pixel and the rotation can be any of the possibility. Fig. 3 illustrates some examples of them. Next, kx neighbors are checked according to their numerical values. Then the pattern of bs in 0's and 1's is obtained:



Center pixel is not in the scope of interest but the all neighbors are processed. For example, in Fig. 3, the first square tells the starting point from k1 till the k8 same as previous pattern. Whereas, the pattern can be obtained from any other starting pixel in counter-clockwise direction as in the second operator. The crucial point is to use the fixed approach for all pixels of raw face image and save parameters like rotation, starting pixel etc. as the key elements.

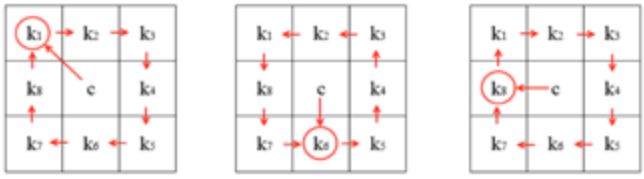


Fig. 3. Some RBO examples of possible starting pixels and rotation.

Numerically, it is considered if k1 > k2, the pattern is taken as 0 where it represents that there is a decreasing behavior, otherwise if it is k1 < k2 then the related bit is taken as 1. Fig. 4 shows a numerical example related to the Relational Bit Operator.

211	71	13
58	67	110
42	9	98

Fig. 4. A real valued piece of image that has 8-bit gray level pixel values to be processed via RBO, k1 is the starting pixel and the rotation is clock-wise.

k1k2k3k4k5k6k7k8 = 211 71 13 110 98 9 42 58
b1b2b3b4b5b6b7 = (0010011)2 = (19)10

The number 19 is now stored for operator based matrix instead of center pixel value 67. By that way, the feature of face image is preserved but it cannot give any meaningful information to attackers. Even, there will be some other security increasing approaches that are explained in the section of proposed system.

IV. OTHER CONCEPTS

A. QR Code

Coding some information into a visual material like barcode brings some remarkable advantages as seen in shopping environment. All products in supermarkets have barcode on them to be scanned by the cashier. The barcode has 1 dimensional approach, whereas QR code is the 2D which was invented by the Japanese corporation Denso Wave. QR image has white and black colors mostly, even though recent QR codes can have some colorful visual pictures on it. The encoding operation is done vertically and horizontally, therefore more data than 1D barcoding can be processed. Data can be text, number, URL, or even Kanji characters, too. In QR image there are 3 finder pattern that the camera can understand the horizontal or vertical positions of the image. The greatest amount of data is 2953 bytes as binary (8 bits) format that can be stored in QR. There is error correction possibility in QR code with 4 levels that inspires this work to use some areas of QR code for steganography [9].

B. Steganography vs. Cryptography

Steganography and cryptography are quite close concepts whereas there are slight differences between them. Securing data against malignant actions is the basic aim to be considered. One survey underlines the relation between cryptography and steganography in a well-structured way. It puts all answers between the similarities and differences of cryptography and steganography concepts [10].

Steganography uses an ordinary digital media, as if there is a way to embed a special message into it. The message can be embedded by using special techniques like LSB as in this work. On the other hand, cryptography is the skill of altering the secret message. If the attacker reaches the secret message, then the cryptographical system is broken. For steganography, the attacker first needs to understand whether there is a stegoimage or not which can be visually hard.

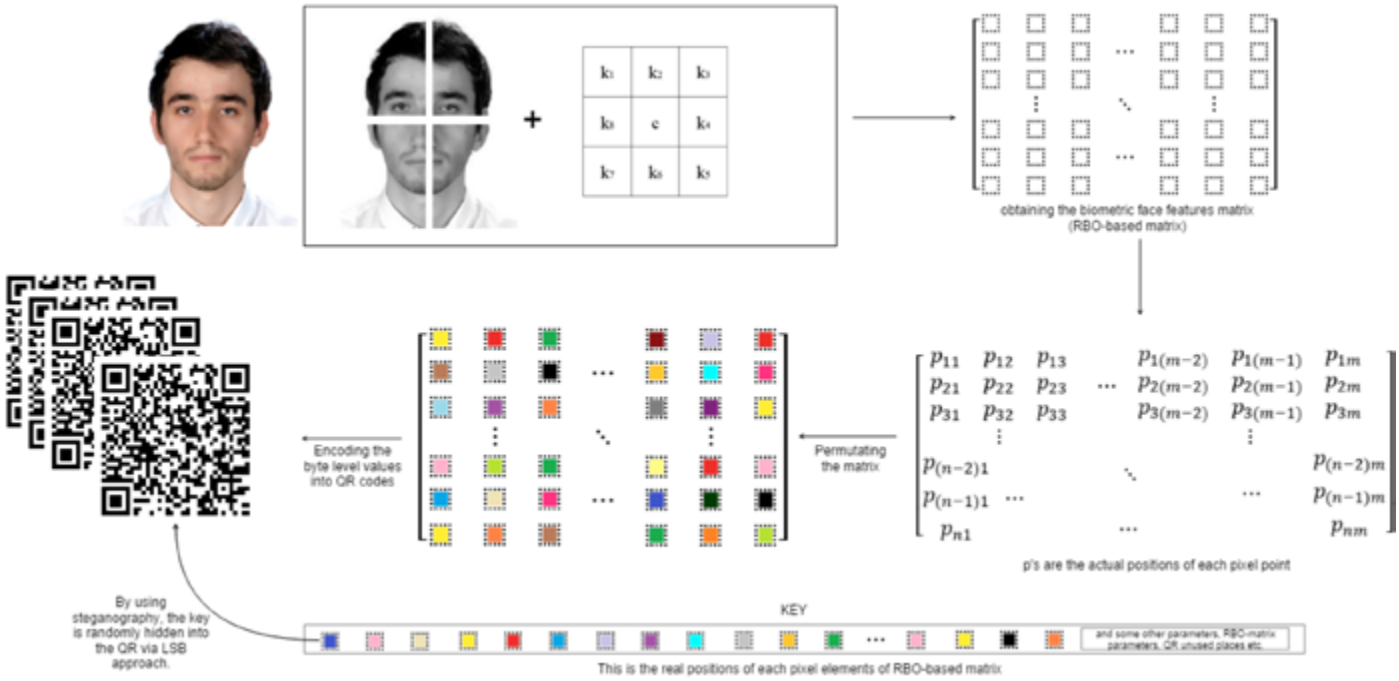


Fig. 5. The whole proposed system that secures biometric face image.

V. PROPOSED METHOD

Using QR code in the biometric data transmission is a bright idea to confuse intruders. The raw data of biometric face image is not shared openly, but its extracted data is sent. Also, it is mixed up randomly by using permutation cipher that the actual positions are stored as key for deciphering process on the destination side. After, this complex data is encoded into QR code where the key of one other image tile is randomly put into QR, too. The key is hidden via steganography. Therefore, the actual positions of featured biometric data are hidden.

As in Fig. 5, first the face image is used for pre-processing like gray level conversion and tiling. Then, the feature extraction by using RBO is handled. The matrix format like an image is constructed and each p is the actual position of the related pixel value. The extracted features are then mixed up randomly by saving their actual positions for the decoding process. The colorful matrix is the new complex data to be encoded in the QR image. The important point in the proposed system is getting the features. The method as Relational Bit Operator is also proposed both by considering the image processing and the cryptography together. During mixing up extracted data, permutation is employed. By keying the actual positions of first nxm sized matrix illustrated with p's is embedded into QR code regular area by LSB while rest of the QR holds the permuted feature data. The face image is actually tiled and each part is embedded into different QRs because of quick response code memory.

A. Vulnerability and Limitation Analysis

The proposed method can be analyzed in the sense of attackers, thus the complexity of the proposed method can be thought in terms of the several dependents:

i. Noticeability of the steganography by an attacker
ii. RBO method
iii. Permutating the matrix
iv. QR code unused areas (key placement)
v. Key management of different image parts

All these concepts have their own difficulty to be issued by an attacker. Even, whomever obtains the QR code of this biometric data should probably first concentrate on the data encoded itself, however it is not the only valid data for authentication.

In practice, from different input images getting the same RBO matrix is not quite possible where it relates to the biometric features. Somehow, it is assumed to have identical outputs because of the physical effects on the images like illumination, noise, acquisition issues etc. If two matrix have the same values, then the permutation acts. Fig. 6 shows that the case when even two different images get the same RBO output. Each input image is also a parameter for the permutation. Plus, the Gray Level Co-occurrence Matrix - GLCM is other input. Each image input has its gray level value occurrences and that differ in different images leading to obtain GLCM, then even though the results are the same, because input images and GLCM differ, the output of permutation changes, too. This preserves to have non-identical results for biometric issues, also giving an advantage to use GLCM for authentication.

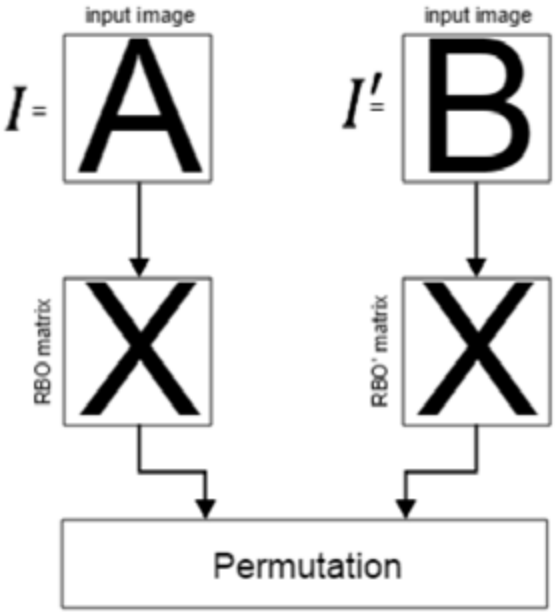


Fig. 6. The worst case scenario that the different images have the same RBO matrix output.

If the overall complexity for the attacks is measured, the exhaustive trials through the database can be checked. For 200x240 pixels and 8-bit gray level face image given in Fig. 5, there are (2gray-level)row*column = (28)48000 brute force trials. For the proposed method, each pixel is revalued by considering the neighbors. This is then converted a decimal value such as gray level value. The occurrences between neighbors cannot be easily inferred from just a decimal number. Even more, the starting pixel and rotation of RBO are other secrets for the attacker. Moreover, the image could be in tiles because of the QR memory. QR code size has a remarkable concern. Version 40 QR code has maximum capacity of 2953 (~3000 will be assumed) bytes which has 177x177 modules. This brings a limitation and the whole data cannot be embedded into the just one QR directly, but the tiles of the image can be generated and they can be used for different QR images. The tiles can be sent in a random order which must be figured out by only the receiver. This puts one other complication for the intruders. Limitation itself brings an advantage, too. For parallel processing issues, tiles can be computed in parallel. Fig. 7 shows the overall reverse operation of reaching back to keys. In this figure, first the image parts are obtained in random order. For instance, 200x240=48000 pixels can be put into 48000/(~3000)=16 QR items approximately. All possible image parts (n=16) can be reordered accurately by n! amount of brute force trials. Then, the attacker needs to understand the steganography. The risk of steganography is the same as all other scenarios: the embedded data should not be captured from its place. If the attacker can understand the steganography, then it can pose a threat but the embedded key itself is like a data package, there are several sub-blocks inside that cannot be easily understood. Exhaustive key search then begins for the malicious attack. The problem can be more complex if the keys are not sent via the QR to which belongs. Keys can be bound to different pairs to make it complex as n! again for each key. The n parts of the image is taken into account and n*n! is reached. Finally, expected key is tried to be extracted from the ~1800x1800 pixels of QR image (version 40). Order of these pixels can give the key as (QR_Row*QR_Column)! times attempts. All in all, the two scenario is compared: obtaining the biometric data by a brute force attack and reversing the proposed method to get the keys to make the system broken. On the condition that the number of image tiles n is large enough, then the proposed scheme is stronger than a brute force attack.

$n*[n!*(QR\_Row*QR\_Column)!*n] >> (2^{gray\text{-}level})^{row*column}$

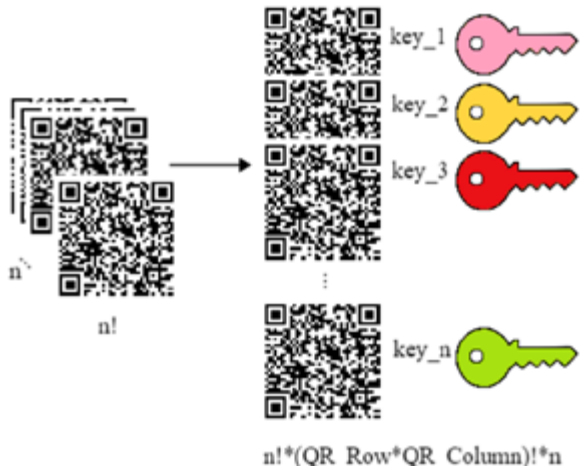


Fig. 7. Hardness of the recovering information in brute force sense.

VI. CONCLUSION

This study proposes a secure, biometric based data hiding technique by using two interesting work space: QR coding and steganography by considering the cryptography. Therefore, this paper aims at combining cryptography which can be useful for image processing techniques. The person who gives biometric data may not want to share his or her data in public openly. Therefore, just the features are sent in a secure way.

There may arise a question that why to choose QR codes for steganography. First, document authentication systems have been becoming quick response code related and QR codes can be used for data compression and web search issues related to image based queries. Besides, QR code has its own correction algorithm and even there can be some losses because of embedding the secret key, but the QR should be recovered successfully.

Security level of this study as discussed in Section V concludes a trade-off where the size of original biometric face image can be in some interval. The tests have shown that as the image size increases, computation load of each image tile – or QR code numbers – increases, too. Whereas, increased n makes the system more resistant to attacks. Finally, it is arrived that the biometric image standards by some institutions like International Civil Aviation Organization-ICAO are successfully met in terms of sufficient image size requirement of this study.

This work is the part of a master thesis, therefore there are some other approaches as future work. For instance, image feature extraction process can be in parallel during the slicing the image into parts. Moreover, for authentication purposes there can be a multi-modal approach like adding fingerprint sensor which was previously realized in [11].

All in all, this paper achieves to use QR codes in a different manner by embedding biometric data features as byte level values into them. In the literature, general trend is just placing text data and after using QR to be hidden into one other image. Therefore, this paper differently uses steganography to hide a key into QR itself via error correction flexibility.

REFERENCES

[1] W. Y. Chen, J. W. Wang, "Nested image steganography scheme using QR-barcode technique," Optical Engineering vol. 48(5), May 2009.

[2] A. Zigomitros, C. Patsakis, "Cross format embedding of metadata in images using QR codes," Intelligent Interactive Multimedia Systems and Services, vol. 11 of the series Smart Innovation, Systems and Technologies, pp 113-121.

[3] C. H. Chung, W. Y. Chen, C. M. Tu, "Image hidden technique using QR-barcode," 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 522 – 525, Kyoto, Sept. 2009.

[4] Y. Y. Chang, S. L. Yan, P. Z. Lin, H.B. Zhong, J. Marescaux, J. L. Su, M L. Wang, Pei-Yuan Lee, "A mobile

medical QR-code authentication system and its automatic FICE image evaluation application," Journal of Applied Research and Technology, vol. 13, pp. 220–229, 2015.

[5] S. U. Maheswari, D. J. Hemanth, "Frequency domain QR code based image steganography using Fresnelet transform," AEU - International Journal of Electronics and Communications, vol. 69, pp. 539–544, Feb. 2015.

[6] M. Ramesh, G. Prabakaran, R. Bhavani, "QR- DWT code image steganography," International Journal of Computational Intelligence and Informatics, vol. 3, pp. 9–13, April 2013.

[7] A. Ross and A. Othman, "Visual cryptography for biometric privacy," IEEE Transactions on Information Forensics and Security, vol. 6, no. 1, pp. 70–81, March 2011.

[8] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognition, vol. 29, no. 1, pp. 51–59, 1996.

[9] W. Islam and S. alZahir, "A novel QR code guided image stenographic technique," 2013 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, Jan. 2013.

[10] A. J. Raphael, V. Sundaram, "Cryptography and Steganography – A Survey," Int. J. Comp. Tech. Appl., vol. 2 (3), pp. 626–630.

[11] S. Aygün, M. Akçay, and E. O. Güneş, "Bulut sistemler için önerilen biyometri tabanlı güvenlik sistemine genel bakış," The Third International Symposium on Digital Forensics and Security (ISDFS 2015), May 2015.

BULUT BİLİŞİMİN KURUMSAL ZORLUKLARI VE ÇÖZÜM ÖNERİLERİ

Yasin İNAĞ, Eyüp Burak CEYHAN, Şeref SAĞIROĞLU

Gazi Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü, ANKARA
yasininag@gmail.com, eyupburak@gmail.com, ss@gazi.edu.tr

Özet — Son yıllarda veri saklama, sistem yönetme ve veriye ulaşma bulut bilişim teknolojisi ile sağlanmaya başlanmıştır. Yeni nesil cihazlarda donanımsal maliyet ve gereksinimleri azaltmak için sanal sunucular kullanılmaktadır. Kurumsal şirketler maliyet, yönetim ve bakım kolaylığı gibi sebeplerden dolayı bulut bilişimi tercih etmektedir. Bulut bilişim kullanıcı odaklı güvenlik açıklarını ortadan kaldırmaktadır fakat farklı güvenlik sorunları da mevcuttur. Fiziksel sunucular üzerinde kurulan sanal sunucuların yönetimi ve güvenliği önemlidir. Çoklu erişim ile birden fazla kullanıcının aynı anda farklı yerlerden erişim sağladığı sunucuların güvenliği ve kullanıcılara sağlanmış olan sanal sunucuların yönetimi farklı sorunlar meydana gelmesine sebep olmaktadır. Bellek yönetimi, sanallaştırma, kullanıcı veri güvenliği, veri gizliliği ihlalleri gibi yönetimsel sorunlar farklı yöntemler ile çözümlenmektedir. Bulut bilişimde, sanallaştırma, ortak altyapı kullanma, bellek yönetimi ve farklı ülkelerin farklı yasalarından kaynaklı zorluklar bulunmaktadır. Karşılaşılan sorunlar, sebepleri ve çözüm önerileri çalışmamızda sunulmuştur.

Anahtar Kelimeler — Bulut bilişim, Bellek yönetim, Çözüm önerileri, Karşılaşılan zorluklar, veri güvenliği

Abstract — In recent years, cloud computing is used for data storage, system management and accessing data. Virtual servers are being used on new generation devices to reduce hardware cost and requirements. Corporate companies prefer cloud computing for its cost, management and easy maintenance reasons. Cloud computing eliminates user-oriented sense of vulnerability. However, it has different security issues. Managing and securing virtual servers on physical servers are crucial. Accessing to these servers at the same time by multiple users from different locations might cause some security issues and management problems. Administrative issues such as memory management, virtualization, user data security, data privacy violations are resolved by different methods. In cloud computing, there are some difficulties in virtualization, using shared infrastructure, data management and different laws in different countries. This study presents current problems, its reasons and possible solutions.

Index Terms — Cloud computing, Storage management, solution recommendations, the difficulties encountered, data security.

I. GİRİŞ

Veriye ulaşmanın ve iletişimin hızla arttığı günümüzde erişimin mekandan bağımsızlaşması için farklı teknolojiler geliştirilmiştir. Bulut bilişim bu teknolojilerden biridir. İnternet alt yapısının gelişmesi ve geniş bant iletişimin sağlanması, yaygınlaşması ve ucuzlaması ile kullanımı yaygınlaşmıştır. Yapısal olarak, sabit fiziksel sunuculara internet yardımıyla erişim sağlanmaktadır. Fiziksel sunucularda kurulan ve yönetilen sanal sunucular ile kullanılır. Çoklu erişim yöntemi ile aynı anda farklı yerlerden farklı kullanıcıların erişimi sağlanmaktadır. İnternet erişiminin olduğu her yerden erişim mümkündür.

Bulut bilişimin daha ucuz ve güvenilir bir sistem olması, bakım onarımının ve geliştirilmesinin kolay olması kurum ve kuruluşlar tarafından tercih edilmesinin sebeplerindendir. Yönetimin kolay, maliyetin ucuz olduğu bulut bilişimde kullanıcı taraflı güvenlik açıkları azaltılmaktadır. Fiziksel sunucular üzerinden kurulan sanal sunucular ile teknoloji bağlantısı 7/24 sağlanmaktadır.

Farklı kullanıcı özellikleri ve aynı güçte tek sunucu üzerinden sanallaştırma kullanılarak hizmet verilmesi, kullanılmayan donanımsal ve yazılımsal maliyetleri engeller.

Bulut bilişimin yönetimi kolay olmasına karşın, güvenlik yönetimi tecrübe ve dikkat gerektirmektedir. Yetki, sanallaştırma, bellek üniteleri yönetimi, saldırı tespit ve sistem geri yükleme gibi durumlarda farklı sorunlar ortaya çıkmaktadır. Mimari yapı ve yetkilendirme, yazılımsal ve süreklilik, dokümantasyon ve yasal eksiklikler karşılaşılan başlıca sorunlardır.

Çalışmada karşılaşılabilecek sorunlar ve çözüm önerileri açıklanmıştır. Çalışmanın ikinci bölümünde bulut bilişimin teknik ve yazılımsal alt yapısı açıklanmıştır. Üçüncü bölümde literatürde karşılaşılan sorunlar üzerinde durulmuştur. Dördüncü bölümde karşılaşılan sorunlara karşı alınacak önlemler belirtilmiştir. Sonuç bölümünde ise kurumların bulut bilişimden yararlanırken dikkat etmesi gereken hususlar sunulmuştur.

II. BULUT BİLİŞİM

Teknolojinin hızla gelişmesi ile birlikte iletişim ve veriye ulaşmak da kolaylaşmaktadır. Veri, program ve yönetimsel araçlar artık sabit sunucularda saklanmamaktadır. Bulut bilişim olarak adlandırılan yapı ile kurumsal ve kişisel veriler sanal sunucularda tutularak zaman ve mekân kısıtlaması sorun olmaktan çıkmıştır. Bulut bilişim için farklı tanımlamalar yapılmıştır. Bulut bilişim, hesaplama ve bilgi hizmetleri iş modelidir. Birçok farklı yerde ve sayıda fiziksel veya sanal sunucular ile bilgiye ulaşmayı sağlayan bir disiplindir. Bulut bilişim daha az enerji harcanarak, daha kolay yönetim ve servis destekleyen bir teknolojidir [1]. Berkeley'e göre bulut bilişim, veri merkezleri aracılığı ile birçok farklı modülde internet uygulamaları, servisler, donanım ve yazılımsal destek sağlayan teknolojidir [2]. Foster ve arkadaşlarına göre bulut bilişim, internet alt yapısı ile ulaşılan geniş ölçekli, dinamik, ölçeklenebilen, büyük veri depolama alanına sahip, sanallaştırılabilen ve daha az maliyetli bir iletişim ve ortak alan teknolojisidir [3].

Bulut bilişim, sanal ortamda ulaşım sağlanan fiziksel sunucularda tutulan verilere ve uygulamalara bilgisayar, mobil cihazlar veya diğer teknolojiler ile ulaşılmasıdır. Geniş bant ağ teknolojisinin ve ülke genelinde internet alt yapısının gelişmesiyle birlikte kullanımı artmakta ve kolaylaşmaktadır. Birçok kurum ve kuruluş bulut teknolojisinden

faydalanmaktadır. Lisans ve bakım işlemlerinde kolaylık sağladığı gibi maliyeti de azaltmaktadır [3].

Sanallaştırma makineleri ile bir sunucu üzerinden yayın yapılarak tek bir lisansla kullanıcıya hizmet sağlanabilmektedir. Eğitim kurumlarındaki bilgisayar laboratuvarlarında, sanal makine ile sunuculara bağlanılarak kasasız, daha geniş depolama alanına sahip, daha güçlü bilgisayarlar olarak kullanılmakta, lisans ve bakım maliyetlerinden de önemli ölçüde kazanç sağlanmaktadır.

İnternetin ve bilgi toplumunun gelişmesi ile bulut bilişim kaçınılmaz bir teknoloji olmuştur. Büyük şirketler, kuruluşlar ve kişiler tarafından kabul edilen ve dağıtık bilgi işleme teknolojisinin yerine kullanılan yeni teknolojidir [4].

Bulut bilişim uygulamaları ve alt yapısı sanallaştırma teknolojisi üzerine inşa edilmiş bir yapıdır. Sanallaştırma, donanımların daha etkin kullanılmasını, aynı anda birden fazla uygulamanın birden fazla kullanıcı tarafından kullanılmasına olanak tanır [5].

Bulut bilişim dinamik yapısı ile kullanıcı ihtiyaçlarına göre etkin kullanım sağlar. Kullanıcının ihtiyacına uygun verileri kullanmasını sağlayarak gereksiz işlemci yükü ve enerji kaybını önler [5].

Büyük veri setlerini ve güçlü sistemleri tek elden kontrol ederek ve erişim sağlayarak büyük bir ekonomik tasarruf sağlar. Her kullanıcı için ayrı depolama alanı, işletim sistemi, uygulamalar, lisanslamalar ve donanımların yerine, doğrudan yönetim sağlayan ve ihtiyaca göre dinamik alt yapı hizmeti sunan ve zamandan ve maliyetten tasarruf sağlayan sistemlerdir.

A. Bulut Bilişim ve Diğer Gelişmiş Teknolojiler

Bulut bilişim, bilim ve teknoloji ışığında bilişim alanlarına uygulanmaktadır. Paralel hesaplama, dağıtık bilgi işleme, kiralık sunucu hizmeti, yaygın hesaplama, yazılım servisleri ve sanallaştırma teknolojileri gibi gelişmiş teknolojilerin bulut bilişim ile ortak çalışma alanları vardır. Çalışmanın bu bölümünde gelişmiş teknolojiler ile bulut bilişim arasındaki benzerlikler ve farklılıklar açıklanmıştır.

Paralel Hesaplama, karmaşık ve çözümü zor olan bazı problemlerin, birçok küçük bilgisayarın bir araya gelerek eş zamanlı çalışması sonucu kısa zamanda çözülmesini sağlayan teknolojidir. Genellikle çözümü için yüksek performans gereken problemlerin çözümünde kullanılır.

Dağıtık bilgi işleme, sanallaştırmayı ve birden çok kullanıcının ortak bir yapıda çalışmasını sağlayan sistemdir. Dağıtık bilgi işleme, problemin çözümü için kullanılan veri setini farklı işletim sistemlerinin veya cihazların ortak kullanımını sağlayan sanal servis sağlayıcısıdır. Sanallaştırma daha güçlü ortak bir işlemci sistemi oluşturur. Dağıtık bilgi işleme teknolojisinde tek bir veri seti etrafında işlem yapan farklı kullanıcılar varken, bulut bilişimde büyük veri setinden faydalanan bağımsız kullanıcılar bulunmaktadır [4].

Kiralık sunucu hizmeti, kullanıcılara ait oluşmuş verileri incelemek için kurulmuş bir alt yapıdır. Kiralık sunucu hizmeti genellikle su, elektrik veya doğalgaz kullanımına göre oluşan faturaları oluşturmakta faydalanılan ortak kullanımlı kaynak sağlayıcısıdır. Kurumların hesap işlemleri, ortak sunuculara bağlanan kullanıcıların tüketim miktarı ve ödemesi gereken tutar gibi verileri sağlayan sanallaştırma teknolojisinin de kullanıldığı yapıdır. Bulut bilişim sadece kaynak sunmaktan ziyade veri iletimi, uygulama geliştirme ve yönetme gibi servisleri de içermektedir [5].

Dağıtık sistem birden fazla bilgisayarın aynı ağ üzerinde birbirleri ile etkileşimidir. Belirlenen amaç doğrultusunda ortak çalışan sistemlerdir [4].

Yazılım destekli servisler, paket yazılımların ortak kullanımına imkân sağlayan altyapı hizmetleridir. Sunucularda tutulan yazılımların kullanım ihtiyacına göre ortak kullanılmasına olanak tanıyan sistemdir. Uygulamalara erişimde sunuculara erişim ile olanak sağlandığından internet altyapısı olan her yerden ve her zaman ulaşılabilir [4].

Sanallaştırma, bilgisayar programlarının veya sistemsel yazılımların aynı donanım üzerinde çalıştırılabilmesidir. İşlemci sanallaştırma, tek işlemciyi çoklu işlemci gibi kullanmaya yarar ve aynı anda birden fazla uygulama veya işletim sisteminin kullanılmasını sağlar. Bu uygulamalar bellek ünitelerini birbirlerinden bağımsız kullanmaktadır. Bu yapılar bilgisayarın daha verimli kullanılmasını sağlar [5].

B. Bulut Bilişimin Sınıflandırılması

Şekil 1'de bulut bilişimin kullanım şekline ve verdiği hizmetlere göre sınıflandırılmıştır. Bu sınıflar alt başlıklarda detaylı açıklanmıştır.

Şekil 1. Bulut bilişimin sınıflandırılması [6].

Sağladığı hizmetlere göre sınıflandırma

Servis hizmetleri (Infracture as a service, IaaS), kullanıcıya sunulan donanımlardır. Fiziksel ve sanal sunucular, ağ, bant genişliği gibi servislerin son kullanıcıya ulaştırılmasını sağlayan yapıya IaaS denir. Kullanıcıların uygulamalarını ve

düzenlerini sağlayabildiği yapıdır [6].

Platform hizmetleri (Platform as a service, PaaS), kullanıcıların uygulamalarını ve verilerini sakladığı alandır. Kişisel kullanım ayarlarının yapıldığı alandır [6].

Yazılım Hizmetleri (Software as a service, SaaS), kullanıcıların direk kullanabildiği yazılım ve program hizmet sağlayıcı birimidir. Veri saklamak için gerekli olan bellek ve yönetimi bu hizmet katmanı altında yer alır. Güvenlik ve sistem yönetimi gibi teknik alt yapılara destek sağlayan hizmet birimidir [6].

Kullanım şekline göre bulut bilişim

Bulut bilişim, güvenilirliği, kullanılabilirliği ve doğruluğu sağlamaktadır. Bunu her bir kullanıcı için ayrı ayrı yapılmasına gerek kalmaksızın gerçekleştirebilir [7]. Bu problemin çözümü için bulut bilişim genel (public), özel (private) ve hibrit(hybrit) olmak üzere üç ayrı kategoriye ayrılmıştır [8]. Genel (Public) bulut, kullanıcılar bulut sağlayıcıları tarafından her türlü bilgi ve belgeyi genelin kullanımına açarlar. Halka açık bir ağ topluluğu olarak da tanımlanabilir. Özel (Private) bulut, kurum veya kuruluşlara özel yapılandırılmış sistemlerdir. Üçüncü şahısların kullanımına kapalıdır. Hibrit (Hybrit) bulut, özel ve genel bulut mimarilerinin birlikte kullanıldığı bir yapıdır. Sistemin belirli kısımlarına sadece belirli kullanıcıların erişimine izin verildiği yapıdır.

C. Bulut Bilişimin Avantajları

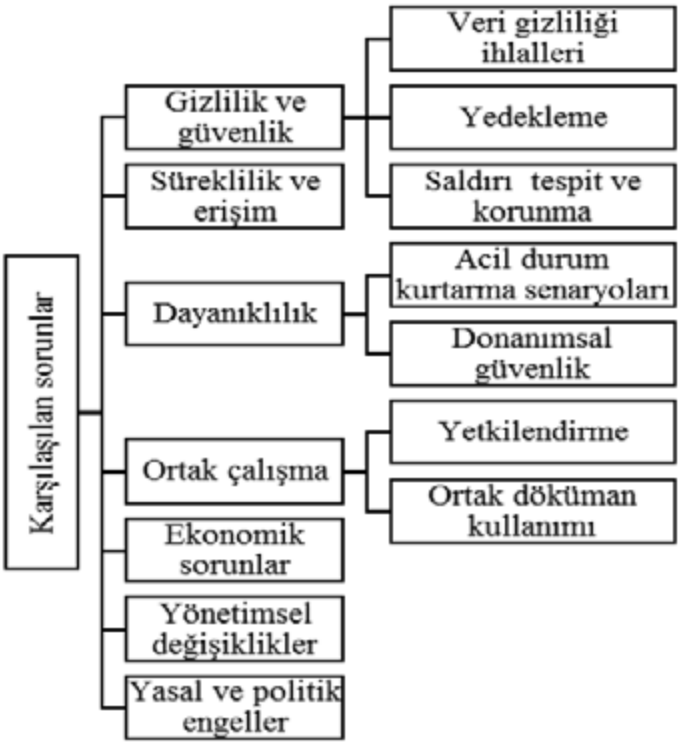
Kullanım alanına ve ihtiyaca göre dinamik yapı sergileyen bulut bilişim birçok kurum ve kuruluş için ekonomik tasarruf sağlamaktadır. Kişisel bilgisayarlar ile yapılan tüm işlemler bulutta da yapılabilmektedir. Bulut yapısı doğru yönetildiği sürece daha iyi performans elde edilebilir [9]. Her kullanıcı için ayrı ayrı alınması gereken programlar ve lisanslamalar tek bir yapı üzerinde oluşturularak ortak kullanıma açılır. Performans kaybı yaşamadan sanallaştırma yardımıyla aynı programdan birden fazla kullanıcı aynı anda yararlanabilir [8].

Yönetimde, sürdürülebilirlikte ve bakım çalışmalarında da büyük kolaylık sağlamaktadır. Herhangi bir yazılımın güncellenmesi gerektiğinde tek elden buluttan yapılan güncelleme ile kısa sürede ve kolayca yeni sürüm kullanıma sunulmuş olur. Yazılımsal ve donanımsal bakımın tek sunucu üzerinden yapılması büyük avantaj sağlamaktadır. Özellikle bilgisayar teknolojisinden anlamayan kullanıcılara sahip şirketlerde kullanılan uygulamalara gelen yenilikleri eklemede kullanıcı bazlı sorunlar yaşanmaktadır [10]. Bilgisayarların tek tek düzeltilmesi veya eğitimler ile yapılması gereken basit değişikliklerin bile anlatılması gerekmektedir. Bulut bilişim bu sorunu ortadan kaldırmaktadır Kurum veya kuruluşların farklı bölümleri için gerekli depolama alanı ihtiyaca göre farklılık göstermektedir. Bulut bilişimde dinamik bellek yapısından kaynaklı sorun yaşanmamaktadır [11].

Bir sunucuda bulunan verilerin farklı bir sunucuda da kopyası tutulmaktadır. Bu sebeple veri güvenliği sağlanmaktadır. Kullanıcı farkındalığı eksikliğinden kaynaklı oluşan hatalar ve sistem açıkları da bulut bilişimde engellenmektedir. Antivirüs ve benzeri uygulamalar etkin bir şekilde kullanılarak sistem ön güvenlik işlemleri gerçekleştirilir [12]. Bulut bilişimde kullanıcılar farklı işletim sistemleri kullanabilmektedir [7].

D. Bulut Bilişimdeki Zorluklar

Şekil 2'de bulut bilişimdeki karşılaşılan sorunlar sınıflandırılmıştır.



Şekil 2. Karşılaşılan sorunlar [10, 13-17]

Süreklilik ve erişim kullanıcılara sunulması gereken öncelikli hizmetlerdendir. Bulut mimarisinin 7/24 erişilebilir olması, sistemin süreklilik arz etmesi ve bu sürekliliğin sadece erişimde değil etkin performansta da olması gerekmektedir [14].

Sistemin olası bir saldırı veya afet durumunda dayanıklılığı, kurtarma senaryoları ve acil durum senaryolarının geçerliliği karşılaşılan zorluklardandır. Sel, deprem gibi doğal affet durumlarında alınmış önlemlerin güvenilirliği ve kurtarma stratejileri önem arz etmektedir [15].

Yetkilendirme, sisteme erişim hakkı olan kullanıcı veya profillerin sistemde müdahale sınırlarının ve önceliğinin belirlenmesidir. Ortak çalışmada veri üzerinde değişiklik hakkına sahip farklı kademede birden fazla kişinin yetkisi bulunabilir [16]. Ortak doküman kullanımında aynı anda yapılan değişikliklerin gerçekleştirilmesi bulut bilişimin zorluklarındandır.

Ekonomik sorunlar kurumların bulut bilişimi tercih etmesinde önemli etmenlerden biridir. Bulut bilişim, kullanıcı odaklı kullanımdan dolayı daha düşük maliyetlidir. Ancak kendi yerel sistemini kullanan bir kurumun bulut bilişim mimarisine geçişi için gerekli maliyet gereksiz gelebilmektedir. Bu sebeple yapılacak olan analizlerde uzun vadeli planlar oluşturulmalıdır. Kısa vadede bir yük gibi görünse de uzun vadede kendini amorti edecek bir sistem olduğu gözlemlenmektedir [15].

Yönetimsel değişiklikler, kurumlarda değişen yöneticilerin farklı kararlar almasından veya gerekli dokümantasyon hizmeti sağlanmamasından projelerin değiştirilmesine sebep olmaktadır. Yönetimin kapsama alanına giren her

türlü yaklaşımda dokümantasyon işleyişin devamlılığı için önemlidir. Yönetimde bireylerin değişmesi farklı stratejileri beraberinde getirebilir. Önceki yapı bilinmeden yapılan değişiklikler veya ne yapıldığı bilinmeden uygulanan farklılıklar sorunlara sebep olabilir [17].

Veri gizliliği, kullanıcıların bulut ortamında sakladığı verileri üçüncü şahıslardan korumaktır. Kullanıcılara ait özel verilerin kötü niyetli kişilerden korunmasıdır. Sistemin geri bildirimini güçlendirmek veya sistemin reklam amaçlı kullanımı bile bazen kullanıcıların verilerinin gizliliği ihlal edilerek yapılmaktadır. Ülkelerin veri gizliliği ve korunması doğrultusunda aldığı kararlar farklılık gösterebilmektedir. Kullanılacak sunucuların bulunduğu ülke yasaları ile kullanıma açılan ülkelerin yasalarındaki farklılıklar sorun olarak gözlemlenmektedir [17].

III. BULUT BİLİŞİMDE GÜVENLİK SORUNLARI

Bulut bilişim, alt yapısı ve sağladığı olanaklar ile kullanıcılara daha ekonomik, kullanılabilir ve güvenilir sistemler sağlamaktadır. Bu yeni teknoloji birçok farklı mimariyi kullanıcı veya IT uzmanlarına sunmaktadır. Kullanılan sistem alt yapısı ve bulut teknolojisine göre farklı güvenlik riskleri oluşmaktadır. Hizmetin yanıt vermemesi, IP (Internet Protocol) çakışması, ortak donanım kullanılmasında yetki ve kullanım alanının sınırlandırılmaması, sanal ağlarda uygulamalara erişim, bellek ünitelerinin donanımsal bağımlılığının sanal ortamda ayrılmaması, saldırıların sunucuya bağlanan kullanıcılar tarafından gerçekleştirilebilmesi, web uygulamalarının ve web servislerinin yazılımsal açıklıkları karşılaşılan sorunlardır. Sanallaştırmada çoklu kullanıcı yapısı kullanıldığında farklı riskler ortaya çıkmaktadır. Birden fazla kullanıcının aynı veri havuzunu kullanıyor olması güvenlik risklerini artırırken sistemin kullanılabilirliği ve dinamik yapısını olumsuz etkileyebilmektedir. Web tabanlı kullanım sağlayan karakteristik yapıdan dolayı izinsiz erişim olasılığı normal web teknolojilerinden daha fazladır. Sanal sunucular arası geçişler ve paylaşılan platformun farklı yetkilere sahip kullanımlardan kaynaklı yönetimsel güvenlik açıkları da bulunmaktadır. Servis sağlayıcı hizmetleri platform tabanlı hizmetler ile bağımlı bir şekilde yazılım hizmetleri sunmaktadır [20, 21].

A. İletişimde Güvenlik Problemleri

Bulut servislerine internet altyapısı kullanılarak erişim sağlanmaktadır. Standart internet protokolleri kullanılarak iletişim gerçekleştirilir. İletişim, kullanıcılar ile bulut ve bulut bilişim sanal sunucuları arasında gerçekleşmektedir [22].

Kullanıcı ile bulut bilişim arasındaki iletişimden kaynaklı sorunlar; hizmetin yanıt vermemesi, ağ yönlendirme sorunları, ağın dinlenmesi, IP çakışması ve maskeleme gibi normal bir bağlantıda karşılaşılan sorunlardır. Ağ yönlendirme, kurulan altyapının yanlış konfigürasyonlar sonucu kısır döngüye girmesi veya yanlış sunuculara yönlendirilmesidir. Ağ dinlemesi, kullanılan alt yapıyı üçüncü şahısların gizlice izlemesi kişisel bilgi, görüşme ve ağ üzerinde gerçekleşen işlemleri elde etme çabasıdır. IP çakışması, kullanıcıların ağ üzerinde tanımlandırılan isimlerin aynı olması ve bu yapılandırmanın yanlış şekillendirmesinden kaynaklı oluşan sorunlardır [23-26, 28-29]. Bu problemlere çözüm önerileri; güvenlik soket katmanı, internet güvenlik protokolleri, şifreleme algoritmaları, denetleme, dijital sertifikalar ve trafik kontrolü gibi çözümlerdir.

Paylaşılan iletişim altyapısı

Ortak kaynak havuzu sadece veri ve uygulamaların kullanımı değil aynı zamanda ağ alt yapısının da ortak kullanımı demektir. Ağ bileşenlerinin ortak kullanılması çapraz kiracı (cross tenant) saldırılarına olanak sağlar. Çapraz kiracı saldırıları, ortak havuzda platform tabanlı servis sağlayıcıların kullanıcılara sağladığı, veri iletişim kanalları üzerinde gerçekleştirilen saldırı türüdür. Bulut bilişim kullanıcıları genellikle sanal sunucularına bağlanırken belirli bir kapasite ile sınırlıdır. Aşırı kullanım veya kötü niyetli kullanıcılar bu şekilde tespit edilmektedir. Sistem MAC ve IP adreslerini kayıt altına alarak sisteme sızmaya çalışan kötü niyetli ataklar engellenmektedir [26].

Sanal ağlar

Bulut bilişim sistemlerinde bağlantı sadece fiziksel ağ yapıları ile değil aynı zamanda sanal ağlar ile de gerçekleştirilmektedir. Sanal ağlar, sanal sunucular arasında iletişim gerçekleştirmek için tasarlanmıştır. Aynı hizmet sağlayıcısı üzerinde kurulu sanal sunucular arasındaki konfigürasyon, köprü kurma ve yönlendirme işlemlerini gerçekleştirmek için kullanılır. Sanal ağlar üzerindeki trafik fiziksel ağ üzerinden izlenememektedir. Bu sebeple sanal ağların güvenliği de fiziksel altyapı üzerinden gerçekleşmektedir. Saldırı tespit ve önleme mekanizmaları genellikle trafik izleme verisi üzerinden veya sistemsel anormallikler üzerinden anlaşılabilmektedir. Sanal ağlara saldırılar genellikle DoS, sızma veya dinleme ile yapılmaktadır. Bu saldırılarda sanal ağ kullanımı izlenerek tespit ve önlem gerçekleştirilir [23].

Hatalı güvenlik yapılandırılması

Bulut bilişimde kullanıcıya güvenli bir altyapı sağlamak oldukça önemlidir. Hatalı yapılandırma güvenlik ihlallerine sebep olabilir. Yaygın olarak yapılan hatalı yapılandırmalardan biri, yöneticinin almış olduğu güvenlik tedbirinin bütün sistemde geçerli veya bütün sistemle uyumlu olmamasıdır. Bulut sistemi güvenlik standartlarına uygun yapılandırılmalı ve denetlenmelidir. Dinamik yapıda yönetilmesi gereken, sistem yapılandırmadan kaynaklı dar boğaza girilmesi veri kaybına veya verinin ihlaline sebep olabilmektedir [28].

B. Mimari Güvenlik Problemleri

Bulut bilişimin donanımsal yapısından kaynaklı farklı sorunlar ortaya çıkmaktadır. Temelinde sanallaştırma teknolojisi kullanılan teknolide sanallaştırmanın yapılandırılması ve sanal sunucuların kullanıcı yetki sınırında kullanılması gerekmektedir. Bellek ünitelerinin güvenliği için, herhangi bir saldırı veya kayıp durumuna karşı geliştirilmiş kurtarma senaryoları geliştirilmelidir. Mimari yapının hataları veya eksikliklerinden oluşan sorunlar çalışmanın devamında sanallaştırma, bellek üniteleri, kimlik yönetimi ve denetimi ve yasal sorunlar olmak üzere dört alt başlıkta açıklanmıştır.

Sanallaştırma

Sanallaştırma bulut bilişimin temel yapılarından biridir. Kullanıcılar kendi sanal makinelerini oluşturabilirler.

Oluşturulan bu sanal makinede örneğin resim paylaşımı gerçekleştirilebilir. Ancak sisteme yüklenen resimlerin içerisinde kötücül yazılımlar yerleştirilmiş olabilir. Sanal sunucuda denetlenmesi zor olan bu yapı sebebiyle sanal makine zarar görebilir. Sanal makinenin gördüğü zarar, fiziksel donanım ve sanal ağlar ile tüm bulut sistemini de etkileyebilir.

Sanal makineler birbirinden bağımsızdır ve ayrı çalışabilmesi gerekmektedir. Aynı donanım üzerinde kurulu farklı sanal makinelerin bellek, işlemci ve diğer donanımların kullanımını ayırmaması durumunda sistem darboğaza girebilir. Bu da daha önce anlatıldığı gibi çapraz kiracı saldırısı olarak algılanıp sistemin çalışmasını engelleyebilir [21, 26].

Sanal sunucular sızıntı veya kaçış kötü niyetli kullanıcıların oluşturduğu ve bütün sanal makineleri kontrol eden mekanizmanın kontrolünden kurtulma amaçlı oluşturulan yapılardır. Makineler arası iletişim ve güvenlik protokollerini kontrol altında tutan mekanizma devre dışı bırakılarak veya bir makine gizlenerek sistemde güvenlik zafiyeti oluşturulabilir. Bu sızıntı ile fiziksel bellek ünitelerine erişim sağlanabilir [19].

Sanal makine göç sorunu güvenlik açıklarından biridir. Göç sorunu, sanal makine başka bir fiziksel makineye taşınırken sanal makinenin durdurulmamasından kaynaklı oluşabilecek ihlaldir. Sanal sunucu üzerindeki veriler, kodlar ve protokoller taşınma sırasında saldırıya uğrarsa her iki sanal makine de eksik kalacağından gerekli güvenlik prosedürü işletilemez olacaktır. Sanal makinelerde yapılan işlemleri geri alma daha önceden gerçekleşmiş güvenlik sorunlarını tekrarlayabilmektedir. Protokollerde yapılmış değişiklikler de etkilenebileceği gibi bu yapısal değişiklik bir ihlale sebep olabilir [22-25].

Bellek üniteleri

Bulut bilişim, sistemdeki verilerin erişim ve kontrolünü tam anlamıyla kullanıcılara bırakmaz. Sanal sunucuların etkin kullanımı ve yönetimi sağlanırken aynı durum veriler için geçerli değildir. Bulut içindeki verinin bütünlüğü, erişilebilirliği, gizliliği ve ihlalleri normal sistemlere göre daha az dayanıklıdır. Kullanıcı sayısı ve erişim miktarı arttıkça riskler de artmaktadır. Sadece sistemi bozmaya çalışan veri ihlali sağlayan yapılar değil, normal kullanımdan kaynaklı oluşabilecek darboğazlar da verinin bütünlüğüne zarar verebilir. Normal sistemlerden farklı olarak aynı bellek ünitesine birden fazla yoldan erişimin olması ve aynı anda farklı kişilerin bu belleklere erişebilmesi, normal fiziksel belleklerden daha riskli bir yapı oluşturmaktadır [28].

Veri kurtarma seçeneklerinden kaynaklı oluşan ihlaller bulut bilişimin veri güvenliği noktasında en önemli açıklarından biridir. Ortak bellek yapısını farklı kullanıcılar kullanmaktadır. Sanal sunucu üzerinden silinen verileri kurtarmak için uygulanacak yöntem ile işletim sistemi üzerinden görülmeyen ancak bellek adresleme tablolarında tutulan veriler kurtarılabilir. Ancak geri getirmek için gidilen adres başka şirket veya kişi tarafından kullanılan bellek bölgesi olabilir. Bir kullanıcıya veri kurtarma opsiyonu sağlayan sistem aslında bir başka kullanıcıya ait özel veriye erişim olanağı sağlamış olabilir.

Uzun zaman kullanılmayan veri setleri veya hasar görmüş bellek ünitelerini temizlerken, farklı kullanıcıların uygulamalarının buralarda açık olmasından kaynaklı silinemiyor ve temizlenemiyor olması, bir süre sonra farklı güvenlik ihlallerine yol açabilmektedir. Yedekleme işlemi veri bütünlüğünü korumak için önemlidir. Bulut bilişimde bellek üniteleri çoklu erişime ve kullanıma açık olmasına rağmen yedeklemenin direk erişime kapatılması olası bir saldırı veya zarar durumunda veri kurtarma olasılığını artırmaktadır [29].

Uygulamalar ve arayüz güvenliğinde karşılaşılan zorluklar

Bulut bilişime erişim web arayüzleri ile gerçekleşmektedir. Aynı web arayüzüne birden fazla kullanıcı eş zamanlı erişim sağlayabilmektedir. Hiyerarşik yapı ve yetki yönetiminin doğru yapılması güvenilirlik açısından önemlidir. Farklı kullanıcıların çoklu erişim desteği sayesinde bulut bilişimde kullandığı web uygulamaları sızıntılara ve saldırılara daha dayanıksız olmaktadır. Uygulamaların bulut bilişimde işleyişi ve çalışma prensibi kullanılan sistem hakkında bilgi verebilir [30].

Kimlik yönetimi ve erişim kontrolünde karşılaşılan zorluklar

Bulut bilişimde veri güvenilirliği ve bütünlüğü, kimlik doğrulama ve erişim kontrolü ile sağlanmaktadır. Sistemde yetkisiz erişimi engellemek ve girişimleri kayıt altında tutup analizler gerçekleştirmek oldukça önemlidir. Kullanıcı, yönetici ve sanal sunucularda, yöneticilerin farklı olmasından kaynaklı kontrollerin gerçekleşmesi zorlaşmaktadır. Aynı anda çoklu erişimin sağlanması, tek fiziksel yapı üzerine birden fazla sanal makinenin kurulması, dinamik bir karaktere sahip olması ve IP yapılandırması farklılığından kaynaklı kompleks bir yapı oluşmaktadır. Zayıf kimlik doğrulama ve erişim kontrolünden kaynaklı, hizmetlerin geç cevap vermesi, sanal makinelerin yeniden başlatılması, doğrulama kontrolünün yapılamaması, sistem günlüğü tutma ve izleme işlemlerinin yanlış veya gecikmeli yapılması gibi sorunların oluşmasına sebep olabilmektedir [23].

Sözleşme ve Yasalardan Kaynaklı Karşılaşılan Sorunlar

Bulut bilişim, altyapısı ve kullanım amacı doğrultusunda farklı coğrafyalardan kullanıcıların farklı ülkelerde bulunan sunuculara erişim sağlayarak gerçekleştirilen bir hizmettir. Bulut desteği sağlayan sunucuların, kullanıcılara önermiş olduğu sözleşmelerin kabul edilebilirliği, olası veri kaybı veya zararda nasıl bir yöntem uygulanacağı ve kimin nasıl sorumluluklar üstleneceği sözleşmede açıkça beyan edilmelidir. Sunucuların bulunduğu ülkelerin yasaları ile kullanıcıların kendi ülkelerindeki yasaların farklılığından kaynaklı oluşacak anlaşmazlıklarda hangi ülke yasalarının geçerli olacağı bulut bilişim sorunlarındandır. Bu sebeple uluslararası geçerliliği olan ve bilirkişiler tarafından oluşturulmuş ortak bir yasanın geliştirilmesi ve kullanımı hızla gelişen ve yaygınlaşan hizmetin önündeki engellerden birini daha kaldıracaktır [17].

IV. GÜVENLİK SORUNLARINA ÇÖZÜM ÖNERİLERİ

A. İletişim Sorunlarına Karşı Önlemler

Bulut bilişimde uluslararası oluşturulmuş bulut güvenlik birliği ile sanal ve yerel ağlar, güvenlik duvarı ve IP yapılandırması gibi karakteristik altyapılar standartlaştırılmıştır. Bu standart, kullanıcı verilerini ve kullanılan sistem alt yapısının güvenliğini sağlamak için de kullanılır [31].

İleri bulut koruma sistemleri, bulut verilerini korumak için geliştirilmiş yüksek güvenlikli sistemlerdir. Kullanıcı odaklı ağ ataklarını engellemek için kullanılır. Servis sağlayıcıları tarafından sanal sunuculara yapılan saldırılar tespit edilerek önlenir. Bulut koruma sistemi iki farklı modülle çalışır. Birinci modülde saldırı izleme ve kayıt altına alma işlemleri gerçekleşirken ikinci modülde saldırı önleme işlemleri gerçekleştirilir. Sunucuların normal işleyişinde işlem süreleri ve kullanıcı sayıları kayıt altına alınır. Olası bir saldırı durumunda oluşan farklılık hemen kayıtlardaki MAC ve IP kayıt defterinden kontrol edilerek saldırı tespiti gerçekleştirilmiş olur. Saldırı tespiti yapıldıktan sonra ikinci modül devreye girer yapılan tüm saldırılar uyarı havuzunda kayıt altına alınır. Sistemin devamlılığı ile birlikte saldırının engellenmesini sağlamak için, kayıt defterinden tespit edilen anormal kullanıcıların sisteme erişimi engellenir [32].

İletişimde güvenlik için farklı yardımcı programlar kullanılmaktadır. Bunlardan biri CyberGuarder uygulamasıdır. CyberGuarder, sunucu üzerinde kurulu farklı sanal sunucular arasındaki iletişimin güvenliğini sağlamaktadır. Sanal sunucu yöneticileri üzerinden yapılan koruma işleminde sunucular arası iletişim uçtan uca (peer-to-peer) yaklaşımı ile gerçekleşmektedir. Bu yardımcı program ayrıca sanal sunucuların güvenliğinde de etkin rol almaktadır [33].

Bulut bilişim güvenliğinde kullanılan bir diğer yaklaşım safeguard güvenlik yapısıdır. Bu yapı ile fiziksel ve sanal sunuculara karşı gerçekleştirilen sızma ve izleme saldırılarına karşı önlemler alınmaktadır. Sanal sunucu yapılandırmasında direk olarak fiziksel sunucu ile birebir iletişim gerçekleştirilerek aracı yapılar engellenmektedir. Güvenlik senaryosunda yönlendirme, güvenlik duvarı ve paylaşım katmanı yapılandırmaları gerçekleştirilmektedir. Sanal sunucular üzerinden fiziksel sunucu ve ekipmanlarına gerçekleştirilen iletişim tek yönlü ve kontrollü gerçekleştirilerek güvenlik sağlanmaktadır. İletişim sağlanan her kanal belirli sabit tanımlayıcılar kullanılarak veri iletişim kontrolü paket iletişim ve doğrulama algoritmaları ile gerçekleştirilmektedir. Güvenlik duvarı ile paylaşım katmanına sızma engellenmektedir. Literatürde sanal sunucuların bağımsız çalışmasından kaynaklı oluşan sorunların çözümü için geliştirilmiş ve DCPortalsNg olarak adlandırılan sistem ile sanal sunuculara ait veriler haritalanmaktadır. Bu sayede ortak bellek ünitelerinden veri kullanımı eş zamanlı gerçekleşmesine rağmen veri paylaşımı ve erişimi sağlanmaktadır. Veri iletişimi paket olarak gerçekleştirilerek iletişimin doğruluğu kontrol edilebilmektedir. İletişimde herhangi bir ihlal durumunu tespit etmek için SnortFlow yaklaşımı kullanılmaktadır. Bu yapı kayıt defterlerini ve veri trafiğini izleyerek sakıncalı bağlantıların tespitini sağlamaktadır [31].

B. Mimari Güvenlik Sorunlarına Karşı Çözümler

Mimaride karşılaşılan sorunlara çözüm önerileri sorunlara göre sınıflandırılmış ve dört alt başlık altında toplanıp açıklanmıştır.

Sanallaştırmadan kaynaklı sorunlara karşı çözümler

Kurulan her sanal makineye güvenilir bir işletim sitemi kurulmalı, yardımcı güvenlik teknolojileri kurulmalı, sanal sunuculara bekleme veya başlatma şifresi oluşturulmalı, sanal sunuculara görüntü yüklendiğinde gerekli önlemler ve düzeltmeler yapılmalı ve bulut yapısı ile sanal sunucularda güvenlik araçları dahil edilmeli ve uygulanmalıdır [31].

Sanallaştırmada görüntü içerisine saklanan kötücül yazılımları engellemek için Mirage yaklaşımı kullanılmaktadır. Bu yaklaşım ile görüntü dosyaları hata ayıklama filtrelerinden geçirilerek kullanılır. Erişim kontrolü ile sağlanan yapıda görüntü dosyalarının bozuk kısımları veya içerisine saklanmış gizli programları temizleyerek veya engelleyerek sanal sunuculara yüklenmesine olanak tanınır. Görüntü içerisindeki kötücül yazılımlara karşı bir diğer yöntem görüntü dosyalarının AES şifreleme algoritması ile şifrelenmesidir. Oluşturulan şifre yönetim mekanizması ile şifrelenen metin sunuculara yüklendikten sonra yeniden şifre çözümleyerek elde edilir. Farklı bir yaklaşımda çevrimdışı olarak sanal sunuculardaki görüntüler kontrol edilmektedir. Yeniden yüklenerek filtrelemeden geçirilmektedir [34].

Fiziksel ve sanal sunucuların kullanım yetkileri farklı olmalıdır. Sistem erişimi için kullanılan yapı ve güvenilirliği ile sanal sunucuda kullanıcıların farklı yetki, şifre ve korunma yöntemlerine sahip olması güvenilirliği artırmaktadır. Bellek yönetimi ve sistem yönetimi servis sağlayıcı tarafından sağlanmalıdır. Kullanılabilirlik ve hız ile birlikte veriye erişimin saldırılara karşı korunmasının sanal sunucularda değil fiziksel sunucularda sağlanması gerekmektedir. Bunun için sanal sunucuların yönetimi fiziksel sunucu güvenlik ayarları ile kısıtlanmalıdır [35]. Sanal sunucular arası yönetimi sağlayan yapıda, sunucular arası veri iletişimi şifrelenerek gerçekleştirilir [36].

Cloudvisor, sanal sunucu yönetimi içerisinde iç içe sanallaştırma gerçekleştirerek hafif düzeyde güvenlik sağlayan nesne tabanlı sanallaştırma yöntemidir. Ayrışma (decoupling) modeli sanal sunucular çalışır durumdayken ortak kullanım ve donanımın yönetimini sağlayan yöntemdir. CPU, bellek üniteleri ve giriş-çıkış birimleri sanal sunucular arasında ortak kullanılan donanımlardır. Sanal sunucu yönetim birimi ile sanal sunucular arasındaki tüm iletişim ve paylaşım ve ayrışımı sağlayan ve güvenlik önlemlerini alan yapıya cloudvisor denir. Örneğin, sanal sunucu fiziksel sunucuda kaydını şifreli tutar ve gerektiği zamanlarda onaylamak için devreye koyar. Sanal sunucuların kullanım alanını, erişim yetkilerini ve bellek tablolarının yönetimini sağlayarak çakışmaları önler. Şifreleme sanal sunucu üzerinde gerçekleştirilerek içerik koruma gerçekleştirilir. Merkel ağaç ve MD5 şifreleme algoritmalarını kullanır [37].

Kullanım aşamasında sanal sunucularda HyperCoffer güvenlik önlemi alınır. HyperCoffer sadece çekirdek tabanlı yapılan işlemleri güvenli görür ve geriye kalan donanımsal

ve web arayüzlü kullanımlarda önlemler alır. Böylelikle HyperCoffer hem donanımsal hem yazılım olarak güvenliği sağlar. HyperCoffer sanal sunucuların birbirleri arasında ve fiziksel sunucular ile sanal sunucular arasındaki iletişimin kayıtlarını tutarak cloudvisor'dan daha geniş bir güvenlik sağlamaktadır [37].

Cloudsec yöntemi ile sanal sunucular üzerinden fiziksel bellekler ve donanımlar gözlemlenmektedir. İç gözlemeli bu yöntem ile saldırı tespiti sunucu çalışma zamanı içinde gerçekleştirilmektedir. Çekirdek tabanlı izlemede gerekli verilerin işleme alınması sağlanmaktadır. Aşırı değişim gösteren bellek ve işlemci değerlerine karşı önlem almada kullanılmaktadır [30].

Bir başka iç gözlemeli sanal sunucu mimarisi dışsal (exterior) olarak tanımlanmaktadır. Dışsal, özellikle misafir yani geçici sanal sunucular ile güvenli sanal sunucu arasındaki bağlantıyı sağlamada kullanılmaktadır. Misafir sanal kullanıcıların sürücü, bellek ve sistem dosyalarına erişimi engellenir [37]. Sanal sunucuların çalışma zamanında güvenliği sağlamak için bir başka yaklaşım, iletişim ve kopyalamada önlemler alan sanal güvenli platform modulüdür. (virtual trusted platform modula-vTPM). Hosting sağlayan sunucu ile sanal sunucu arasında kullanılan iletişim kanalının güvenilirliğini sağlamak için kullanılır. Aynı kanal birden fazla sanal sunucu tarafından aynı anda kullanılabilir. Şifreleme yapısı, süreklilik, güvenlik ve doğru iletişimi sağlamak için kullanılır [37].

Bellek ünitelerinden kaynaklı sorunlara karşı çözümler

Sanal sunucuların ortak kullandıkları bellek ünitelerinde, güvenliğin sağlanması için anahtarlama ile şifreleme yöntemleri kullanılmaktadır. Verilere erişimi ve sanal sunucuların bütünlüğünü bozmadan; güvenilir şifreleme algoritmaları, bellek yönetim yapısı ile kullanılabilirliği ve doğruluğu kabul görmüş güçlü şifreleme yöntemleri kullanılmalıdır [21].

Seccloud sadece sunuculara veri yüklerken değil sunucu içindeki veriler için de güvenlik sağlamaktadır. Arşivlenmiş verilerin güvenliği de şifrelenerek sağlanmaktadır. Yapısal olarak kullanıcı, bulut ve üçüncü kişiler için farklı anahtar yapısı kullanmaktadır. Sanal sunuculara veri yüklendiği zaman kişiye ait alana verinin kaydedilmesi önemlidir. Sisteme veri yükleyen kullanıcı kendine ait oluşturulan anahtarı ve bulut fiziksel sunucusu anahtarı ile sisteme veri yükleyebilir. Kişiye ait tasarlanmış alana kayıt gerçekleştirildiği zaman şifreli veri anahtarlar ile açılarak doğru adrese ve doğru kişinin erişimine açılır [38].

Sanal sunucularda tutulan veriler genel ise herhangi bir doğrulama ve güvenliğe ihtiyaç duyulmaz. Özel veri ise verinin saklandığı bellek indeksleri ve veri şifrelenerek saklanır. Sunucular bellek yönetimini genel ve özel olmak üzere iki farklı şekilde tanımlamaktadır. Üçüncü bir bölüm indekslerin tutulduğu yerdir ve erişim izni çok kısıtlıdır. Verilerin saklandığı alan da anahtarla şifrelendiği için olası veri kurtarma seçeneklerinde farklı kullanıcıların verilerine erişim engellenmektedir. Kişiye özel alan ve veri şifreleme gerçekleştiğinden, bellekten silinmiş verinin kurtarılmaya çalışılması sonradan o alana yüklenmiş veriye erişimi engeller [39].

Uygulamalardan kaynaklı sorunlara karşı çözümler

Bellek yönetiminde kullanılan anahtarlama yöntemi uygulamaya erişim ve kullanımda da geçerlidir. Uygulamaların güvenilirliği kullanıcıdan çok sistem yöneticileri tarafından sağlanan hizmetlerin güvenilirliği ile sağlanmaktadır. Saldırı modelleri sürekli yenilenmeli ve sistem güncel tutulmalıdır. Kullanılan yazılım ve uygulamaların güvenilirliği kontrol edilmelidir. Normal bilgisayar kullanıcılarının güvenlik duvarı ve antivirüs yazılımlarını kullandığı gibi uygulama hizmeti sağlayan servisler de kontrol edilmelidir. Sızma testleri sistem yöneticileri tarafından yapılarak sistemin güvenlik açıkları tespit edilip düzeltilmelidir [40].

Kimlik yönetimi ve erişim kontrolünden kaynaklı sorunlara karşı çözümler

Kimlik yönetimi ve erişim kontrolü bulut bilişimin yaygınlaşması ve kurumsal anlamda kullanılması için sağlanması gereken bir yapıdır. Şifreleme ve kişiye özel anahtar kullanılarak sisteme erişim ve kimlik denetimi farklı yöntemler ile gerçekleştirilmektedir [41]. Literatürde kişiye özel anahtar yerine kişiye özel ID veya şifrelenmiş kullanıcı adı gibi farklı tanımlar kullanılmıştır. Olası erişim sorunları, sunucu hataları ve uygulamaların çalışmaması gibi sorunlara karşı bulut bilişim, verileri farklı sunucularda yedekleyerek erişim sorununu çözmeyi amaçlamaktadır. Fiziksel sunuculardan kaynaklı oluşacak hataları ortadan kaldırmak için yedekleme ve acil durum yönetim hizmetleri sağlanmaktadır.

Sözleşme ve yasalardan kaynaklı sorunlara karşı çözümler

İnternet tabanlı oluşabilecek suçlara karşı dünya genelinde ortak kabul edilen bir yasa bulunmamaktadır. Bazı ülkelerin oluşturdukları ortak yasalar ve işbirliği geneli kapsamamaktadır. Bu sebeple hem internet ortamında işlenen suçlar hem de bulut bilişimde sunucuların yasa dışı saldırılarda hangi yasalara tabi olacağı bir muammadır. Ülkemizde yayın yapan internet sitelerinin yasa gereği sunucularını ülkemiz sınırları içerisinde tutması ve ülkemiz yasalarına uyması gerekmektedir. Bulut bilişimde veri yüklediğimiz ve kullandığımız uygulamalar için oluşturulmuş ortak bir yasa mevcut değildir. Bu sebeple bazı bulut hizmeti sağlayan kuruluşlar ülke yasalarına uygun altyapı ve fiziksel sunucu özellikleri belirlemektedir. Bulut bilişimde yaşanabilecek herhangi bir yasa dışı saldırı, normal internette işlenen suçlar kapsamında değerlendirilip Türk Ceza Kanunu hükümlerine göre cezalandırılır [42].

V. SONUÇ

Bulut bilişim hızla yaygınlaşan ve gelişen bir teknolojidir. Yakın zamanda bilgisayarlar bellek ünitelerinden bağımsız sanal sunucu bağlantısı sağlayan elektronik devreler ile yönetilecektir. Yerelde sanallaştırma ile kullanılan laboratuvarların yakın gelecekte tüm bilgisayarlar için geçerli olması kaçınılmazdır.

Gelişmekte ve yagınlaşmakta olan yeni teknolojinin karşılaşılan sorunlarının çözümü ve bulut yöneticilerine dikkat edilmesi gereken sorunlar karşısında farkındalık

oluşturmak amacıyla sunulan çalışmada, karşılaşılan sorunlar alt başlıklar halinde sunularak karşılaşılabilecek sorunlara çözüm yolu bulmada kolaylık sağlanması amaçlanmıştır. Literatürde karşılaşılan sorunlara üretilen farklı çözüm önerileri derlenmiştir.

Bulut bilişimin kurumsal şirketler tarafından kullanılmasının ekonomik ve yönetimsel faydalarına ek olarak hizmet standartları, iletişim ve bilişsel faydaları sağlanmalıdır. Kurum çalışanlarını mekân ve zaman probleminden bağımsızlaştıran yeni teknolojinin stratejik faydalarının sağlanmasının güvenliği kurumlar için önemlidir.

Bulut bilişimin maddi ve yönetimsel olarak sağladığı kolaylık ve doğru yönetildiğinde üst düzey güvenlik sağlaması sebebiyle kurum ve kuruluşlarda da hızla yaygınlaşmaktadır. Google ve Windows gibi sektörün önde gelen şirketleri bulut teknolojisine büyük önem vermektedir. Windows Azure, Microsoft Office ve Google Drive gibi uygulamalar ve yatırımlar bunun en büyük kanıtıdır.

Bulut bilişimin temelinde sanallaştırma ve bellek yönetimi yatmaktadır. Sanallaştırmadan kaynaklı karşılaşılan güvenlik sorunları aslında ağ alt yapısında karşılaşılan sorunlar ile benzerlik taşımaktadır. Ancak bulut bilişim ve terimlerinin yeni olması farklı sorunlar gibi algılanmaktadır. Ortak ağda farklı işlemciler arkasında ve farklı güvenlik duvarları ve korunma yöntemlerine sahip bellek ünitelerinde ihlal durumu bulut bilişimden çok da farklı değildir. Sunulan çalışmada belirtildiği gibi bellek ünitelerinde ve veri kurtarma ve sanal sunucuya ait alan tablosu oluşturmada karşılaşılan zorluklar mevcuttur. Her sistemde olduğu gibi eğitimli ve donanımlı kişiler tarafından doğru kurulmuş ve yönetilen sistemlerde bu sorunlar çözümlenebilmektedir.

Bireysel kullanıcıların kiralama yöntemiyle kullandıkları sunucuların kontrolü farklı birim ve kişilerin elindedir. Oysa ki kurumsal bulut bilişim kendi sunucularını kullanabilmektedir. Maliyet politikası, iş zaman adam yönetimi ve teknolojiyi takip etmenin verdiği prestijlerin getirisi çok daha fazladır.

Bulut bilişimde ortak fiziksel sunucu kullanılması ve günümüzde tuş kaydedici (keylogger) gibi sunucu üzerinden kullanıcı bilgilerine ve kullanımlarına erişim sağlayan programların varlığı tehdit oluşturmaktadır. Kurumsal şirketlerin tamamının sunucu tabanlı ve yerel ağ tabanlı çalıştığı düşünüldüğünde, bulut mimarisinin daha güvenli olduğu gözlemlenmektedir. Bağlanmış olunan ağ kişinin bilgisayarının dünya ile bağlantısında mutlak uğrak noktası olduğu için her geçiş kaydedilip izlenebilmektedir. Oysa ki bulut bilişimde kullanıcı, sunucu üzerinde direk çıkış almakta ve güvenlik seviyesini kontrol etme yetkisi sunulmaktadır. Bu sebeple bellek yönetimi, veri güvenliği ve ihlal durumları açısından kurumsal bulut yapısının daha güvenli olduğu gözlemlenmektedir.

Kişilerin verilerini sakladığı ve büyük şirketlerin kontrolünde olan bulut teknolojisinin kötü amaçlı kişiler tarafından ele geçirilmesi ve veri gizliliği ihlallleri teknolojiye olan güveni azaltmıştır. Ancak kurumsal anlamda kullanılan bulut teknolojisinde sunucuların kontrolünün kurum görevlilerinde olması güven sorununu azaltmaktadır.

Sunucu kiralama yöntemi ile kullanılan bulut erişiminde sunucuların güvenliğinin kurumlarda olmaması, kiralama teknolojisine olan güveni azaltmaktadır. Bellek ve sunucu teknolojisinin gelişmesi ve ucuzlaması ile donanıma erişimin kolaylaşması, kurumların kendi alt yapısını kurmasına sebep olmakta ve karşılaşılan donanımsal sorunlara karşı kurum ihtiyaçlarına göre çözüm üretilebilmektedir.

Karşılaşılan sorunların teknolojinin doğru yönetilmesi ve kullanılması ile çözülebilecek sorunlar olduğu tespit edilmiştir. Kurumun ihtiyaçlarına göre tasarlanmış alt yapı ve bellek yönetimi ile maddi fayda sağlandığı gözlemlenmiştir. Bulut bilişimin sağladığı faydalara ek olarak doğru yönetildiği ve gerekli güvenlik önlemleri alındığında kullanıcı odaklı hataların azaldığı tespit edilmiştir.

KAYNAKÇA

[1] U. A. Kashif, Z. A. Memon, A. R. Balouch, J. A. Candio,"Distributed Trust Protocol for IaaS Cloud Computing", 12th International Bhurban Conference on Applied Sciences & Technology (IBCAST) Islamabad, Pakistan, 275-279, 2015.

[2] M. Armbrust, A. Fox, R. Griffith, , A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, "Above the Clouds: a Berkeley view of Cloud computing" Technical report, 2009.

[3] I. Foster, Y. Zhao, I. Raicu, S. Lu, "Cloud computing and grid computing 360-degree compared", Grid Computing Environments Workshop, Austin, 1-10, 2008

[**4]** Internet: HP Cloud research, http://www.hpl.hp.com/research/ cloud.html Erişim Tarihi: 20.05.2015

[5] J. Gantz, D. Reinsel, "IDC's digital universe study (sponsored by EMC)", 2012.

[6] Internet: Amazon Elastic Compute Cloud, http://aws.amazon.com/ec2 Erişim Tarihi: 12.05.2015.

[7] Internet: Microsoft Azure, http://www.microsoft.com/windowsazure Erişim Tarihi: 10.05.2015.

[8] Internet: Eucalyptus Public Cloud, http://open.eucalyptus.com/wiki/ Documentation Erişim Tarihi: 12.02.2015.

[9] J. D. Lasica, "Identity in the Age of cloud computing: The Next-generation Internet's Impact on Business, Governance and Social Interaction" The Aspen Institute, 2009.

[10] S. Hackett, "Managed Services: An Industry Built on Trust", IDC, 2008.

[11] J. Staten, "Hollow Out The MOOSE: Reducing Cost With Strategic Rightsourcing", Forrester Research Inc., 2009.

[12] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, A. Ghalsasi, "Cloud computing - The business perspective" Decision Support Systems, 51 (1), 176-189, 2011.

[13] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, I. Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility" Future Generation Computer Systems, 25(6), 599-616, 2009.

[14] L.M. Vaquero, L. Rodero-Merino, J. Caceres, M. A. Lindner, "Break in the clouds: Towards a cloud definition", SIGCOMM Computer Communications Review, 39, 50-55, 2009

[15] M Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, "Above the Clouds: A Berkeley View of Cloud Computing" UC Berkeley Reliable Adaptive Distributed Systems Laboratory White Paper, 2009.

[16] G. Motta, N. Sfondrini, D. Sacco, "CLOUD COMPUTING: A business and economical perspective", International Joint Conference on Service Sciences, Shanghai, 18-22, 2012.

[17] W. Voorsluys, J. Broberg, R. Buyya, "Cloud Computing Principles and Paradigm", John Wiley and Sons, 2011.

[18] D. M. Parrilli, "Legal Issues in Grid and cloud computing, Grid and Grid Computing", Grid and Grid Computing, 97-118, 2010.

[19] M. G. Avram, "Advantages and challenges of adopting cloud computing from an enterprise perspective", Procedia tecnology, 12, 529-534, 2014.

[20] A. Corradi, M. Fanelli, L. Foschini, "VM consolidation: a real case based on openstack cloud", Future Generation Computer Systems, 32, 118–127, 2014.

[21] D.A.B. Fernandes, L.F.B. Soares, J.V. Gomes, M.M. Freire, P.R.M. Inácio, "Security issues in cloud environments: a survey", International Journal of Information Security, 13(2), 113–170, 2014.

[22] K. Hashizume, D.G. Rosado, E. Fernndez-Medina, E.B. Fernandez, "An analysis of security issues for cloud computing", Journal of Internet Services and Applications, 4(5), 1–13, 2013.

[23] W.A. Jansen, "Cloud hooks: Security and privacy issues in cloud computing", 44th Hawaii International Conference on System Sciences (HICSS), USA, 1–10, 2011.

[24] B. Liu, E. Blasch, Y. Chen, A.J. Aved, A. Hadiks, D. Shen, G. Chen, "Information fusion in a cloud computing era: a systems-level perspective", IEEE Aerospace and Electronic Systems Magazine, 29(10), 16–24, 2014.

[25] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, A. Ghalsasi, "Cloud computing the business perspective", Decision Support Systems, 51(1), 176–189, 2011.

[26] K.S. Rao, P.S. Thilagam, "Heuristics based server consolidation with residual resource defragmentation in cloud data centers", Future Generation Computer Systems,50, 87-98, 2015

[27] M. Song, "Analysis of risks for virtualization technology" Applied Mechanics and Materials, 539, 374–377, 2014.

[28] S. Subashini, V. Kavitha "A survey on security issues in service delivery models of cloud computing", Journal of Network and Computer Applications, 34(1), 1–11, 2011.

[29] J. Szefer, E. Keller, R.B. Lee, J. Rexford "Eliminating the hypervisor attack surface for a more secure cloud", in Proceedings of the 18th ACM Conference on Computer and Communications Security, 401–412, Chicago USA, 2011.

[30] Internet: Open Web Application Security Project Top 10-2013, The ten most critical Web application security risks, https://www.owasp.org/index.php/Top10 Erişim Tarihi: 08.05.2015.

[31] Internet: Cloud security alliance, security guidelines for critical areas of focus in cloud computing v3.0, https://cloudsecurityalliance.org/ guidance/csaguide.v3.0.pdf Erişim Tarihi: 08.05.2015.

[32] F. Lombardi, R.D. Pietro, "Secure virtualization for cloud computing", Journal of Network and Computer Applications, 34(4), 1113–1122, 2011.

[33] J. Li, B. Li, T. Wo, C. Hu, J. Huai, L. Liu, K.P. Lam, "Cyber-guarder: a virtualization security assurance architecture for green cloud computing", Future Generation Computer Systems, 28(2), 379–390, 2012.

[34] J. Wei, X. Zhang, G. Ammons, V. Bala, P. Ning, "Managing security of virtual machine images in a cloud environment", in Proceedings of the ACM Workshop on Cloud Computing Security, USA, 91–96, 2009.

[35] M. Kazim, R. Masood, M.A. Shibli "Securing the virtual machine images in cloud computing", in Proceedings of the ACM 6th International Conference on Security of Information and Networks, 425–428, Türkiye, 2013.

[36] D. Jeswani, A. Verma, P. Jayachandran, K. Bhattacharya, "ImageElves: rapid and reliable system updates in the cloud", IEEE 33rd International Conference on Distributed Computing Systems (ICDCS), 390–399, Philadelphia, 2013.

[37] F. Zhang, J. Chen, H. Chen, B. Zang, "Cloudvisor: retrofitting protection of virtual machines in multi-tenant cloud with nested virtualization", in Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles, 203-216, 2011.

[38] L. Wei, H. Zhu, Z. Cao, X. Dong, W. Jia, Y. Chen, A.V. Vasilakos, "Security and privacy for storage and computation in cloud computing,", Information Sciences, 258, 371-386, 2014.

[39] Y. Tang, P.P. Lee, J.C.S. Lui, R. Perlman, "Secure overlay cloud storage with access control and assured deletion" IEEE Trans.on Dependable Secure Computing, 9(6), 903–916, 2012.

[40] S.K. Sah, S. Shakya, H. Dhungana, "A security management for cloud based applications and services with diameter-AAA", IEEE International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), Ghaziabad, 6–11, 2014.

[41] S. Ruj, M. Stojmenovic, A. Nayak, "Decentralized access control with anonymous authentication of data stored in clouds", IEEE Transactions on Parallel & Distributed Systems, 25(2), 384–394, 2014.

[42] M.L. Hale, R. Gamble, "Secagreement: advancing security risk calculations in cloud services" IEEE Eighth World Congress on Services, Honolulu, 133–140, 2012.

KİŞİSEL, KURUMSAL VE ULUSAL BİLGİ GÜVENLİĞİ FARKINDALIĞI ÜZERİNE BİR İNCELEME

S.E. Erol, E.B. Ceyhan ve Ş. Sağıroğlu

Özet — Bilgi sistemleri son yıllarda kamu ve özel sektörün en fazla yatırım yaptığı alanlardan biri olarak göze çarpmaktadır. Kritik altyapı sistemleri olarak kabul edilen altyapılar (elektrik, su, telekomünikasyon, bankacılık vb.) da dahil olmak üzere hayatın hemen her alanı bilişim sistemleri ile yönetilmektedir. Bu gelişmeler siber saldırıların da hızla artmasına ve çok farklı yöntemlerle uygulanmasına ortam sağlamaktadır. Bilgi sistemlerinin güvenliğinin sağlanabilmesinde en önemli unsur insandır ve ancak insanların bilgi güvenliği farkındalığının yükseltilmesi ile bilgi sistemlerinde güvenlikten söz edilebilir. Bu çalışmada öncelikle saldırganın bir sonraki adımının tahmin edilebilmesi için bir siber saldırı yaşam döngüsü ortaya konulmuştur. Tahmin edilen saldırıların ulusal, kurumsal ve kişisel düzeyde önlenebilmesi için ise bilgi güvenliği farkındalık döngüsü dinamik bir süreç olarak ortaya konulmuş ve örnek olaylar kapsamında değerlendirilmiştir.

Abstract — Information systems have been developed in last years and public/private companies have been investing on much in this area. Especially, critical infrastructures that are affecting public life such as banking, barage systems, hydroelectric plant systems, telecomunication systems etc. are managing by information systems. This type of management systems allow hackers to reach and damage critical infrastructures so the number of cyber attacks increasing in this area. To defend information systems from hackers the most important component is human being. To be able to secure information systems all people should have high security awareness. In this paper firstly, cyber attack life cycle modelled to be able to predict attackers next step. Then, an information security awareness cycle modelled and evaluated with examples to prevent attacks that can be effective in national, institutional and personel level security.

Anahtar Kelimeler — bilgi güvenliği; farkındalık; bilgi; siber saldırı; bilgi güvenliği farkındalık döngüsü; siber saldırı yaşam döngüsü; güvenlik.

Keywords — information security; awareness; information; information security awareness cycle; cyber attack life cycle; cyber attack; security.

I. GİRİŞ

Günümüzde bilgi teknoloji sistemlerinin bilgiye her yerden ve merkezi olarak erişimi mümkün kılmasından dolayı bilgi sistemleri hayatın tüm alanlarında etkin olarak kullanılmaktadır. Günümüzde e-devlet uygulamalarının da kullanıma sunulması ile resmi işlemlerden yasal işlemlere, eğlenceden eğitime kadar bir çok alanda bilgi teknolojileri gündelik hayat içerisinde kendisine yer edinmiştir.

Bu denli hızlı gelişen bilgi teknoloji sistemlerinin gelişimine ve tanımlarına göz attığımızda öneminin her geçen gün arttığı görülmektedir. Özellikle sistemlerin birbirine bağımlılığının artmasıyla iş dünyası ve kamu hayatı hızla artan sayılarda ve çok çeşitli saldırılara maruz kalmaktadır. Bilginin basılı, elektronik ortamda, tabelalarda, konuşmalarda vb. birçok şekilde bulunması, bilgi paylaşımlarının yaygınlaşması ve farklı bir çok yöntem ile gerçekleştirilmesi saldırıların yöntem ve çeşitliliğinin artmasına da imkan sağlamaktadır. Verinin paylaşımı ve sürekli erişime açık olması nedeniyle bilginin gönderen kaynaktan alıcıya kadar gizlilik içerisinde, bozulmadan, yok edilmeden, değiştirilmeden, başkaları tarafından ele geçirilmeden ve bütünlüğü sağlanmış bir şekilde iletilmesi bilgi güvenliğinin sağlanması için temel kriterlerdir. Kamu hayatını düzenleyen sistemler açısından bakıldığında bu kriterlerin önemi çok daha açık bir biçimde karşımıza çıkmaktadır.

Dünyada giderek yaygınlaşan e-devlet uygulamaları ve gelişen web ortamı dünyayı çok hızlı bir dönüşümden geçirmektedir. Bu yaklaşımların yaygınlaşması siber ortamdaki saldırganların motivasyonlarını da yükseltmiştir. Genel olarak politik, ticari ve bireysel olarak sınıflandırılan motivasyonlar siber ortamın sağladığı erişim imkanlarıyla da birleşerek siber saldırının sınırlarını yok etmiştir. Dünyanın herhangi bir bölgesinden başka bir bölgeye çok düşük maliyet ve tanınmazlık perdesinin arkasına sığınarak saldırılar yapmak ve ülke veya ticari kurumları zarara uğratmak günümüzde mümkün hale gelmiştir. İnternet ortamının getirdiği faydaların yanı sıra dış tehdit kavramı da oldukça yaygınlaşmıştır. Kritik bilgi altyapısına sahip SCADA sistemlerini etkileyen olayların 2001 öncesi %50'si kullanıcı hatası olarak etiketlenmiş ve sadece %29'u dış etkilere bağlanmıştır. 2004 yılında, bu oranlar aniden değişmiş ve %66 dış olaylar ve %22 kaza olarak rakamlar değişmiştir [1].

Günümüzde saldırılar ve saldırganların yöntemleri çok çeşitlenmiştir. Dolayısıyla takip etmek neredeyse imkansız hale gelmiştir. Ancak bilginin değerinin her geçen gün daha da artması ve sistemlerde oluşan hasarların giderilmesi maliyetlerinin koruma maliyetlerinden çok daha yüksek olması sebebiyle kurumlar ve kişiler bilgi güvenliğine dikkatlerini yöneltmişlerdir. Bu çalışmada bilgi güvenliği farkındalığının davranışa dönüştürülebilmesi için siber saldırı yaşam döngüsü ve bilgi güvenliği farkındalık döngüsü modellenerek dinamik süreçler olarak ortaya konulmuştur. Tekrarlayan ve birbirini izleyen adımlardan oluşan bu döngülerin anlaşılabilmesi ve içselleştirilebilmesi için öncelikli olarak bilgi ve bilgi güvenliği kavramları, siber saldırı yaşam süreci ve sınıflandırılması, bilgi güvenliği tehditleri açıklanmış, bilgi güvenliği farkındalığının eksik olması durumunda ortaya çıkan kişisel, kurumsal ve ulusal bilgi güvenliği ihlallerine ilişkin örnek olay değerlendirmeleri yapılarak bilgi güvenliği farkındalık döngüsünün önemi ortaya konulmuştur.

II. BİLGİ VE BİLGİ GÜVENLİĞİ KAVRAMLARI

Bilgi; fiziksel veya sanal ortamlarda yer alan, hayatımızı kolaylaştırma, düzenlenme, saklanabilme ve çeşitli iletişim araçları vasıtasıyla hedeflenen alıcılara iletilme özelliklerine sahip anlamlı, işlenmiş veriler bütünü [2] olarak tanımlanmaktadır. ISO-IEC 17979'da bilgi, iş dünyasının önemli varlıklarından biri olarak, kurumun iş yaşamını devam ettirmesi için korunması gereken varlıklardan biri olarak tanımlanmıştır. Özellikle sistemlerin birbirine

bağımlılığının artmasıyla iş dünyasında hızla artan sayılarda ve çok çeşitli saldırılara maruz kalmaktadır. Yine ISO-IEC 17979'da bilginin filmlerden konuşmalara kadar bir çok farklı şekilde bulunabileceği, paylaşım ya da depolanma yöntemi gözetilmeksizin her zaman uygun şekilde korunması gerektiği belirtilmiştir. Bilgi güvenliğinin kamu ve özel sektör arasında yoğun veri alış verişi olmasının bilgiye erişimde kontrolleri zorlaştırdığı, bilginin kamu, özel sektör ve kritik altyapılar için çok önemli olduğu aktarılmıştır. Literatürde bilgi güvenliği bilginin bir varlık olarak doğru teknoloji ile doğru amaç ve yöntemler kullanılarak tüm platfomlarda başkaları tarafından elde edilmesinin engellenmesi, oluşabilecek zararlardan korunması [3], sayısal ortamda bilgilerin saklanması ve iletilmesi esnasında güvenliğin sağlanabilmesi için bilginin güvenli bir ortamda işlenmesine yönelik yapılan tüm çalışmalar [2] gibi ifadelerle tanımlanmaktadır. Genel bir değerlendirme yapıldığında; gizlilik, bütünlük, erişebilirlik ve önceden tahmin edip önlem alma kavramlarının ortak paydalar olduğu görülmektedir. Verinin paylaşımı ve sürekli erişime açık olması bilginin gönderen kaynaktan alıcıya kadar gizlilik içerisinde, bozulmadan, yok edilmeden, değiştirilmeden, başkaları tarafından ele geçirilmeden ve bütünlüğü sağlanmış bir şekilde iletilmesi çok büyük önem arz etmektedir. Bu şartların sağlanabilmesi için ise ortaya çıkabilecek tehditlerin önceden tespit edilerek önlemlerinin alınması gerekmektedir.

Kamu hayatını düzenleyen sistemlere kişisel, kurumsal ve ulusal bilgi güvenliği açısından bakıldığında bilgi güvenliğinin günümüzde hangi noktalara ulaşabildiğini görmek mümkündür. Estonya'nın 26 Nisan 2007'de Bronz Asker heykelini kaldırmasıyla dünyada siber savaş kavramı bir gerçekliğe dönüşmüş ve Rusya yanlısı gruplar tarafından gerçekleştirilen DoS saldırılarıyla Estonya Hükümeti, kamu kurumları ve bankacılık hizmetlerine ait bir çok internet sitesi hizmet dışı kalmıştır. Estonya konuyu NATO'nun gündemine taşımış, dünya bir savaşın eşiğine gelmiştir. Benzer olarak Rusya Gürcistan ile savaşırken eş zamanlı olarak siber saldırıları da başlatmış ve Gürcistan devlet kurumlarının bilgi sistemlerini uzun süre erişilemez hale getirmiş ve zarara uğratmıştır [4].

Türkiye'de halihazırda Gürcistan ve Estonya örneklerinde belirtilen uluslar arası ölçeklerde bir siber saldırı yaşanmamış olmasına karşın, UYAP'ta henüz yargılaması başlamamış gizli bir kovuşturmaya ilişkin verilerin deşifre olması durumunda kararlara tesir edecek delillerin karartılması, zanlıların kaçması gibi sonuçlar doğurabileceği, ya da askeri bir bilgi sisteminden harekata ait bilgilerin sızdırılması veya değiştirilmesi gibi senaryolar değerlendirildiğinde sonuçlarının çok vahim seviyelerde olabileceği açıktır.
Bahsedilen risklerin varlığı, ortaya çıkardığı dezavantajların yanı sıra kurumların varlıklarını gözden geçirip açıklıklarına yönelik önlemler üzerinde düşünmesine ve tedbirler almaya yönelik çalışmalara başlamasına sebep olmuş, kurumların bilgi sistemlerine yönelik farkındalığın artması sürecini de hızlandırmıştır.

III. BİLGİ GÜVENLİĞİ FARKINDALIĞI KAVRAMSAL DEĞERLENDİRME

Bilgi güvenliği farkındalığı kişisel ya da kurumsal güvenliğin sağlanabilmesi için bilgi güvenliğine yönelik tehditlerin ve sonucunda oluşabilecek durumların kavranmasıdır. Bu farkındalık sayesinde kullanıcıların karşısına çıkan kötücül uygulamalara, linklere ve yazılımlara karşı davranışlarında bilinçli bir tutum sergileyerek saldırganların bilgi sızıntısı yapmasına ya da kullanıcının kendini zorda bırakabileceği, veri, itibar kaybedebileceği durumlara karşı kendini korumasıdır. Bir başka deyişle bireylerin bilgi güvenliğinin ne olduğunu ve neden önemli olduğunu bilmeleri teknolojik tüm önlemlere rağmen insanın bilgi güvenliğinin en uç noktasında bulunduğunun kavranması açısından önemlidir.

Kurumlarda çalışanların kurumun bilgi güvenliği politikalarına uyumluluğu önemli bir sosyo-organizasyonel kaynak olarak ortaya çıkmıştır [5,6]. Çünkü çalışanlar bilgi güvenliği konusunda en zayıf halkadır [7,8]. Benzer şekilde internet, akıllı cihaz vb. kullanımlarında da kişi ancak bilgi seviyesi ile orantılı şekilde önlemler alabilir. E-devlet uygulamalarının getirisi olarak ulusal sistemler de artık dış tehditlere açık hale gelmiştir. Tehditlerin kişisel, kurumsal ve ulusal bilgi varlıklarını hedef alacak şekilde farklılıklar göstermesi, bilgi güvenliği farkındalığının 3 ana başlık altında değerlendirilmesini gerektirmektedir. Tüm başlıkların da hedef kitlesi insan olmasına rağmen izlenmesi gereken yol ve yöntemler ciddi anlamda farklılaşmaktadır.

Kurumların uzun yıllar yoğun çaba ve emek harcayarak sahip oldukları en değerli varlıkları olan bilginin, güvenliğin temel bileşenleri olan; gizlilik, bütünlük ve erişilebilirliğinin sağlanması için etkin bir şekilde korunması gerekmektedir. Kurumsal düzeyde bilgi güvenliği, kurumun sahip olduğu ürün ya da sunduğu hizmetin devamlılığının sağlanabilmesi için bilgi varlıklarının muhtemel saldırı ve tehditlere karşı korunması olarak ifade edilmektedir [9]. Bilgi varlıklarının bahse konu tehditlere karşı güvenliğinin sağlanabilmesi için üç temel sürecin bütüncül bir yaklaşımla uygulanması gerekmektedir [2]. Bu süreçlerden ilki, planlama, strateji ve politikaları kapsayan yönetsel süreç, ikincisi, virüsten koruma, yedekleme gibi teknik işlemleri kapsayan teknolojik önlem süreci, üçüncüsü ise kullanıcı eğitimlerini kapsayan bilgi güvenliği farkındalık sürecidir. Kurumlarda sistematik bir yaklaşımla bilgi güvenliği sağlanamadığı durumlarda kurumun saygınlığını kaybetmesi, borçlanması ve maddi zarara uğraması gibi sonuçlar ortaya çıkabilmektedir [10].

Kurumlarda çalışanların bilgi güvenliği farkındalığı, etkili bilgi güvenliği yönetim sistemlerinin çok önemli bir parçasıdır [11]. Kişisel ve kurumsal bilgi güvenliği olarak iki ana başlık altında yapılan değerlendirme sonrasında kurumlarda bilgi güvenliği farkındalığı da kendi içinde genel farkındalık ve bilgi güvenliği politikaları farkındalığı olarak ele alınabilir [12]. Genel bilgi güvenliği, bir çalışanın bilgi güvenliği konusunda temel bilgiler ve potansiyel problemler ve bunların etkileri hakkında fikir sahibi olmasıdır. Bilgi güvenliği politikası farkındalığı ise çalışanın kurumun güvenlik politikasını bilmesi, politika içinde yer alan gereksinimleri ve hedeflerini anlamasıdır [12]. Örnek olarak; çalışanlar parolalarının güvenli olması gerektiğini bilebilirler ancak kurumsal olarak parola belirleme ve yönetim politikalarını ve nasıl uygulayacakları konusunda bilgi eksiklikleri olabilir. Bu durumda politikaların eyleme dönüştürülmesi yönünde eksiklik olduğu görülmektedir. Kurumsal bilgi güvenliğini arttırmak amacıyla verilen eğitimlerde, kurumların en değerli varlığı olan bilginin korunması konusunda kurum çalışanları ve bilgi etkileşiminde bulundukları kişilerin de sorumluluklarını anlamaları hedeflenmelidir [13]. Bir

kurumda çalışan bireylerin düzenli olarak farkındalık eğitimleri almaları gerektiği ISO 27001:2005 standardında da Ek A 8.2.2 maddesinde, Bilgi Güvenliği Farkındalığı Eğitimi ve Öğretimi başlığı altında, "Kuruluştaki tüm çalışanlar ve ilgili olan yükleniciler ve üçüncü taraf kullanıcılar, kendi iş fonksiyonları ile ilgili kurumsal politikalar ve prosedürler hakkında gerekli farkındalık eğitimini düzenli olarak almalıdırlar." [14] şeklinde belirtilmektedir.

Kurumsal bilgi güvenliği farkındalığı daha çok kurumun standart ve politikalarının kullanıcılara öğretilmesi ve ortaya çıkabilecek riskli durumlar ve karşılaşılan tehditlere yönelik olarak yapılması gerekenlerin üzerinde durularak gerçekleştirilebilmektedir. Ancak, kişisel bilgi güvenliği farkındalığı denildiğinde internet ortamına çıkılan andan itibaren her türlü tehdit kapsam alanına girmektedir. Genelde kurum çalışanları bir şekilde denetleme, belge imzalatma gibi uygulamalar ile güvenliğe ilişkin en azından fikir sahibi olmaktadır. Ancak, internet ve bilgisayarı kullanan her bir birey için kişisel bilgi güvenliği farkındalık eğitimleri kesinlikle verilmelidir. Özellikle sosyal paylaşım sitelerini (facebook, twitter, blogger, linkedin vb.) kullanan insanlar kendi istekleriyle farkında olmadan iş bilgileri, kişisel bilgiler gibi özel bilgileri internet ortamında herkesin erişebileceği şekilde paylaşmaktadır. Bu durum bilgisayar korsanlarına sosyal mühendislik yöntemlerini de kullanarak dolandırıcılık gibi suç faaliyetlerini gerçekleştirmelerine olanak sağlamaktadır.

Ulusal bilgi güvenliği ise Ulusal Bilgi Güvenliği Teşkilatı ve Görevleri Hakkında Kanun Tasarısında; "Ulusal güvenliği ilgilendiren, yetkisiz ellere geçtiği taktirde devletin güvenliğini tehlikeye sokabilecek veya devlet aleyhine kullanılabilecek her türlü bilgiyi, üretim, kullanım ,işlenme saklanma, nakledilme ve imha sırasında yetkisiz kişilerin erişimine ve olası her türlü fiziksel ve elektronik müdahaleye karşı korumaya; bilgiye erişim ve kullanıma ait usulleri açık şekilde belirlemeye ve bilgiyi gerektiğinde hazır bulundurmaya yönelik tedbirler [15]" olarak tanımlanmıştır. Genel olarak değerlendirildiğinde en büyük tehdit kişisel bilgi güvenliğine yönelik olarak gerçekleşmektedir. Farkındalık seviyesi arttıkça tehdit miktarı da azalmakta, ancak tehditin gerçekleşmesi durumunda ortaya çıkan kayıp ters orantılı olarak artmaktadır. Bu nedenle kritik altyapılara, ülkenin ulusal güvenliğini ilgilendiren alanlarda yapılacak bir saldırı kaotik bir ortama sebep olacak olmasından dolayı dikkatle ele alınmalı ve en üst seviyede önlemler alınmalıdır.

Sosyal mühendislik saldırılarının odak noktasında insan vardır. Saldırganlar tarafından sistem güvenlik önlemleri aşılamadığı zaman en etkili yol sosyal mühendislik saldırıları ile erişim hakkı elde etme yöntemleri olmaktadır. Sosyal mühendislik saldırılarındaki başarı riskini indirgemenin en önemli yolu kişilerin bu konudaki farkındalığını, bilgi ve becerisini artırmaktır.

Kişisel, kurumsal ya da ulusal seviyede bilgi güvenliği farkındalığı oluşturabilmek için sistematik bir döngü dahilinde hareket edilmelidir. Bu döngü insanların davranışlarında değişiklik oluşturmalıdır. Dijle'nin Türkiye'de eğitimli insanların bilişim suçlarına yaklaşımına ilişkin yaptığı bir çalışmada, yazılımlar aracılığıyla verilerinin çalındığını düşünen kullanıcı sayısı %51,7 iken, açıklıklara ilişkin firmalar tarafından yapılan güncellemelerin kullanılamadığı ve savunmasız bir ortamı meydana getiren lisanssız yazılım kullanım oranı %75 olarak ortaya konulmuştur [16]. Farkındalık seviyesi ile davranışın ayrı yönlerde hareket edebildiği göz önüne alındığında farkındalık üzerine yapılan eğitim ve çalışmalarda davranış değişikliği hedeflenmelidir. Davranış değişikliği ise ancak tekrarlayan ve dinamik bir döngü ile güvenlik kültürü olarak kullanıcılar tarafından içselleştirilebilir.

Şekil 1'de gösterilen bilgi güvenliği farkındalık döngüsü doğrultusunda kullanıcılar öncelikle karşılaşabilecekleri tehditlere ilişkin bilgi sahibi olarak tehditlerin yeteneklerini kavramalıdır. Dönemsel olarak ortaya çıkan ya da sanal ortamda sürekli var olan tehditlerden kendisine karşı harekete geçebileceğini değerlendirdikleri tespit edilerek, bilgi varlıklarını koruyabilmek adına bu tehditlerden korunma yöntemlerini öğrenmeli ve en uygun olanını seçerek uygulamalıdır. Tehditin alınan korunma önlemlerine rağmen gerçekleşmesi durumunda kullanılan cihaz ve bilişim ortamı üzerinde nasıl etkiler bırakacağı, kötücül ortamı ve oluşturduğu hasarı nasıl saptayacaklarını bilmeli ve hangi düzeltici işlemleri uygulamaları gerektiği konusunda deneyim sahibi olmalıdırlar. Bu adımdan itibaren olası saldırılarda zarara uğramamak için güncel olası tehditler sürekli olarak izlenmelidir. Ancak bu şekilde dinamik bir farkındalık döngüsünü içselleştiren kullanıcıların sahip olduğu bilgi güvenliği farkındalık seviyesinin davranışlarını şekillendirilebileceği değerlendirilmektedir. Aksi taktirde edinilen bilgi ve tecrübeler kalıcı olmayacak, kısa bir süre sonra yeni tehditlerle beraber bilgi güvenliği ihlalleri ortaya çıkacaktır.

Şekil 1: Bilgi Güvenliği Farkındalık Döngüsü

Bilgi güvenliği farkındalığına ilişkin bir diğer önemli konu eğitimlerin genel kapsamlı yapılmamasıdır. Ağırlıklı olarak teknik, bilgi sistemlerinde çalışan personelin eğitilmesinin bilgi güvenliği açısından yeterli olduğu düşünülmektedir. Ancak bilgi sistemleri ile ilgili görevlerde çalışan kişilerde görevleri gereği farkındalık seviyesinin daha yüksek olduğu ortaya konulmuştur [17]. Bu nedenle bilgi sistemleri ile alakalı olmayan personelin daha çok risk altında olduğu, çalıştığı kurum ve birimle alakalı olarak da kurumsal ya da ulusal seviyede bilgi güvenliği için bir tehdit unsuru olabileceği de göz önünde bulundurulmalıdır.

IV. TEHDİTLERİ ANLAMAK: SİBER SALDIRI YAŞAM SÜRECİ VE SALDIRI SINIFLANDIRMASI

Kullanıcı ya da çalışanların bilgi güvenliği farkındalığının arttırılabilmesi için saldırıların hangi kaynaklar tarafından gerçekleştirildiği, bir saldırının hangi adımlardan oluştuğu, saldırıların türleri ve tasniflendirilmesi ile ilgili genel bilgi birikimi oluşturulmalıdır. Her ne kadar teknik personel kadar konuyu kavrayamasalar da içerik olarak neler olduğunu anımsayabilecekleri bilgi birikimine sahip olmalıdırlar. Günümüzde siber saldırılar sistemlerin sahip olduğu bilgi varlıklarının öneminin ve miktarının artmasıyla kabuk değiştirmiştir. Yabancı devletler, terörist gruplar, endüstriyel siber casuslar ve organize siber suçlular, siber eylemciler (Hacktivist), bilgisayar korsanları (Hackers) saldırılara kaynak teşkil etmektedir [18]. Saldırı ve tehditlerin artmasına paralel olarak güvenlik önlemleri de artmakta, sistemlere erişim eskiye nazaran takım çalışması ve karmaşık bilgilerin çözülmesine yönelik işlemlerle gerçekleştirilmektedir. Sisteme erişimin zorlaşması saldırganların sistematik bir yöntem izlemesini de gerektirmektedir. Tehditleri anlamak ancak siber saldırılar gerçekleştirilirken kullanılan metotların ve adımların anlaşılması ile mümkün olabilir. Bu kapsamda 6 adımlı siber saldırı yaşam süreci sistematik yaklaşımları tanımlayan ve en doğru ifade eden sınıflandırmalardan biri olarak karşımıza çıkmakta ve uygulanmaktadır [19]. Bu süreç sonucunda saldırganlar kazanç elde etmekte ve bir sonraki hedefi belirlemek için harekete geçmektedirler. Bu sayede siber saldırılar Şekil 2'de görülen dinamik bir döngüye dönüşmekte bu da saldırıların sürekli olarak kendini yenilemesini ve geliştirmesini sağlamaktadır.

Şekil 2 : Siber Saldırı Yaşam Döngüsü

Günümüzde Şekil 2'de belirtilen saldırı adımlarının gerçekleştirilebilmesine olanak sağlayan bir çok araç bulunmaktadır. Dolayısıyla artık çok yüksek bilgi seviyesinde olmayan kişiler bile rahatlıkla hedef sistemlere saldırı gerçekleştirebilmektedir.

Siber saldırı yaşam sürecinin ilk iki adımını, çoğu zaman iç içe geçen bilgi toplama ve keşif oluşturmaktadır. Saldırganlar, sistemlere sızabilmek için Host Sweep ve port tarama yöntemlerini kullanırlar. TCP Echo, UDP Echo, ICMP Sweep ile bir ağda bulunan hostları belirlenirken; port tarama atakları ise, açık portların belirlenerek servisler üzerinden saldırı yapılmasına imkan tanır [20].

Açık kaynak kodlu araçlarla kolaylıkla port taraması yapılabilmektedir. NMAP bir çok farklı port tarama tekniğini tek komutla yapmaya imkan tanımaktadır. Port tarama dışında sosyal mühendislik, whois sorguları, pasif saldırılar(ağa yerleşerek trafiğin izlenmesi, ağ topolojisinin çıkarılması vb.), ping okuma, google hacking, shodan gibi arama motorları bilgi toplama ve keşif için etkin olarak kullanılmaktadır.

Zafiyetlerin taranması adımında ise sistemin açıklıklarının bulunmasına yönelik uygulamalar yapılmaktadır. Açıklık; istemeden/kazayla başlatılabilen ya da bilerek suistimal edilebilen zaaflar [21] ya da "bir varlığı tehditlere karşı korumasız hale getiren her türlü unsur (sistem bileşenlerinden, güvenlik politika ve prosedürlerinin yokluğundan, yetersizliğinden veya uygulanmayışından, eksik veya hatalı sistem tasarım ve uygulamalarından, organizasyon yapısı, yönetici ve çalışanların bilgi birikimi ve tutumundan kaynaklı nedenler)" olarak tanımlanmaktadır [22].

Saldırı yaşam sürecinin zafiyet tarama, açıklıkları istismar etme ve sistemin ele geçirilmesi adımlarının içeriğini oluşturan saldırıların, genel olarak 5 ana başlık altında sınıflandırıldığı görülmektedir. Bunlar [23];

1. Amaca dayalı saldırılar: Keşif saldırısı, erişim saldırısı ve servisin engellenmesi saldırısı olarak 3 kategoride toplanmıştır.
2. Yasal sınıflandırma: Siber suç, siber casusluk, siber terörizm ve siber savaş olarak 4 kategoride toplanmıştır.
3. Dahil olma şiddetine göre: Aktif ve pasif saldırılar olmak üzere 2 kategoride toplanmıştır.
4. Kapsama göre: Kötü niyetli büyük ölçekli ve iyi niyetli küçük ölçekli olarak 2 kategoride toplanmıştır.
5. Ağ türüne göre: Mobil adhoc ağları ve kablosuz sensör ağları olarak 2 grupta kategorize edilmiştir.

Sisteme erişime imkan sağlayan açıklıklar KALI benzeri araçlar kullanılarak ya da manuel olarak tespit edilebilmektedir. Bu adımı takiben sistemin açıklıkları istismar edilerek hedeflenen verilerin elde edilmesine ve sistemin ele geçirilmesine yönelik işlemler gerçekleştirilir. Örneğin, web sitelerindeki kodun açıklıklarından faydalanarak kullanıcı girdilerinden anlamlı komutlar türeterek bilgi sızdırma olarak tanımlanan SQL enjekte yapılabilir.

Bir saldırının son adımı ise, sistemde yer alan saldırıya ilişkin izlerinin silinmesidir. Bir ağa sızmayı başaran saldırgan, elde ettiği bilgiler ya da sisteme verdiği zararlar anlaşıldığında kendisine ulaşılmasını engellemek için sisteme bıraktığı izleri silmek isteyecektir.

Tüm bu saldırı adımları uygulanarak dışarıdan yapılan saldırıların yanı sıra insider olarak tanımlanan iç saldırılar da gerçekleşebilmektedir. İçeriden gelen saldırıların bir kısmı aslında kötü niyetli olmayan ancak bilgisayarına zararlı yazılım bulaşmış masum kullanıcılardan kaynaklanırken, kötü niyetli saldırıların %80'inin ise sistemde görevli teknik personel tarafından yapıldığı gözlemlenmiştir [24].

V. ÖNEMLİ TEHDİTLER

Saldırı yaşam süreci ve bu sürecin adımlarında gerçekleştirilen işlemlerin ardından kullanıcı ve çalışanların kurumsal ve kişisel bilgi güvenliğine etkin katılan bir unsur olabilmeleri için öncelikle bütün tehditleri olmasa da en azından genel geçer ve yaygın olan tehditler konusunda bilgi sahibi olmaları gerekmektedir. Bu tehditler alt başlıklar halinde verilmiştir.

A. Hizmet Dışı Bırakma Saldırıları (DoS/DDoS, Botnet)

Siber savaşın en etkili ve en verimli yöntemlerinden birisi Dağıtık Hizmet Aksattırma (DDOS/DOS) ataklarıdır. Bu saldırıda bilgi güvenliğinin özelliklerinden biri olan "erişilebilirlik" hedef alınmaktadır. Bu saldırı sonucu sadece bilgi, para, zaman kaybı değil bazen daha önemli olabilecek itibar kaybı ortaya çıkmaktadır [25]. Bu saldırıda amaç, hedef sisteme cevap veremeyeceği miktarda fazla istek göndererek, bant genişliği, CPU zamanı veya disk alanı gibi kaynakların tüketilmesi ya da yapılandırma bilgileri ile fiziksel ağ bileşenlerinin bozulmasıdır. Dağıtık hizmet engelleme (DDoS, Distributed Denial of Service) saldırılarında ise DoS saldırılarından farklı olarak botnetler kullanılmaktadır. Botnet, yönetici tarafından kontrol edilen savunmasız sistemler, köle bilgisayarlar olarak tanımlanmaktadır. Köle bilgisayarlar vasıtasıyla farklı konumlarda bulunan kaynaklardan tek bir hedef sisteme istekler gönderilmektedir. Kullanılan botnetin büyüklüğüne göre çok büyük ve iyi bağlantılı web sayfalarının bile hizmetlerinin engellenmesi mümkündür [26]. Bunun örneklerini şu ana kadar gerçekleşen iki büyük DDOS saldırında görmekteyiz. Şimdiye kadar tespit edilen en büyük DDoS saldırısı Apple Daily ve PopVote sitelerine 500 gbps trafik ve saniyede 250 milyon DNS sorgusu ile yapılmıştır. İkinci DDoS saldırısı ise NTP kullanılarak gerçekleştirilmiştir. ABD ve Avrupa'yı hedef alan bu saldırıda 400 Gbps seviyesinde bir zirve yakalamıştır [27].

DDoS ile mücadele edebilmek için öncellikle teknik altyapılar ön planda olsa da kurumsal düzeyde oluşturulan güvenlik politikaları ve yapılan risk analizleri sayesinde bu saldırıların önlenmesi mümkün olabilmektedir. İlave olarak ağ üzerindeki bütün cihazlar dikkate alınmalı, zaman ve kaynak tüketen fonksiyonlar belirlenmeli ve gerçek DoS saldırılarına benzer ortam oluşturularak, ağ test edilmelidir. DDoS saldırısına karşı alınabilecek proaktif önlemler şu şekilde sıralanabilir [28]:

- Kullanılmayan servisler kapatılmalı,
- IDS imzası oluşturulmalı,
- DNS zaman aşımı kısa tutulmalı,
- ISS ile irtibata geçilerek ek bant genişliği talebinde bulunulmalı,
- Statik web sayfası kopyası bulunmalı,
- IP ve portu paket özelliklerine göre aktif edebilme, engelleyebilme veya kapatabilme özelliği bulunmalı,
- Default TCP timeout değerlerinin yüksek olması nedeniyle bu değer 1/10 a düşürülmeli veya oturum dolmaya başladıkça timeout değerleri otomatik azaltılabilmeli,
- Bir IP adresinden örneğin 500'den fazla istek geldiyse engelenecekler listesine eklenebilmeli ve IP adresine ait oturum tablosu boşaltılmalı,
- Yasal trafik geçirmiş IP adresine göre beyaz liste/kara liste uygulanmalı,
- Ülkelerin IP bloklarına göre erişim izni verilmeli, (Örneğin 20-30 farklı ülkeden DDos saldırısı ile karşılaşıldığında sadece Türkiye kaynaklı servis sağlayıcılarına giriş izni verilebilir)
- DNS round robin ve TTL değerleriyle oynama yaparak engelleme yapılmalı,
- Sniffer kullanılmalıdır.

B. SQL Enjekte Saldırıları

SQL enjekte açıklıkları web uygulamaları için en ciddi tehdit olarak tanımlanmaktadır. SQL enjekte açığı bulunan web uygulamaları saldırgana uygulamanın tüm veri tabanına erişim imkanı verir [29]. Bu veri tabanlarında kullanıcıların ve tüketicilerin çok hassas bilgileri yer almaktadır ve güvenlik ihlallerinde bu bilgiler çalınma, ifşa edilme, satılma veya dolandırıcılık amacıyla kullanılabilmektedir.

C. Sazan Avlama (Phishing) Saldırıları

Kimlik hırsızlığı olarak adlandırılan bu yöntem, banka, telekomünikasyon şirketleri gibi resmi bir kaynaktan geldiği izlenimi verilerek hazırlanmış e-postalar aracılığıyla kişisel bilgilerin elde edilmesi olarak tanımlanmaktadır. Sosyal mühendisliğin hareket alanında yer alan bu saldırı tipinde kullanıcı farkında olmadan kişisel bilgilerini kötü niyetli kişilere göndermekte ve sonrasında bu bilgiler kullanılarak gerçekleştirilen saldırılara maruz kalmaktadır [3]. Bu saldırı tipine karşı kişisel bilgi güvenliği farkındalık seviyesinin arttırılması ve dikkat en önemli silahlardır.

D. XSS Saldırıları

Kullanıcının veri girişi yapabileceği alanlarda genelde javascript kodları kullanılarak zararlı kod gönderilmesiyle yapılan saldırı türüdür [30]. XSS saldırıları günümüzde hakerlar tarafından web uygulamalarına sızmak için en fazla kullanılan saldırı türlerinden biridir. XSS saldırılarında saldırganın amacı, kullanıcının çerez bilgilerinin veya web sitesi tarafından kullanıcının kimlik doğrulamasının yapılmasına imkan sağlayan hassas bilgilerin çalınmasıdır [31]. Tablo 1'de web ortamında her dört saldırıdan birinin XSS saldırısı olduğu SQL injection saldırısının % 7 olduğu görünmektedir.

Saldırı Türü	Yüzde
XSS	25
Bilgi Sızdırma	23
Kimlik Doğrulama/Yetkilendirme	15
Oturum Yönetimi	13
SQL Enjekte	7
CSRF	6
Diğer	11

Tablo I - 2013 Yılı web uygulama saldırıları [32]

D. Zararlı Yazılımlar

Kötücül yazılımlar bulaştığı bilgisayar sistemindeki donanıma ve dosyalara zarar veren, yazılımların ayarlarını değiştiren,

bilgisayardaki verileri izinsiz olarak başka kişilere gönderen veya bilgisayarı yabancı kişilerin erişimine açan istenmeyen zararlı yazılımlardır. En genel kötücül yazılımlar; Virüsler, solucanlar (worm), Truva atları (Trojan horse), arka kapılar (backdoor), mesaj sağanakları (spam), kök kullanıcı takımları (rootkit), korunmasızlık sömürücüleri (exploit), klavye dinleme sistemleri (keylogger), görüntü yakalama sistemleri (screen logger), tarayıcı soyma (browser hijacking) ve casus yazılımlar (spyware) olarak belirtilebilir [33].

E. Sosyal Mühendislik

Günümüzde bilgi teknolojilerindeki hızlı ilerlemeye paralel olarak güvenlik alanında da teknolojik olarak kullanıcıyı koruyan uygulama ve donanımlar geliştirilmekte, kullanıcılar bilgi güvenliği eğitimleri ile bilinçli hale getirilmeye çalışılmaktadır. Bu noktada saldırganların başarıya ulaşmak için başvurdukları etkili yöntemlerden biri olarak sosyal mühendislik karşımıza çıkmaktadır. Sosyal mühendislik saldırıları, saldırganların isteklerini gerçekleştirmek için insan davranışları ve iletişimlerindeki açıklıkları kullanması [34, 35] ya da kullanıcıların normalde paylaşmayacağı kişisel verilerini kendiliğinden vermelerini sağlamak [36] olarak tanımlanmaktadır. Bu saldırıların en önemli riskleri oluşturduğu ve diğer bilinen saldırılara göre kontrolünün daha zor olduğu değerlendirilmektedir [37]. Bu saldırı türünün önlenebilmesi için kurumsal olarak güvenlik politikalarında düzenlemeler yapılmalı, eğitimlerde kullanıcılara anlatılmalı ve olay gerçekleştiğinde müdahale yöntemlerine yönelik önlemler alınması gerekmektedir.

En sık kullamılan sosyal mühendislik yöntemleri şunlardır:

- Karşı tarafı, güvenilir bir kaynak olduğuna inandırmak,
- Hedef sistemin atıklarını (çöpler, eski donanımlar vb.) bilgi bulmak maksadıyla karıştırmak,
- Ortak tanıdıklar vasıtasıyla yakınlık kurmak,
- Başkasını taklit ederek aldatmak (özellikle telefonda),
- Gizlice, düzmece, zor bir durum oluşturarak yardım ediyormuş izlenimi vermek.

Kurumsal düzeyde sosyal mühendislik saldırılarına karşı sistemin en zayıf halkası olan insanın eğitilerek hata payını en aza indirmek için bazı tedbirler alınabilir. Bu tedbilerin alınmasında amaç insan unsurunun;

- Bilgisini (İnsanlar ne biliyor)
- Tavırlarını (İnsanlar ne düşünüyor)
- Davranışlarını (İnsanlar ne yapıyor) değiştirebilmek ve geliştirebilmek [38] olmalıdır.

VI. BİLGİ GÜVENLİĞİNİN KURUMSAL VE KİŞİSEL OLARAK ÖNEMİ VE ÖRNEK OLAY DEĞERLENDİRMELERİ

Bilgi güvenliği farkındalığı oluşturulması kapsamında en önemli adımlardan biri de dünyada ve ülkedeki güncel gelişmelerin takip edilmesidir. Böylelikle kullanıcılar değişen ve gelişen teknolojik ortamda gerçekleşen saldırılar hakkında bilgi sahibi olarak, bu saldırılardan dersler çıkarmalı ve kendi bilgi sistemleri ya da cihazlarında saldırılara maruz kaldığında korunabilmeli ya da en az zararla kurtulabilmelidir. Bu kapsamda ulusal, kurumsal ve kişisel olarak farkındalık yaşam döngüsünün ihlal edilmesi sonucunda ortaya çıkan örnek saldırılar ve değerlendirmeler sunulmuştur.

A. Stuxnet

İran'ın Buşehr ve Natanz'da konuşlu nükleer tesislerinde gerçekleştirilen nükleer çalışmalarını sekteye uğratmak amacıyla ABD ve İsrail tarafından oluşturulduğu düşünülen, Haziran 2010'da tespit edilen solucan yazılımdır. Dış dünyaya kapalı sistemlerin ve endüstriyel kontrol sistemlerinin de tamamen güvende olmadığını, çeşitli saldırı yöntemleri kullanılarak bu sistemlerin de zarara uğratılabileceğini gösterdiğinden büyük yankı uyandırmıştır. Bu solucan İran'da 62.867 bilgisayara bulaşmış ve bunun için herhangi bir ağa gereksinim duymamıştır [39]. Bu örnekte ulusal düzeyde bir saldırı ile karşılaşılmıştır. Genel anlamda bilgi güvenliği farkındalık süreci doğrultusunda değerlendirildiğinde sistemin değerinin farkında olunduğu bu doğrultuda kapalı bir sistem kurularak bir çok saldırının doğrudan devre dışı bırakıldığı, tehditleri izleme ve anlama adımlarının gerçekleştirildiği görülmektedir. Bu saldırının usb bellek, cd, dvd veya başka harici cihazlarla yayıldığı düşünülmektedir. Kullanıcının harici cihazı sistemde kullanması, sistem yöneticilerinin harici cihazın kullanımına izin vermesi ve sistemin zararlı yazılımı tespit edecek güvenlik yazılımlarına sahip olmaması korunma yöntemlerinin uygulanması adımında yaşanan aksaklıklar olarak karşımıza çıkmaktadır. Solucanın sisteme dahil oluşuna ilişkin, bir çalışanın bedava aldığı veya yerde bulduğu bir usb belleği bilgisayarına takmasıyla başladığı düşünceleri bulunmaktadır. Bu örnek korunma yöntemleri öğrenmenin tek başına yeterli olmadığı, bilginin davranışa dönüştürülerek uygulanmadığı takdirde sistemin güvenliğini tehlikeye attığını göstermesi dolayısıyla son derece güzel bir örnektir.

B. Rusya Kaynaklı Sitenin Kamera Kayıtlarını İzinsiz Yayınlaması

Kasım 2014 tarihinde çıkan bir habere göre Rusya kaynaklı bir internet sitesi dünyanın 250 noktasında bulunan evlerdeki kameraların görüntülerini canlı yayınlamaktadır. Habere göre bu siteye girenler, sadece Türkiye'de ev, bebek odaları, hastane ve devlet daireleri de dahil çeşitli yerlere kurulan 170 kameranın canlı görüntülerini izleyebilmekte ve kamera sahiplerinin bu durumdan haberleri bulunmamaktadır [40]. Bu duruma sebep olarak insanların kamera ilk kurulduğunda otomatik belirlenen default şifreleri değiştirmedikleri dolayısıyla hakerların sistemlere kolayca erişebildiği tahmin edilmektedir. Şifrelerin ya boş bırakıldığı ya da "12345678", "0000000", "88888888" gibi şifreler olduğu görülmektedir. Yaşanan bu durum bilgi güvenliği farkındalık döngüsü kapsamında değerlendirildiğinde kullanıcıların tehditleri izleme ve anlama adımlarında eksiklik gösterdiği görülmektedir. Kişilerde evlerinde bulunan kamera sisteminin özel hayatlarına ilişkin bilgileri deşifre edebileceğine ve saldırganların öncelikle standart şifreleri deneyeceğine yönelik bilgi eksikliği olduğu görülmektedir. Bu ve benzeri bilgi güvenliği ihlalleri ile karşılaşmamak için default şifreler ürün/sistemler kullanılmaya başlamadan önce değiştirilmeli, hatırlanması kolay tahmin edilmesi zor parolalar tercih edilmelidir.

C. Cryptolocker Virüsü

2015 yılında da dünyada ve ülkemizde fidye yazılımları olarak da adlandırılan "cryptolocker" virüsü gündemdeki yerini korumuştur. Bu virüs TTNET tarafından gönderildiği izlenimi verilen bir e-posta ile kullanıcılara ulaştırılmaktadır. Fatura meblağının yüksek olması gibi insan zaafiyetlerinin de kullanıldığı e-postada ekli olan fatura dokümanı PDF dosyası ikonuna sahip olmasına rağmen .exe uzantısına sahiptir. Faturayı görüntülemek amacıyla .exe uzantılı dosyanın çalıştırılmasıyla zararlı yazılım aktif hale gelmekte ve kullanıcının bilgisayarındaki tüm dokümanları şifreleyerek kullanılamaz hale getirmektedir. Şifrenin anahtarı karşılığında kullanıcıdan veya kurumdan hatırı sayılır miktarda para talep edilmektedir. Günümüzde internetin yaygınlaşması ile bir çok insan bu virüsten haberdar olmuştur. Ancak çok sayıda insan dokümanlarının şifrelenmesine engel olamamıştır. Bu örnekte bilgi güvenliği farkındalık döngüsünün tehditleri izleme ve anlama adımlarının gerçekleştirildiği, ancak korunma yöntemlerini öğrenme adımında eksiklik olduğu değerlendirilmektedir. Kullanıcının sahte e-postanın uzantısının kontrol edilerek tıklamaması yeterli olabilecekken cyrptolocker virüsü bilgi eksikliği dolayısıyla kişisel bilgisayarlarda çok etkili olmuştur [41]. Farkındalık dögüsünün etkin olarak kullanıldığı kurumsal bilgisayarlarda belirli uzantıdaki dosyaların erişime kapatılması nedeniyle korunma yöntemleri iyi uygulanmış ve güvenlik ihlallerinin yaşanılmasının önüne geçilmiştir.

Şekil 3: Cyrptolocker Virüsü için Kullanılan E-Posta Örneği [41]

Şekil 3'te gönderen adrese dikkatli bakıldığında e-postanın TTNET'ten gelmediği görülmektedir. Ayrıca ekteki dosyada .exe uzantılı dosya olması yine şüphe ile bakılması gereken bir durumdur. Fatura meblağının da normalden yüksek olması dikkati çekmesi gereken konulardan birisidir. Cyrptolocker virüsü kurumsal ve kişisel alanda etkili olmuş, ancak kurumsal anlamda başarılı olamamıştır.

D. Ankara'da Tapu Bigilerinin Sızdırılması

Ankara'da 1 milyon 568 bin kişinin tapu bilgileri çalınmıştır. Bir emlakçı tarafından bir vatandaşa ilişkin tüm bilgilerin söylenmesi üzerine ortaya çıkmıştır. Organize Suçlarla Mücadele polisi, vatandaşın tapu bilgilerini çalarak 500 TL karşılığında satan suçluları gözaltına almıştır. Bilginin içerden sızdırıldığı tahmin edilmektedir. Bu olayda milyonlarca insanın bilgilerinin sızdırılmasına sebep olan ana unsurun güvenlik eksiklikleri olduğu değerlendirilmektedir [42]. Bilgi güvenliği farkındalık döngüsünün tehditleri anlamak adımında yaşanan eksiklik dolayısıyla bu durumla karşılaşıldığı değerlendirilmektedir. Tapu bilgilerinin siber saldırı yaşam döngüsünün kazanç adımında etkin olarak kullanabileceği bir veri olması nedeniyle tehdit altında olduğu tam olarak tespit edilememiş dolayısıyla farkındalık döngüsünün takip eden adımları da uygulanmamıştır. Bu örnekte çözüm olarak işlenen bilgi önemine göre tasnif edilmeli ve her bir veri uygun seviyede güvenlik önlemleri ile korunmalıdır.

E. TEİAŞ Kurumuna Siber Saldırı

15 Kasım 2014 tarihinde Türkiye'nin kritik altyapılarından biri olan TEİAŞ kurumu siber saldırıya uğramış ve bu durum bir paylaşım sitesinde paylaşılarak kurumun prestiji sarsılmıştır. Bir müdür yardımcısının şifreleri çalınarak sisteme yönetici yetkisiyle girilmiştir. Bu saldırı çalışanların daima sosyal mühendislik saldırısına maruz kalabileceğini göstermiştir. Bu nedenle çalışanların özellikle kişisel bilgilerini paylaşmamaları, sosyal ağlarda kullandıkları şifreleri ve e-posta adreslerini iş yerinde kullanmamaları gerekmektedir. Ayrıca donanımsal olarak da bazı güvenlik zafiyetleri olduğu yapılan araştırma sonucu ortaya çıkmıştır [43]. Bu örnekte kişisel bilgi güvenliği farkındalık eksikliğinin kurumsal sonuçlar doğurduğu görülmektedir. Dolayısıyla kurumların bilgi güvenliğinden söz edebilmesi için çalışanlarının kişisel bilgi güvenliği farkındalık seviyesinin de yüksek olması gerekmektedir. Farkındalık döngüsü kapsamında değerlendirildiğinde kişisel ve kurumsal olarak sisteme uzaktan erişimin bir sorun sahası yaratabileceği ve yetkilendirme seviyelerinin iyi değerlendirilerek uygun önlemlerin alınmadığı, dolayısıyla tehditleri anlama adımında eksiklikler olduğu değerlendirilmektedir.

F. Apple ICloud Skandalı

2014 yılında bazı yabancı ünlülerin hesaplarının heklenmesi sonucu kişisel özel fotoğrafları basına sızdırılmıştır. Bu olayda Apple şirketine sınırsız sayıda parola deneme imkanı sunduğu ve fotoğrafların silinmesine rağmen hesaplarda tutulduğu gibi gerekçelerle dava açılmıştır. Fakat bu vakada diğer bir zafiyet yine kullanıcıların zayıf parola kullanması sonucu sözlük saldırı ile parolalarının kırılabilmesidir. Bir kez daha bilgi güvenliği farkındalık eksikliği sonucu kullanıcılar prestij ve manevi yönden ağır zarara uğraşmıştır [44]. Günümüzde en çok üzerinde durulan ve vurgulanan konuların başında parola güvenliği gelmektedir. Bu ve benzeri saldırılardan kişisel anlamda korunabilmek için parolaların belirlenmesi konusunda karakter sayısı, kombinasyon ve tüm hesaplar için farklı parolalar belirlenmesi hususlarına dikkat edilmeli, periyodik olarak parolalar değiştirilmelidir. Kurumsal bazda güvenliğin sağlanabilmesi için fazla sayıda parola girişi önlenmelidir. Kullanıcıların parola belirlemesi esnasında uyulacak politikalar belirlenerek kullanıcıların politikaya uygun parola belirlemesi zorunlu kılınmalıdır. Bu olay kurumsal ve kişisel bilgi güvenliği kavramlarının iç içe geçtiği bir örnek gibi görünse de farkındalık döngüsünde yer alan korunma yöntemlerinin uygulanması adımında firmanın etkin önlemleri alması durumunda kullanıcıların hata yapmasının da önüne geçebileceği ve bilgi güvenliği ihlallerinin gerçekleşmesini önleyebileceği değerlendirilmektedir.

G. HSBC Bankasına Siber Saldırı

13 Kasım 2014 tarihinde HSBC bankası 2.7 milyon kullanıcısının kullanıcısının kredi kartı ve banka kartı bilgilerinin çalındığını duyurmuştur. İlk kez bir banka siber saldırıya uğradığını kamuoyuyla paylaşmıştır. Bu olayın detayları hakkında fazla bilgi bulunmamasına rağmen, üçüncü parti yazılımların kullanımından kaynaklanan zafiyetlerin sebep olduğu konusunda teyit edilmemiş bilgiler bulunmaktadır. Bu yüzden eldeki zayıf bilgiler ışığında saldırının bireysel kullanıcı hatalarından çok kurumsal politikalardaki eksiklik veya hatalardan kaynaklandığı söylenebilir [45]. Dolayısıyla farkındalık döngüsünün korunma yöntemlerinin uygulanması adımında eksiklikler olduğu görülmektedir.

Örnekler genel olarak değerlendirildiğinde bilgi güvenliğinin %100 sağlanmasının mümkün olmadığı bir gerçeklik olarak karşımıza çıkmaktadır. Bilgi güvenliği farkındalık döngüsünün içselleştirilmesi durumunda kişiler tehditleri izleme, anlama ve korunma yöntemlerini öğrenerek uygulama adımlarını gerçekleştirme konusunda daha başarılı olacaktır. Ancak gene de saldırılar gerçekleşecek ve kişiler zarar görebilecektir. Bahse konu adımların etkinlikle uygulanmasına rağmen ortaya çıkan siber saldırılarda zararın asgari düzeyde tutulabilmesi için kişiler döngünün bir sonraki adımı olan saldırıların etkilerini giderme yönünde harekete geçmelidir. Siber saldırıların bir yaşam döngüsü olduğu göz önüne alındığında değişim ve dönüşüm geçirerek tekrar karşımıza çıkacağı unutulmamalıdır. Bir sonraki saldırı öncesinde tehditler izlenerek proaktif önlemler alınmalıdır.

VII. SONUÇ VE GELECEK ÇALIŞMA

Kurumlar tarafından bilgi güvenliği farkındalığı ancak iyi hazırlanmış ve etkin kullanılan bir bilgi güvenliği yönetim sistemi kurulduktan sonra etkili olacaktır. Aksi takdirde bilgi güvenliğinin bir diğer sorun sahası olan sistemsel ve teknolojik eksiklikler sebebiyle saldırılar ve bilgi sızıntıları meydana gelecektir. Bilgi güvenliği yönetim sistemleri içeriğinde yer alan standart ve politikaların kurum dinamiklerini de göz önüne alacak şekilde oluşturulması ve yönetim kademesi tarafından uygulanması aşamasında destek görmesi büyük öneme sahiptir.

Teknolojik ve yönetsel açıklıkların giderilmesini takiben bilgi güvenliğinin yumuşak karnı olan insan faktörü üzerine yoğunlaşılmalıdır. Bu doğrultuda kullanıcı ve çalışanların etkin bilgi sistemi ve cihaz kullanımı sağlanarak bilgi güvenliği farkındalığını kaybetmeyecekleri şekilde eğitim ve bilgilendirme faaliyetlerine tabi tutulmaları gerekmektedir. Aksi takdirde yapılan yatırımlar küçük hatalar nedeniyle çok büyük maddi zarar ve itibar kaybına sebep olarak kaynakların etkin kullanımını engellemektedir. Buna rağmen verilen eğitimlerin ve sağlanan farkındalığın ortaya çıkan yeni gelişmeler ile tamamen geçersiz ya da yetersiz hale gelebileceği unutulmamalıdır.

Günümüzde farkındalık eğitimleri dönemsel olarak kullanıcıların bilgilerini tazelemektedir. Ancak saldırıların yaşadığı dönüşüm ve gelişim çoğu zaman iki eğitim arasındaki zamandan çok daha hızlıdır. Dolayısıyla kullanıcı bir önceki gün aldığı eğitime rağmen ertesi gün hiç duyulmamış bir saldırının kurbanı olabilir. İşte bu noktada çoğu zaman sosyal mühendislik saldırılarına hareket noktası oluşturduğu için sıklıkla eleştirilen sosyal medya ve mobil uygulamalar bir fırsata dönüştürülmelidir. Farkındalık kavramı kullanıcılara planlanan eğitimlerden ziyade günlük hayatta yaptıkları işlemlerin içerisine entegre olarak sunulmalıdır. Bu kapsamda ülkemizde de bir hareketlilik söz konusudur. Gazi Üniversitesi Bilgi Güvenliği Mühendisliği Ana Bilim Dalı Yüksek Lisans öğrencileri tarafından www.guvenlikicinbirdakika.org alan adı ile bir sosyal sorumluluk projesi başlatılmış ve sosyal medya üzerinden bir dakikalık spotlarla insanların bilgi güvenliği konuları ile etkileşim içerisinde tutulmasına yönelik yayınlar yapılmaktadır. Bu ve benzeri çalışmalar arttırıldığı taktirde kullanıcılar gündelik yaşamlarının içinde bilgiye, istediği zaman ve hızlı bir şekilde erişim sağlayabilecek ve farkındalık kavramı dinamiklik kazanacaktır.

Günümüzde bilgi güvenliği ihlallerinin büyük bir bölümü eğitimlerle, deneyimlerle kazanılan farkındalık seviyesine uygun olarak davranılmamasından kaynaklanmaktadır. Farkındalık sahibi olmanın ancak davranışlarda değişiklik yapıldığında güvenlik konusunda fayda sağlayacağı unutulmamalıdır. Bu da ancak siber saldırı yaşam döngüsü ve bilgi güvenliği farkındalık döngüsünün içselleştirilmesi ile sağlanabilecektir. Ancak bu sayede saldırganın davranışsal olarak nasıl bir yol izleyeceği tespit edilerek, bu yolda saldırganı engellemek için gerekli önlemler alınabilecektir.

Türkiye'nin 1994 yılında tanıştığı internet şuan bir çok kurumsal işlemlerin gerçekleştirildiği bir platforma dönüşmüştür. Bu nedenle kurumlar tarafından son 10 yıldır CIO (Chief Information Officer) adında pozisyonlar oluşturarak hem bilgi sistemlerini yapılandıran ve mevcut operasyonlara entegre eden ve hem de bilgi sistemleri ile ilgili güvenlik ve farkındalığı yöneten ve direk CEO ile çalışan pozisyonlar açılmıştır. Bilgi güvenliği ile ilgili olarak bahse konu yönetimsel düzenlemelerin dışında dünya genelinde iki yönlü bir çalışma hızla sürdürülmektedir. Bilgi güvenliği kavramına ilişkin kanuni düzenlemeler, standartlar ve iş birlikleri çalışmaların bir yönünü oluştururken diğer yandan da teknolojik olarak güvenliğin sağlanmasına yönelik çalışmalar sürdürülmektedir. Tüm bu çalışmalara rağmen bilgi güvenliğinde açıklıklar, saldırılar, istismarlar artarak sürmektedir. Bilgi güvenliği riskleri insan faktörü olayın içine dahil olduğu andan itibaren tekrar ortaya çıkmaktadır. Bunun en temel nedeni güvenlik tehditleri için birçok donanımsal ve yazılımsal yatırımlar yapılmakla birlikte zaman zaman insana yatırım yapılmasının unutulmasıdır. Eğitimler, yayınlar ve sosyal medya yolu ile insanlar bilinçlendirilmeli, bilgileri içselleştirilmeleri ve öğrendiklerini uygulamaları sağlanmalıdır.

Gelecek çalışma olarak, ortaya konulan bilgi güvenliği farkındalık döngüsünün sosyal mühendislik, sazan avlama, zararlı yazılımlar gibi bilgi güvenliği tehditlerinin gerçekleşme süreçlerine uygun olarak test edilmesi ve elde edilecek bulgular doğrultusunda modelin geliştirilerek, güncel hayatta kullanılan uygulama ve araçlarla bütünleştirilmesine yönelik çalışmalar yapılmasının faydalı olacağı değerlendirilmektedir.

KAYNAKÇA

[1] Dumont, D., "Cyber Security Concerns of Supervisory Control and Data Acquisition (SCADA) Systems", IEEE HST 2010 Conference.

[2] G Öztemiz, S., Yılmaz, B. (2013). Bilgi Merkezlerinde Bilgi Güvenliği Farkındalığı: Ankara'daki Üniversite Kütüphaneleri Örneği. Bilgi Dünyası, 14 (1) syf. 87-100.

[3] Canbek, G., Sağıroğlu, Ş. (2006). Bilgi ve Bilgisayar Güvenliği: Casus Yazılımlar ve Korunma Yöntemleri. Türkiye: Grafiker. syf. 168-169

[4] ALTUNDAL, Ö., F., "DdoS nedir?Ne değildir?, http://www.siberguvenlik.org.tr/makaleler/ddos-nedir-ne-degildir, Ağustos 2012.

[5] S. R.Boss, L. J. Kirsch, "The Last Line of Defense: Motivating Employees to Follow Corporate Security Guideliness,"in Proceedings of the 28th International Conference onInformation Systems, Montreal, Aralık 9-12, 2007.

[6] M. T.Siponen, S. Pahnila, A. Mahmood, "Employees' Adherenceto Information Security Policies: An Empirical Study," in New Approaches for Security, Privacy and Trust In Complex Environments, H. Venter, M. Eloff, L. Labuschagne, J.Eloff, and R. vonSolms, Boston: Springer, syf. 133-144, 2007.

[7] K. D. Mitnick, W., L., Simon, "The Art of Deception: Controlling the Human Element of Security", Indianapolis, IN:Wiley Publishing, Inc., 2002.

[8] M. Warkentin, R. Willison, 2009, "Behavioral and Policy Issues in Information Systems Security: The Insider Threat, "European Journal of Information Systems (18:2), syf. 101-105

[9] T., K., Bensghir, (2008). Kurumsal bilgi güvenliği yönetim süreci. URL: www.erzincan.edu.tr/userfiles/file/stratejidb/guvenlik.ppt. Son Erişim Tarihi: 27.03.2015.

[10] H. Cavusoglu, Raghunathan, "Economics of IT Security Management: Four Improvements to Current Security Practices" Communications of the Association for Information Systems (14), syf. 65-75, 2004

[11] H. Cavusoglu, J., Son I., Benbasat, "Information Security Control Resources in Organizations: A Multidimensional View and Their Key Drivers," working paper,Sauder School of Business, University of British Columbia, 2009.

[12] B., Bulgurcu, H., Cavusoglu, I., Benbasat, "Information Security Policy Compliance: An Empirical Study Of Rationality-Based Beliefs And Information Security Awareness1", MIS Quarterly Vol. 34 No. 3 pp. 523-548 / Eylül 2010

[13] E. Şahinaslan, A. Kantürk, Ö. Şahinaslan, E. Borandağ, "Kurumlarda Bilgi Güvenliği Farkındalığı, Önemi ve Oluşturma Yöntemleri", Akademik Bilişim'09 - XI. Akademik Bilişim Konferansı Bildirileri 11-13 Şubat 2009 Harran Üniversitesi, Şanlıurfa

[14] TS ISO/IEC 27001, 2006, Syf. 17

[15] Internet: http://www.yasad.org.tr/hakkinda/Sayfa/ulusal-bilgi-guvenligi-kanun-tasarisi, son erişim tarihi: 04.08.2015

[16] Öğütçü, G.,"E-Dönüşüm Sürecinde Kişisel Bilişim Güvenliği Davranışı ve Farkındalığının Analizi.", Yüksek Lisans Tezi, 2010.

[17] Ilkan, M., Iscioglu, E., Egelioglu, F., Doganalp, A., "Information Security Awareness of Academic Staff Members: An Example of Eastern Mediterranean University School of Computing and Technology", 4th Information Security and Cryptology Conference, 2010

[18] İnternet: United States Computer Emergency Readiness Team "Control Systems Security Program(CSSP)" http://www.us-cert.gov/control_systems/csthreats.html (2011).

[19] Yiğit, T., Akyıldız, M., A., "Sızma Testleri İçin Bir Model Ağ Üzerinde Siber Saldırı Senaryolarının Değerlendirilmesi",2014

[20] Al-Jarrah, O., Arafat, A., "Network Intrusıon Detectıon System using Attack Behavior Classification" , 2014, 5th International Conference on Information and Communication Systems (ICICS).

[21] Stoneburner, G., Goguen, A., Feringa, A., "Risk Management Guide for Information Technology Systems","Recommendations of the National Institute of Standards and Technology", Special Publication 800-30, July 2002.

[22] Özbilen, A., "TCP / IP Tabanlı Dağıtık Endüstriyel Denetim Sistemlerinde Güvenlik ve Çözüm Önerileri", Ankara(2012).

[23] Uma, M., Padmavathi, G., "A Survey on Various Cyber Attacks and Their Classification", International Journal of Network Security, Vol.15, No.5, syf. 390-396, Eylül 2013.

[24] Garuba, M., Liu, C., Fraites, D., "Intrusion Techniques: Comparative Study of Network Intrusion Detection Systems", 5th International Conference on Information Technology: New Generations, 2008.

[25] Zargar, S.T., Joshi J., Tipper, D. (2013). A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks. IEEE Communıcatıons Surveys & Tutorıals, Vol. 15, No. 4.

[26] Canbek, G., Sağıroğlu, Ş. (2006). Bilgi, bilgi güvenliği ve süreçleri üzerine bir inceleme. Gazi Üniversitesi Politeknik Dergisi, 9.3.

[27] İnternet: Largest Ever DDoS Cyber Attack Hits US and European Victims, URL: http://www.ibtimes.co.uk/largest-ever-ddos-cyber-attack-hits-us-european-victims-1435973. Son Erişim Tarihi: 27.03.2015.

[28] Bhuyan, M.H., Bhattacharyya, D.K., Kalita, J., K. (2014). An empirical evaluation of information metrics for low-rate and high-rate DDoS attack detection. Pattern Recognition Letters 51, p: 1–7.

[29] Halfond W., G., Viegas, J., Orso, A. (2006). A classification of SQL-injection attacks and countermeasures. Proceedings of the IEEE International Symposium on Secure Software Engineering, Arlington, VA, USA.

[30] Demirez, K. (2011). Linux Backtrack 5, Türkiye: Nirvana Yayınları.

[31] Klein, A. (2002). Cross Site Scripting Explained. URL: https://crypto.stanford.edu/cs155/papers/CSS.pdf

[32] Application Vulnerability Trends Report 2014. URL: https://www.trustwave.com/Resources/Library/Documents/Cenzic-Application-Vulnerability-Trends-2014/

[33] Çifci, H. (2012). Her Yönüyle Siber Savaş. Tübitak Popüler Bilim Kitapları, Ankara.

[34] Bircan, C. (2014). Sosyal Mühendislik Saldırıları. https://www.bilgiguvenligi.gov.tr/sosyal-muhendislik/sosyal-muhendislik-saldirilari-3.html, Son Erişim tarihi: 27.02.2015)

[35] K. D. Mitnick and W. L. Simon, The art of deception: Controlling the human element of security: Wiley, 2001

[36] TÜBİTAK BİLGEM. Tehditler ve Korunma Yöntemleri. http://www.bilgimikoruyorum.org.tr/?b325_sosyal_muhendislik_saldirilarindan_korunmak, Son Erişim Tarihi: 27.02.2015.

[37] C. Hadnagy, Social engineering: The art of human hacking: Wiley, 2010

[38] Allam, S., Flowerday, S., V., Flowerday, E. (2014). Smartphone information security awareness:A victim of operational pressures. Compute r s & security42, 56 -65.

[39] İnternet: "Stuxnet". URL: http://tr.wikipedia.org/wiki/Stuxnet, Son Erişim Tarihi: 11.04.2015.

[40] İnternet: http://www.hurriyet.com.tr/dunya/27620170.asp. Son Erişim Tarihi: 11.04.2015.

[41] İnternet: http://www.hwp.com.tr/2014/12/17/dikkat-cryptolocker-virusu-bu-sefer-de-turk-telekom-faturasi-ile-geliyor/. Son Erişim Tarihi: 12.04.2015.

[42] İnternet: http://www.hurriyet.com.tr/gundem/27662013.asp. Son Erişim Tarihi: 12.04.2015.

[43] İnternet: "Enerji Bakanlığı: Borçlar silinmedi." URL: http://www.hurriyet.com.tr/gundem/27580556.asp. Son Erişim Tarihi: 12.04.2015.

[44] İnternet: "Apple'dan ilk açıklama". URL: http://www.milliyet.com.tr/unlulerin-ciplak- fotograflarina/dunya/detay/1934735/default.htm. Son Erişim Tarihi: 18.04.2015.

[45] İnternet:" HSBC Türkiye'ye Siber Saldırı Şoku!" URL: http://www.milliyet.com.tr/hsbc-turkiye-ye-siber-saldiri-bilisim-1969049/ Son Erişim Tarihi: 27.03.2015.

Salih Erdem Erol Lisans eğitimini Hava Harp Okulu Bilgisayar Mühendisliği Bölümünde tamamlamıştır. Gazi Üniversitesi Bilgi Güvenliği Mühendisliğinde yüksek lisans eğitimine devam etmektedir.

Eyüp Burak Ceyhan Gazi Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümünde araştırma görevlisidir.
Şeref Sağıroğlu Gazi Üniversitesi Fen Bilimleri Enstitüsü Müdürü ve Bilgisayar Mühendisliği Bölüm Başkanıdır.

KURUMSAL EPOSTA SINIFLANDIRMA VE DEĞERLENDİRME SİSTEMİ

A. YILDIZ, E. B. CEYHAN, Ş. SAĞIROĞLU

Özet — Sunulan çalışmada epostaları, içeriklerine göre anlamlandıran ve sınıflandıran bir sistem geliştirilmiştir. Çalışma kapsamında geliştirilen sistem, kurumlarda eposta hizmetlerini kullanan bütün personellere bilgi güvenliği farkındalığına yardımcı olmak için geliştirilmiştir.

Çalışmada temel bir SMTP istemcisi geliştirilmiş ve kurumsal eposta sunucusuna bağlanarak epostalar, geliştirilen uygulama içerisine alınmıştır. Alınan epostaları analiz etmek için epostaların özgün kaynak içerikleri alınmıştır ve bu kaynak bilgilerinden başlık bilgileri ve içerik bilgileri ayrıştırılmıştır. Ayrıştırılan başlık bilgileri ve içerik bilgileri anlamlandırılmak için uygulamada görsel öğelerle ve sayısal değerlerle desteklenmiştir. Epostaların başlık analizlerinin yanında içerik olarak da değerlendirilmesi için kelimeler temel bir sadeleştirme işleminden geçirilip değerlendirilmiş ve sınıflamaya tabi tutulmuştur. Sınıflandırılan içerikler yeni gelen epostalara örnek set oluşturması için veri tabanına sadeleştirilerek kaydedilmiş ve yeni gelen epostalara; ortak veya benzer kelime sayılarını sayarak içerik benzerliği ölçülüp sınıflandırma tahmini yapılmıştır. Bu sayede ortak bir bilinç oluşturup içerik olarak epostaların taşıdığı bilgiler sınıflandırılabilmektedir ve kurumlarda bilginin değerinin ölçülmesi kolaylaşmaktadır. Yeni gelen epostalar için bir tahmin sunulabilmekte ve kullanıcıların aldıkları epostanın taşıdığı bilgi seviyesini ölçmelerinde yardımcı olunabilmektedir.

Sistemin genel sınıflandırma başarısı %80, istenmeyen eposta sınıflandırma başarısı ise %83 olarak elde edilmiştir. Bu sistem yerel ağda çalışabilen bir masaüstü uygulama olabilme özelliği ile literatüre katkı sağlamaktadır.

Anahtar Kelimeler—Elektronik posta, bilgi güvenliği, bilgi güvenliği farkındalığı, sınıflandırma.

Abstract—In this paper, you can see the study about email classification and evaluation system which works according to email contents and header information. This system is developed to assist information security awareness for all staff types of any organization.

In the study, a basic SMTP application is developed for connecting enterprise email server and getting emails. To analyze received emails from server, content information and header information are extracted and examined. After examination of headers and content, result is provided with visuals and numeric values to facilitate understanding of information value of email content. In addition to evaluation by the email header analysis, email content is subjected to a fundamental simplification process and after this process, words are saved to database to be reference for new incoming emails. Similarity of emails is calculated with common or similar words count between database and new incoming mail. Thus, new incoming emails can be evaluated according to common consciousness that created with words has been saved to database. And the system can provide estimates for all staff of enterprise about value of email content like spam, important or classified etc. Through this system aimed to decrease human factor in information security for enterprise security.

Overall classification success ratio of the system was obtained as %80 and spam email classification success ratio was obtained as %83. This system contributes to literature with being a desktop application that could work in local networks feature.

Index Terms—Electronic mail, Information security, Information security awareness, classification.

I. GİRİŞ

Eposta mesajları başlık bilgileri ve gövde (ana kısım) bilgileri olarak yapılandırılmıştır ve standartlaştırılmıştır. Gövde genel olarak düz metin şeklindedir ancak HTML (Hyper Text Markup Language) ve MIME (Multi-purpose Internet Mail Extensions) formatlarını da barındırabilir. Ayrıca eposta gövdelerine bir takım multimedya ekleri de eklenebilir. Başlık bölümünde ise epostanın yönlendirilmesini ve kimliğini oluşturmayı sağlayan birçok özel alan vardır. Başlık bilgilerindeki özel alanlar arasında; Kimden (From), Kime (To), Bilgi (CC), Gizli Bilgi (BCC), Konu (Subject), Tarih (Date), Dönüş Yolu (Return Path) gibi alanlar vardır. Epostalar iletişim protokolü olarak Simple Mail Transfer Protocol (SMTP) kullanırlar. Diğer taraftan epostaların adresler arasında iletilmesini sağlayan Mail Transfer Agent (MTA) sunucuları vardır. Bu sunucuların bilgileri de başlık bilgilerine eklenir [1].

Kimden başlığında epostayı gönderen kişinin eposta adres bilgisi bulunur. Kime başlığında epostayı alacak kişinin eposta bilgisi bulunur. Bilgi başlığı, epostayı alan kişiden ayrıca bir kopya da buradaki adrese göndermek için bulunan başlık alanıdır, buraya birden fazla eposta adresi yazılabilir ve hepsine bir kopya gönderilir. Gizli Bilgi başlığı da Bilgi başlığı ile aynı şekilde çalışır ancak alıcı kişi Gizli Bilgi kısmındaki alıcıyı/alıcıları göremez. Konu başlığı epostanın konusunu belirten başlık alanıdır. Tarih başlığı epostanın gönderilme/alınma tarihini gösteren alandır. Tarih başlığındaki tarih bilgisi kullanılan eposta uygulaması aracılığıyla epostanın kaynağına eklenir. Dönüş Yolu başlık alanı, epostaların gönderimleri sonrasında iletildi mesajının alınması veya iletimin başarısız olması durumunda başarısız bildiriminin kime yapılacağını belirten başlık alanıdır.

Eposta kullanımı internetin ortaya çıkmasından bugüne kadar hızla artmaktadır. Resmi ve kişisel yazışmalarda, birçok internet sitesi aboneliğinde ya da birçok paylaşımda eposta yazışmaları kullanılmaktadır. Günümüzde; bir günde ortalama 200 milyardan fazla eposta gönderilmektedir [3].

Kurumsal yazışmalarda da birçok yazışma, kayıtlı olarak veya kayıtsız olarak epostalar aracılığıyla yapılmaktadır. Yapılan bu yazışmalarda birçok sınıflandırılmış ya da sınıflandırılmamış bilgi paylaşılmaktadır. Kurumlarda eposta ile paylaşılan bu bilgilerin gizli olmasından veya kritik öneme sahip olmasından dolayı genelde güvenlik zafiyeti ortaya çıkmaktadır. Bu duruma da dolaylı olarak kullanıcı sebep olmaktadır. Kurumlarda birçok alanda eğitimli veya eğitimsiz

çalışan bulunmaktadır. Bu çalışanlar ne kadar eğitilse de veya uyarılsa da zaman zaman insan faktöründen dolayı gözden kaçan durumlar olabilmektedir. Bu gözden kaçan durumları en aza indirmek için akıllı yazılımlar devreye girmektedir. Eposta sınıflandırma ile ilgili yazılımlar genel olarak veri madenciliği teknikleriyle veya gönderen adresin uzantılarına göre sınıflandırma yapabilmektedir. Benzer olarak gelişmiş eposta sağlayıcıları epostaların içeriklerini de analiz edip gelişmiş bilgisayarlarla klasörleme şeklinde sınıflandırma yapabilmektedir [4].

Kurumlarda eposta sınıflandırması sadece varlık yönetimini kolaylaştırıyor gibi görünse de epostalar aracılığıyla halen büyük sayılarda kullanıcılar ve kurumlar, iletilen zararlı virüslerden ve istenmeyen epostalardan zarar görmektedir. Bu zarar, donanımsal seviyede olabildiği gibi birçok ilgisiz epostayla ilgilenip zaman kaybından dolayı ortaya çıkan bir zarar da olabilmektedir [5].

Çalışmanın ikinci bölümünde literatürdeki ilgili çalışmalardan bahsedilmiş, üçüncü bölümde geliştirilen sistem hakkında bilgiler paylaşılmış, dördüncü bölümde geliştirilen sistemin işleyişi detaylandırılmış,

II. İLGİLİ ÇALIŞMALAR

Yapılan literatür taramasında, çalışmaların genel olarak eposta klasörleme konulu olduğu görülmüştür. Bu çalışmalarda genelde istenmeyen posta tespiti ve klasörleme üzerine yoğunlaşılmıştır. Geliştirilen algoritmalarla epostanın spam olup olmadığı ya da spor, haber ve sinema gibi konularla ilgili olup olmadığını anlamaya yönelik çalışmalar yapılmıştır. Bu yaklaşımlar da gönderen adresin uzantısından veya tanınan epostaların sınıflandırılmasından yola çıkılarak yapılmıştır. Ayrıca sınıflandırma ve kümeleme için geliştirilmiş algoritmalar da bulunmaktadır [4]. Benzer şekilde daha gelişmiş olarak istenmeyen epostalardan Botnet tespiti yapmayı amaçlayan çalışmalar da literatürde mevcuttur. Bu çalışmalarda toplu olarak gönderilen istenmeyen epostaların göndericileri üzerinde analiz yapılıp Botnet tespiti yapmak amaçlanmıştır [9].

A. Eposta Sınıflandırma Konulu Akademik Çalışmalar

Araştırmacılar eposta sınıflandırma çalışmalarında genel olarak kullanım kolaylığı ve istenmeyen eposta tespiti üzerine yoğunlaşmıştır. Ama bu hizmeti gelişmiş eposta hizmet sağlayıcıları da varsayılan olarak verebilmektedir. Eposta hizmet sağlayıcıları ve gelişmiş eposta yönetim araçları (Outlook v.b.) da benzer şekilde klasörleme için destek vermektedir.

Yapılan çalışmalarda metin kümeleme üzerine VSM, KNN, Ripper, Maksimum Entropy, Winnow ve ANN gibi gelişmiş algoritmalar kullanılmıştır. Bu algoritmalar sayesinde metin içerikleri kategorilendirilmiş ve klasörlenmiştir. Çalışmalarda, gönderilen epostaların önemsiz eposta olup olmadığı sınıflandırmasına ek olarak ilgilenilen, ilgi dışı veya önemli önemsiz gibi sınıflandırmalar da yapılmaktadır [4]. Benzer şekilde içerik sınıflandırmasında yapay sinir ağları algoritmalarını kullanan çalışmalar da mevcuttur. Bu çalışmalarda web içerikleri çok katmanlı yapay sinir ağı modeli kullanılarak sınıflandırılmış ve bilgi güvenliği açıklıklarının giderilmesine katkıda bulunmuştur [19].

B. Eposta Sınıflandırma Konulu Ticari Ürünler

Kurumsal çözümler için ticari profesyonel çözümler bulunmaktadır. Bunlar eposta sınıflandırması yapabilen veya çok daha gelişmiş şekilde değerli bilgi takibi yapabilen Data Loss/Leak Prevention (Veri kaybı önleme) uygulamalarıdır. Ancak bu uygulamalar da açık kaynaklı olmadıkları için kurumlarda yine bir güven problemi oluşturmaktadır.
Eposta sınıflandırma alanında ticari uygulamalar arasında Boldon James firmasının "E-mail Classifier" uygulaması örnek olarak verilebilir. Boldon James E-mail Classifier uygulamasının yetenekleri şöyledir [17]:

- Bilgi güvenliği politikalarını uygular,
- Güvenlik politikaları hakkında kullanıcı bilincini yükseltir,
- Veri kaybı önleme (DLP) önlemleri geliştirir,
- Yapılandırılmamış bilgi kontrolleri,
- İç ve dış veri sızıntısını önler,
- Microsoft, Windows Hak Yönetimi politikaları, artı-posta şifreleme ve imzalama,
- Otomatik olarak en hassas içeriği koruma,
- Kullanıcı davranışı ve uyum pozisyon görünürlüğünü sağlama,
- Düşük dağıtım ve yönetim maliyetleri sağlama.

Bu konudaki bir diğer otorite de DLP (Data Loss/Leak Prevention) uygulamalarıdır. Bu uygulamalar bilgi güvenliği için çok daha gelişmiş olarak ağ katmanında veya son kullanıcı katmanında değerli bilgilerin sızdırılmasını veya paylaşılmasını engelleyici çok yetenekli uygulamalardır.

Eposta sınıflandırma alanında ticari uygulamalar arasında Boldon James firmasının "E-mail Classifier" uygulaması örnek olarak verilebilir. DLP uygulamalarına profesyonel cevap veren uygulamalar arasında GTB Technologies firmasının "GTB's Complete Data Protection Platform" ürünü örnek olarak verilebilir. Bu uygulama DLP konusunda birçok yerde karşılaşılan bir üründür. Bu ürün; bütün portlarda ve protokollerde, dosya paylaşımlarında, veritabanlarında, veri havuzlarında, Outlook kullanımında, büyük verilerde, mobil cihazlarda, dizüstü ve masaüstü bilgisayarlarda ve bazı bulut çözümlerinde DLP hizmetini sağlayabilmektedir [18].

III. GELİŞTİRİLEN SİSTEM

Geliştirilen eposta sınıflandırma ve değerlendirme sistemi, kurumsal eposta sunucularında alınan önlemlerin ve genel güvenlik duvarı politikalarının yetersiz kaldığı durumlarda veya bu politikaların kapsamının dışında bir istisna oluştuğu durumlarda, bilgi varlıklarının korunması için epostada paylaşılan bilginin içeriğindeki kelimelere göre; gizli veya önemli bilgi olup olmadığı şeklinde, değerini ölçüp kullanıcıya bir tahmin sunarak kullanıcının paylaştığı bilginin değerinin farkına varmasını sağlamaktadır. Sistem bu yönüyle veri kaybı önleme sistemlerine benzerlik göstermektedir. Örneğin kullanıcının dışarıya ileteceği bir epostayı içeriğindeki kelimelerin gizli epostalar sınıfına benziyorsa kullanıcı uyarılıp sehven yapılacak hataların önüne geçilebilir.

Uygulamada sınıflandırma işlemi, kullanıcının epostaları okuyup değerlendirdiği sınıflandırma seçimleri hatırlanarak geçmişte yaptığı davranışlardan yeni gelen epostaları

değerlendirmesi şeklinde gerçekleştirilmiştir. Bu sayede epostalara kullanıcının belirlediği uyarıcı etiketler verilmiş ve kullanıcıyla etkileşime geçilmiştir. Etiketler epostanın taşıdığı bilginin değerinin anlaşılmasına yardımcı olacak şekilde tasarlanmıştır ve kullanıma göre yeni etiketler belirlenebilmektedir. Geri bildirim şeklinde doğrulamayla eposta içerikleri sınıflandırılmış ve sınıflandırılan içeriklerle yeni gelen epostaların benzerliğine ve geçmişteki okunup sınıflandırılan epostalara göre veri tabanı oluşturulmuştur.

Saldırı içeren bir epostanın tespiti durumunda sistem, ortak havuzdan tahmin yaptığı için diğer kullanıcılara bir tahmin sunmakta ve kullanıcıyı uyarmaktadır. Bu uyarı da olası saldırıların önlenmesine veya en kötü durumda bir kullanıcı saldırıya uğrayınca bir sonrakilerin uyarılmasını sağlamaktadır. Aynı şekilde eposta içeriğinde paylaşılan bilginin değeri de sınıflandırma etiketlerine ortak veya benzer kelimeler sayılarak ölçülmekte ve sınıflandırması için var olan etiketlerden tahmini etiketler sunulmaktadır.

Literatürdeki yaklaşımlar ve uygulamalar bütün epostaları internette paylaşmayı gerektirdiği için kurumsal politikalarla çelişebilir. Kurumlarda gizlilik ve bilgi güvenliği gereği birçok eposta ve doküman dışarıya kapalıdır. Bu durumda kurumların kendi iç yazışmaları için kendilerinin geliştirdiği ve dışarıya kapalı bir uygulama gereksinimi ortaya çıkmıştır. Sunulan bu çalışmada geliştirilen eposta sınıflandırma ve değerlendirme sistemi ile kurumların epostalarını dışarıya açmadan kendi veritabanlarında saklayacakları bilgilerle, kendi belirledikleri kelime gruplarıyla sınıflandırma yapabilen ve bu kelime gruplarına göre bilgi değerini ölçebilen bir uygulamaya sahip olacaklardır.

A. Geliştirilen Sistemin Amacı

Sunulan çalışmada epostaları içeriklerine göre anlamlandıran ve sınıflandıran bir sistem geliştirilmiştir. Çalışma kapsamında geliştirilen sistemde temel amaç, kurumsal eposta hizmetlerinde alınan ve gönderilen epostaların değerlendirilmesi ve anlamlandırılmasıyla kullanıcılara önermelerde bulunarak paylaşılan bilginin önemi ve bilgi güvenliği seviyesinin anlaşılmasında kolaylık sağlamaktır. Bu sayede kullanıcıların bilgi güvenliği seviyesi ve farkındalığının arttırılması sağlanmıştır. Yapılan çalışma kapsamında Windows Form teknolojisi kullanılarak Microsoft Visual C# yazılım diliyle bir uygulama geliştirilmiştir. Bu uygulama ile elektronik posta sunucusundan epostalar alınarak sınıflandırılmakta ve uygulama tarafından daha önceki sınıflandırma sonuçlarına bakılarak tahmini bir sınıflandırma yapılabilmektedir.

B. Sistemin Çalışması

Geliştirilen sistem genel olarak makine öğrenmesi alt yapısı yaklaşımıyla kurgulanmıştır. Sunucudan alınan epostalar uygulama tarafından önce ön işlemden geçirilip yalınlaştırılmış ve veri tabanına kaydedilecek formatta düzenlenmiştir. Ekler, bağlaçlar ve rakamlar gibi anlam ağırlığı olmayan kelimeler çıkarıldıktan sonra kullanıcıdan epostanın sınıf niteliği hakkında bilgi alınıp, kalan kelimeler veritabanına kaydedilmiştir. Yeni gelen epostaların analizi için de veritabanında oluşturulan kelime havuzuna benzetim yapılarak en çok hangi sınıftaki kelimelere benziyorsa o epostanın sınıf bilgisi tahmini sınıf olarak sunulmuştur. Bu sayede örneğin daha önce gizli olarak sınıflandırılmış bir epostanın benzeri bir eposta alınırsa, kullanıcı epostanın değeri hakkında uyarılmakta ve o epostayla iletişim kurmadan önce bilinçlendirilmesi sağlanmaktadır.

C. Sistemin Etiketleme Yapısı

Çalışma kapsamında epostalar için belirli sınıflar tasarlanmış ve bu sınıflara renkler tahsis edilmiştir. Kullanıcılar bu sınıfları kendilerine göre özelleştirebilirler. Bilgi güvenliği uzmanlarının veya yetkili kullanıcıların daha önceden belirlediği sınıflandırılmış içerikler bu etiketleri taşımaktadır ve yeni gelen epostalar da bu içeriklerin etiketleriyle tahminlenecektir. Bu çalışma kapsamında etiketlerin renk yapıları,

- İSTENMEYEN POSTA: Siyah
- ÖNEMLİ: Turuncu
- TASNİF DIŞI: Yeşil
- GİZLİ: Sarı
- HİZMETE ÖZEL: Mavi
- VİRÜS: Kırmızı

olarak belirlenmiştir.

D. Sistemin Değerlendirme Yapısı

Geliştirilen sistemde epostanın kaynak verisi incelenip tehditlerin algılanması sağlanmaktadır. Kaynak verisiyle mesaj kimliği, epostanın ulaşmasına kadar oluşan gecikme süresi, epostanın üzerinden geçtiği sunucu bilgileri ile ayrıntıları ve kaynağı HTML veya RAW formatında görüntüleme özellikleri sunulmuştur.

Değerlendirme yapısında mesaj kaynağının el ile incelenmesi için de ayrıştırılıp kullanıcıya sunulması sağlanmıştır. Bu sayede kullanıcının değerlendirdiği epostaların içerikleri ön işlemden sonra kaydedilmiş ve tahminleme analizi için referans olarak kullanılmıştır.

Değerlendirilip analiz edilecek olan yeni gelen eposta, önceden veritabanına kaydedilmiş sınıflandırılmış kelimelerle aynı ön işlemden geçirilmiş ve uzaklık hesaplayan algoritmalarla benzetim yapılmıştır. Uzaklık belirleyen algoritmalar için uygulamada tasarlanan algoritmaya ek olarak, aynı kökten gelen veya birbirine benzeyen kelimelerle benzetimi hassaslaştırmak ve doğruluk oranını arttırmak [6] için Levenshtein Distance uzaklık hesaplayıcı algoritma da kullanılmıştır.

Uygulama kapsamında tasarlanan algoritmada, ön işlemden geçirilen eposta gövdesindeki kelime listesiyle veri tabanındaki kelime listeleri karşılaştırılır. Karşılaştırma sonucunda gelen eposta hangi eposta sınıfındaki kelimelere daha çok benziyorsa, o epostanın sınıf verisi tahmini sınıf olarak sunulur. Bu algoritmaya ek olarak Levenshtein Distance algoritmasıyla da opsiyonel olarak destek verilmiştir. Kullanıcının ilgili formdaki Levenshtein algoritmasını da kullanması için gereken kontrolü seçmesiyle benzetim detaylandırılır ve harf benzerlikleri sayılır, en fazla %20 farklı kelimeler aynı sayılır ve benzetim güçlendirilmiş olur.

IV. GELİŞTİRİLEN SİSTEMİN İŞLEYİŞİ

Sistem için bir SMTP istemcisi geliştirilmiş ve bu istemciye bağlanan eposta adresinden epostaların alınması sağlanmıştır. Alınan epostalar gelen kutusu benzeri bir veri yapısında listelenmiştir ve detayların verildiği paneller eklenmiştir.

Bu alt yapı için .NET yapılarından Windows Form uygulamaları seçilmiş ve proje oluşturulmuştur. Oluşturulan bu projede .NET SMTPClient sınıfı kullanılmış, e-posta bağlantısı gerekli protokol çerçevesinde gerçekleştirilmiştir. Alınan epostalar bir gelen kutusu yapısına göre listelenmektedir. Listelemede seçilen e-posta için standart gelen kutusu gösterimlerinin yanında bir de bayrak yapısı ve değerlendirme sonucunu belirten etiketler gösterilmiştir.

Geliştirilen uygulamada POP3 bağlantısı yapan açık kaynaklı kütüphane kullanılmıştır ve standart bağlantı protokolleri bu kütüphane ile sağlanmıştır. Uygulamanın giriş ekranında kullanıcının eposta adresi ve parolası alınıp güvenli bir şekilde epostaların alınması sağlanmıştır. Alınan epostalar sistemin gelen kutusunda kullanıcıya sunulmaktadır.

A. Uygulamanın Çalışma Prensipleri

Uygulamanın giriş ekranında gerekli protokoller sağlandıktan sonra giriş yapılıp, gelen ekranda alınan epostalar listelenmektedir. Listelenen epostalara ait gönderen kişi ve gönderme tarihleri gibi temel bilgileri sunulmaktadır. Detayları görüntülenmek istenen eposta için bu listeye gidilip ilgili eposta tıklandığında sistem seçili eposta için analizleri yapıp detayları sağ tarafındaki panelde göstermektedir. Detayların gösterildiği panelde temel olarak kimden, kime, bilgi ve konu gibi başlıkların yanında mesaj ID, epostanın uğradığı sunucular ve gönderim esnasında yaşanan toplam gecikme gibi başlıklar da detaylı bir şekilde sunulmaktadır. Epostanın iletilirken uğradığı sunucuların ayrıntıları verilirken; sunucunun ismi, IP adresi, metodu ve tarih bilgileri verilmiştir. Mesaj ID alanında ise epostaların tekilliğini sağlayan kimlik bilgisi verilmiştir. Ayrıca buradaki kimlik bilgisi ile gönderen kişinin adresinin uzantıları benzemiyorsa muhtemel bir sahte eposta durumu olduğu anlaşılmaktadır. Bu durum mesaj ID alanında belirtilmektedir. Eposta içeriğinin hem HTML formatında gösterimi hem de işlenmemiş RAW formatında gösterimi sağlanabilmektedir. Bu seçimi yapmak için, mesaj kutusunun hemen solunda bulunan tercih kutularındaki seçimi değiştirmek yeterlidir.

Kategori alanındaki aşağı açılan liste ile seçili epostanın sınıflandırma etiketi belirlenip kaydedilebilmektedir. Belirlenen bu sınıflandırma etiketi kaydedilip o etiket ile ilgili veriseti oluşturulmuştur. Yeni gelen epostalar bu kayıtlardan elde edilen verisetine, bir denetimsiz dinamik öğrenme modeli ile tahminleme yapılıp etiket önerisiyle desteklenir. Örneğin gelen epostanın içeriği daha önce tasnif dışı olarak etiketlenen bir eposta içeriğine benziyorsa bu eposta da tasnif dışı olarak etiketlenmektedir.

B. Uygulamada Kullanılan Değerlendirme Algoritmaları

Çalışmada daha önceki bilgi değeri etiketlenen istenmeyen eposta veya zararlı epostaların tanınmasından dolayı ortak bir hafıza oluşturulmuş ve veritabanına kaydedilerek bir veri seti oluşturulmuştur. Veri setini oluşturacak içerik için anlamlı terimler seçilmiştir. Bu yapı için ön işleme algoritması geliştirilmiştir. Bu algoritma kapsamında noktalama işaretlerinin, eklerin, bağlaçların, rakamların ve gereksiz boşlukların çıkarılması şeklinde epostalar ön işleme tabi tutulmuştur. Ön işlemden sonra kaydedilen içeriklerle ortak bir hafıza oluşturmuştur.

Bu ortak hafıza eğitim setimiz olarak nitelendirilebilir. Ayrıca dinamik bir eğitim seti olduğu için ve geliştiği için denetimsiz bir öğrenme algoritması olarak nitelendirilebilir. Benzetim yapılırken daha önce etiketlendirilen epostaların içeriklerinin oluşturduğu veriseti yardımıyla tahminleme yapılmaktadır.

Benzetim için tasarlanan altyapı Google arama motorundaki “Bunu mu demek istediniz?” özelliğinde kullanılan algoritmalara benzerlik göstermektedir. Bu yapıda kullanıcılar yanlış yazdıkları kelimeler için sonuç bulamayınca düzeltip tekrar aramayı denemektedirler. Google da bu arama süreçlerini kaydedip kullanıcıların dillerinde hangi kelime için hangi yanlış yazımlar yapılabilir şeklinde bağıntılar biriktirmektedir. Bu bağıntıları kullanarak bu hizmeti sunmaktadır [19].

Benzetim yapısı için uygulamaya özel bir algoritma tasarlanmış ve Levenshtein algoritması ile desteklenmiştir. Uygulama için geliştirilen algoritmada ön işlemden geçen kelimeler veritabanındaki kelime havuzlarından hangisine daha çok benziyorsa o sınıfa ait olabileceği tahmini sunulmaktadır. Destekleyici algoritma olan Levenshtein algoritması ise dizi yapıları arasındaki benzerlik uzaklıklarını hesaplayan bir algoritmadır. Bu algoritma ile birbirine benzer ya da birbirinin kökleri olan kelimeler bulunabilmektedir [7]-[8]. Uygulamada %20 benzerlik kabul edilebilir bir oran olarak alınmış ve aynı köklü veya benzer kelime olarak nitelendirilerek benzetim hassaslaştırılmıştır. Levenshtein mesafesi ölçülürken kelimelerin arasındaki uzaklık farkı %20 veya daha az ise kabul edilebilir olarak hesaplanmaktadır.

C. Uygulama Sonuçları

Uygulamada geliştirilen benzetim algoritması, Levenshtein destekli iken farklı, yalın haldeyken farklı sonuçlar üretebilmektedir. Bu farklı sonuçlar iki tahminin değerlerinin birbirine yakın olmasından kaynaklanmaktadır. Benzetim hassaslaştırılınca benzer köklü kelimelerle değerler daha detaylı incelenebildiği için tahminin doğruluğu arttırılmıştır. Yeterli eğitim veriseti oluşturulduktan sonra genel olarak uygulamadaki algoritmanın başarısı; daha önce tanımadığı önemsiz veya istenmeyen epostaları tanıyabildiği için başarılı olarak değerlendirilmiştir.

Sistem etiketleri tanımak için belirlenen sınıflara dahil edilen 10’ar eposta ile eğitildikten sonra tahmin başarısı hesaplanmıştır. Genel olarak 60 eposta içinden sistemin sunduğu tahmin sınıf ve okunup karar verilen sınıf arasında 48 eposta içeriği doğru sınıflandırılmıştır ve 12 eposta içeriği yanlış tahmin edilmiştir. Dolayısıyla sistemin genel sınıflandırma başarısı %80 olarak elde edilmiştir. Istenmeyen eposta etiketi için ise 30 eposta üzerinde sistemin başarısı ayrıca hesaplanmış ve Tablo 1’de sunulan karmaşıklık matrisi elde edilmiştir.

		Tahmin	
		Pozitif	Negatif
Gerçek	Pozitif	8	3
	Negatif	2	17

Tablo I. - İstenmeyen epostalar için karmaşıklık matrisi

Tablo 1'de sistemin analiz yapıp tahmin sunduğu 30 adet epostanın istenmeyen eposta etiketi için karmaşıklık matrisi sunulmaktadır. İstenmeyen eposta etiketi sınıflandırması sonucunda 30 epostadan 25'i doğru sınıflandırılmış, 5'i yanlış sınıflandırılmış, dolayısıyla %83 başarı sağlanmıştır.

V. SONUÇ

Sunulan çalışmada eposta içerikleri ve kaynak verileri incelenip anlamlandırma üzerine bir sistem geliştirilmiştir. Sistem eposta sunucusuna bağlanabilen ve eposta alabilen, sonrasında analizler yapıp %80 başarıyla tahminleme yapabilen bir istemci şeklinde tasarlanmıştır. Geliştirilen sistem ile kurumların eposta sunucularından epostalar alınmakta ve normal şartlarda epostaların işlenmemiş kaynağına bakıldığında anlaşılmayan ve analiz etmesi zor başlık bilgileri düzenlenerek kullanıcılara sunulmaktadır.

Geliştirilen sistem genel olarak; kurumsal eposta sunucusunda oturum açmak için giriş bilgilerini aldıktan sonra standart gelen kutusu benzeri bir yapı ile epostaları almaktadır. Sonrasında seçilen epostayı gösterirken başlık detaylarını anlaşılır biçimde göstermektedir. Bir masaüstü uygulaması olup yerel ağda çalışabilen bir uygulama olarak bu özellik literatürde ilktir. Başlık detaylarında kimden, bilgi ve konu başlığı gibi standart alanların yanında epostanın tekilliğini sağlayan mesaj ID, iletilme sürecinde oluşan toplam gecikme süresi ve epostanın iletilinceye kadar uğradığı sunucuların detayları listelenmiştir. Bu başlık detaylarının yanında epostalar kullanıcının belirlediği sınıflara göre de etiketlenebilmektedir. Bu etiketlerle bilgi güvenliği farkındalığı olan kullanıcıların sınıflandırma seçimleri hatırlanarak, bu sistemi kullanan diğer kullanıcıların da yeni gelen epostaları değerlendirmelerine yardımcı olmaları sağlanabilmektedir.

Sunulan çalışma ile kurumsal epostalarda paylaşılan bilgilerin güvenlik farkındalığını arttırmak ve epostalardaki bilgi değerini daha önceki okunup sınıflandırılan epostalara göre hesaplamak için bir sistem geliştirilmiştir. Bu sayede kullanıcıların sehven paylaştıkları bilgilerin değerleri hatırlatılacak ve bilgi güvenliğindeki insan faktörü azaltılabilecektir.

Literatürdeki eposta sınıflandırma çalışmalarında genel olarak kullanım kolaylığı ve istenmeyen eposta tespiti üzerine yoğunlaşılmıştır. Fakat bu hizmeti gelişmiş eposta hizmet sağlayıcıları da varsayılan olarak verebilmektedir. Kurumsal çözümler üreten firmalar güvenlik üzerine yoğunlaşmış profesyonel uygulamalar sunmaktadır. Bu uygulamaların da ciddi ücretler karşılığında kurulumları yapıldığı için tercih edilme oranları azalmaktadır. Aynı şekilde DLP çözümlerinin de bilgi değeri olan ve kritik işler yapan ya da üst bilgilerini korumak isteyen her kurumun kullanması gerekmektedir. Ancak bu uygulamalar da ciddi maliyetler gerektirdiği için her kurum bu imkanlara sahip olamamaktadır. Maliyet konusunda engeller olmasa bile kurumlar epostalarını dışardan bir uygulama üzerinden değerlendirmek istemeyebilmektedir.

Sunulan çalışmada geliştirilen sistem sayesinde kurumlar kendi kurallarını belirleyebilmekte ve dışarıya açık olmayan bir uygulamaya sahip olabilmektedir. Ayrıca bu sistem sayesinde kurumsal bilgi güvenliğinin sağlanmasındaki birincil ve yönetilmesi çok zor olan insan faktörünün etkilerini azaltmak için yardımcı bir uygulama sunulmaktadır. Aynı şekilde kurumsal veya kişisel bilgi güvenliği farkındalığı eksiği olan kullanıcılar da bu uygulama sayesinde eposta başlık yapıları analizi yetkinliğine sahip olabilmekte ve kişisel incelemelerde de farkındalıklarını arttırılabilmektedir. Geliştirilen sistem sayesinde, kurumlarda kullanılan eposta hizmetlerinde paylaşılan, tasnif edilmesi ve değerinin hesaplanması zor olan bilginin değerini hesaplamada, kurumlara epostalarını dışarıya açmalarına gerek kalmadan yardımcı olunabilmektedir.

KAYNAKLAR

[1] P. Resnick, "Internet Message Format", RFC 2822, 2001.

[2] M.T. Banday, F.A. Mir, J.A. Qadri, N.A. Shah, "Analyzing Internet e-mail date-spoofing", Digital Investigation, 7 (3-4), 145-153, 2011.

[3] Internet: "Dünya İstatistikleri, Bugün gönderilen Eposta Sayısı", http://www.worldometers.info Erişim Tarihi: 22.07.2015.

[4] I. Alsmadi, I. Alhami, "Clustering and classification of email contents", Journal of King Saud University – Computer and Information Sciences, 27 (1), 46–57, 2015.

[5] M. A. Al-Kadhi, "Assessment of the status of spam in the Kingdom of Saudi Arabia", Journal of King Saud University – Computer and Information Sciences, 23 (2), 45–58, 2011.

[6] Internet: "Levenshtein Uzaklığı Algoritması", http://en.wikipedia.org/wiki/Levenshtein_distance Erişim Tarihi: 22.07.2015.

[7] Internet: "OpenPOP .NET Kütüphanesi", http://sourceforge.net/projects/hpop Erişim Tarihi: 22.07.2015.

[8] Internet: "Levenshtein Algoritması Uygulaması", http://en.wikibooks.org/wiki/Algorithm_Implementation/Strings/Levenshtein_distance Erişim Tarihi: 22.07.2015.

[9] L. Zhuang, J. Dunagan, D.R. Simon, H.J. Wang, J.D. Tygar, "Characterizing Botnets from Email Spam Records", Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats LEET'08, 2, 2008.

[10] H. Önal, "E-posta Başlıklarından Bilgi Toplama", Bilgi Güvenliği Akademisi, 2009.

[11] L. Daniel, L. Daniel, "E-mail Evidence", Digital Forensics for Legal Professionals, Bölüm 34, 239-244, 2012.

[12] D. Bradbury, "Can we make E-mail Secure?", Network Security, 2014 (3), 13-16, 2014.

[13] M. N. Marsono, M. W. El-Kharashi, F. Gebali, "A spam rejection scheme during SMTP sessions based on layer-3 e-mail classification", Journal of Network and Computer Applications, 32 (1), 236–257, 2009.

[14] G. González-Talaván, "A simple, configurable SMTP anti-spam filter: Greylists", Computers & Security, 25 (3), 229–236, 2006.

[15] W. Goralski, "SMTP and Email", The Illustrated Network: How TCP/IP Works in a Modern Network, Bölüm 21, 535–558, 2009.

[16] Boldon James, "Boldon James E-mail Classifier, Boldon James Product Datasheet, 1-2, 2015.

[17] GTB Technologies, "GTB's Complete Data Protection Platform","https://www.gtbtechnologies.com/en/products/the-gtb-data-loss-platform" About Product, Erişim Tarihi: 22.07.2015.

[18] Otomatik Tamamlama, https://support.google.com/ websearch/answer/106230?hl=tr Erişim Tarihi: 22.07.2015.

[19] E. N. Güven, H. Onur, Ş. Sağıroğlu, "Yapay Sinir Ağları ile Web İçeriklerini Sınıflandırma", Bilgi Dünyası, Cilt:9, No:1, s.158-178, Nisan 2008.

Şeref SAĞIROĞLU, Gazi Üniversitesi Fen Bilimleri Enstitüsü Müdürü ve Bilgisayar Mühendisliği Bölüm Başkanıdır. İletişim için ss@gazi.edu.tr adresini kullanmaktadır.

Eyüp Burak CEYHAN, Gazi Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümünde araştırma görevlisidir. İletişim için eyupburak@gmail.com adresini kullanmaktadır.

Abdurrahman YILDIZ, Gazi Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği ABD'da Yüksek Lisans öğrenimine devam etmektedir. İletişim için abdurrahmanyildiz35@gmail.com adresini kullanmaktadır.

MOBİL PLATFORMLARDA GİZLİ AĞ SALDIRILARININ ÖNLENMESİ VE MOBİL UYGULAMASI

S. Oyucu, H. Polat, İ. A. Doğru

Gazi Üniversitesi, Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü 06500, Teknikokullar, Ankara-Türkiye, e-posta: saadinoyucu@gazi.edu.tr
2 Gazi Üniversitesi, Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü 06500, Teknikokullar, Ankara-Türkiye, e-posta: polath@gazi.edu.tr
3 Gazi Üniversitesi, Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü 06500, Teknikokullar, Ankara-Türkiye, e-posta: iadogru@gazi.edu.tr

Özet — Günümüzde kötü amaçlı yazılımların tespiti siber saldırılarla mücadelede önemli rol oynamaktadır. Kötü amaçlı yazılımların tespit edilmesi ise genel olarak ilk kurulum aşamasında veya uygulama pazarına yüklenirken yapılmaktadır. Fakat ilk etapta herhangi bir kötü amaca hizmet etmeyen bir uygulama, kullanıcının platformuna yerleştikten sonra kendini yenileyerek zararlı hale gelebilmektedir. Bu çalışma kapsamında ilk etapta zararlı olmayan daha sonra kendini güncelleyerek zararlı hale gelen bir saldırı çeşidi ele alınmıştır. Ayrıca bu tür saldırılar kullanıcıdan habersiz internet ağını kullanarak mobil platform üzerindeki bilgileri başka noktalara aktarmak içinde kullanılmaktadır. Bu saldırı türü bazı durumlarda gizli ağ bağlantılarıyla gerçekleştirilmektedir. Bundan dolayı çalışmada mobil platform ile kurulan gizli ağ bağlantılarının ve bu bağlantılardan yapılabilecek zararlı yazılım güncellemelerinin üzerinde durulmuştur. Çalışma sonucunda gizli ağ saldırılarını tespit ve önleme işlemlerinde kullanılacak bir mobil uygulama geliştirilmiştir.

Anahtar Kelimeler — Mobil cihaz güvenliği, Ağ güvenliği, Bilgi güvenliği, Saldırı tespit sistemleri

Abstract — Nowadays the detection of malicious software plays an important role in the fight against cyber attacks. In overall, detection of malware are made during the initial setup process or loading to the application market. But, even if an application does not serve any evil purpose at the first step, it can renew itself as harmful after settling on the user's platform. In this study, non-harmful in the first step then it becomes a kind of harmful attack by self-update is considered. In addition, such attacks are used to transfer information on mobile platforms to another location by using Internet networks without informing users. In some cases, these type of attacks are carried out by secret network connections. Therefore, the secret network connections established with mobile platform and harmful software updates can be done through these links are analyzed in this study. As a result of this study, a mobile application for detection and prevention of secret network attacks was developed.

Keywords — Mobile device security, Network security, Information security, Intrusion detection systems

I. GİRİŞ

Mobil platformlar için geliştirilen uygulamalar sayesinde kullanıcıların günlük işleri kolaylaştırılmıştır. Kullanıcı, mobil platformların uygulama pazarından istediği uygulamayı rahatlıkla indirip kendi platformuna kurulabilmektedir. Uygulama pazarına erişimde ise ağ bağlantılarının farklı türleri kullanılmaktadır [1]. Erişim kolaylığı ve yüksek kullanılabilirliğe sahip mobil uygulamaların sayısı giderek artmaktadır. Bu durum mobil uygulamaları günümüz insanın vazgeçilmezi haline getirmektedir. Bununla birlikte mobil platformlar kötü niyetli yazılımlar için mükemmel bir yayılma ortamı sağlamaktadır [1]. Özellikle mobil ödeme sistemlerinin kullanımının artması saldırganları mobil platformlara saldırmak için güdülemektedir.

Mobil platformlar için geliştirilen birden fazla işletim sistemi vardır. Bunlardan en çok bilinenleri iOS, Android, BlackBerry OS ve Windows Phone'dur. Bu işletim sistemlerinden Android dünya genelinde en çok kullanılan mobil işletim sistemi olma özelliğini devam ettirmektedir [2]. Bu yüksek kullanım tercihi Android'i kötücül yazılımların/geliştiricilerin hedefi haline getirmektedir. Yapılan araştırmalara göre mobil tabanlı kötücül saldırıların %99'u Android tabanlıdır [2]. Android'in kötücül yazılımlar için hedef haline gelmesinde çeşitli faktörler vardır. Bunlardan biri Android'in resmi uygulama marketi olan Play Store'a yüklenen uygulamalara yönelik pasif koruma sergilenmesidir. Diğer önemli faktör ise Android işletim sistem sisteminin açık kaynak kodlu yapısıdır [2].

Kötü amaçlı yazılımlar mobil güvenlik için en büyük tehditlerden biridir [3]. Mobil kötü amaçlı yazılımlar üç ana kategoriye ayrılır. Bunlar virüs, truva ve casus yazılımlardır [4]. Kötü amaçlı yazılımlardaki virüs saldırılarına karşı mobil platformları koruma konusunda iki yaklaşım vardır. İlk yaklaşım cihaz üzerinde anti virüs yazılımı ile tarama yapmaktır. İkinci yaklaşımda ise tek cihaz ile güvenilir bir bilgi işlem ortamı kurmaktır. İlk yaklaşımda çevrimiçi virüs veri tabanına bağlanılarak işlem gerçekleştirmek gerekmektedir. İkinci yaklaşımda ise cihazda güvenilir ortamın kurulması ve korunması için güvenlik hizmetine ihtiyaç vardır [5].

Kötü amaçlı yazılımın bulaştığı uygulamalar kendi kendini kullanıcının haberi olmadan arka planda güncelleyebilmektedir. Bu işlemi gerçekleştirirken paket içeriklerini gizli tutarak tehlikeli paket içeriklerini okuyabilecek uygulamaların izlemesine de engel olmaktadır. Ayrıca profesyonel kötü amaçlı yazılımların yaptığı ağ bağlantıları da gizli ağ bağlantıları olmaktadır. Kötü amaçlı yazılımlar gizli ağ bağlantıları ile kullanıcıdan habersiz kişisel bilgileri başka noktalara aktarabilmektedir. Aynı yöntemle kendi güncel kötü amaçlı yazılım dosyalarını kullanıcının platformuna indirebilmektedir. Bu durumda ise çoğu güvenlik senaryosu saf dışı bırakılmaktadır.

Literatür incelendiğinde mobil platformların güvenliğini sağlamak adına birçok çalışma yapıldığı görülmektedir. Özellikle kötü amaçlı yazılımların tespiti, mobil ağ ve mobil geçici ağların güvenliği üzerine yapılan çalışmalar oldukça fazladır.

Oberheide ve arkadaşları 2008 yılındaki çalışmalarında kötü amaçlı yazılımların tespitinde farklı bir yöntem uygulamışlardır. Kaynak tüketimi analizine göre tespit edilen kötü amaçlı yazılımlar, mobil sistemlerde sanallaştırılmış bulut hizmeti kapsamında ele alınmıştır. Çalışma sonucunda daha az CPU ve bellek tüketiminin olduğu gözlemlenmiştir [6].

Yu ve arkadaşları 2009 yılındaki çalışmalarında mobil güvenlik erişim sistemi geliştirmişlerdir. Çalışmalarında ağ üzerinden iletilen bilgilerin daha güvenli olması için SSL, VPN ve akıllı kart teknolojisini kullanmışlardır [7]. Üst düzey güvenlik gerektiren işlemler için bir öneri olarak sunulmuştur. Ahmad ve arkadaşları 2013 yılındaki çalışmalarında Andoid ve IOS mobil işletim sistemlerini güvenlik açısından karşılaştırmışlardır. Çalışma boyunca yapılan karşılaştırmalar sonucunda güvenlik bakımından IOS işletim sisteminin Android'e göre daha avantajlı olduğu sonucuna varmışlardır [8].

Sun ve arkadaşları 2014 yılındaki çalışmalarında beş açıdan Android güvenliğini ele almış ve bir mobil güvenlik uygulama denetlemesi önermişlerdir. Bu çalışmalarındaki teoriyi desteklemek için üç tekniği prototip olarak tasarlanmışlardır [9].

Penning ve arkadaşları 2014 yılındaki çalışmalarında kötü amaçlı mobil yazılım tehditleri ve saldırıları, kötü amaçlı yazılıma yönelik siber suçlu motivasyonları, mevcut korunma yöntemleri ve bunların sınırlılıklarını özetlemektedirler. Ayrıca çalışmalarında kötü amaçlı mobil yazılım tespiti için bulut tabanlı bir çerçeve önermektedirler [10].

Wang ve Alshboul 2015 yılındaki çalışmalarında mobil güvenlik testleri üzerine odaklanmış ve mobil güvenlik için dört test yaklaşımını incelenmişlerdir. Bunlar adli mobil yaklaşım araçları, sızma testi, statik ve dinamik analizlerdir [4].

Önceki çalışmaların genellikle güvenlik senaryoları ve kötücül yazılımların tespiti üzerine yapılan çalışmalar olduğu görülmektedir. Fakat bazı çalışmalar ağ güvenliği üzerine yoğunlaşmıştır. You ve arkadaşlarının 2013 yılındaki çalışmalarında yaptıkları TCP ve SSL karşılaştırması [11], Rashwan ve arkadaşlarının 2014 yılındaki çalışmalarında oluşturdukları mobil sistemler için güvenlik senaryosu ve bu senaryo içerisinde sürekli iletişim performansı analizinin yapılması [12] daha önce yapılan önemli çalışmalardandır. Fakat bu çalışmada ele alınan konu daha önce yapılan çalışmalardan farklı bir konuya dikkat çekmektedir.

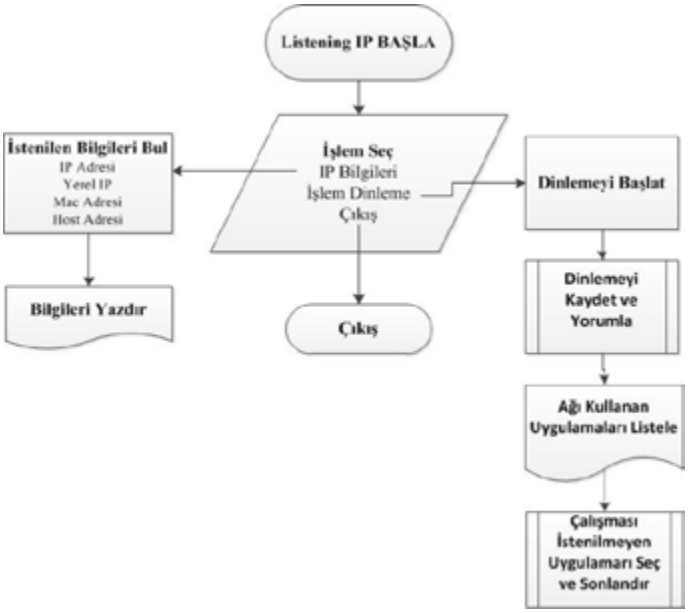
Bu çalışmada ilk etapta hiçbir kötücül unsur içermeyen bir uygulamanın, gizli bağlantılar ile kendini güncelleyip kötücül hale gelmesinin tespiti yapılmıştır. Ayrıca uygulamanın ağ davranış biçimi incelenerek kullanıcıya uyarı verilmesi amaçlanmıştır. Kullanıcı bu uyarılara göre karar verip çalışan uygulamaları kapatabilecektir. Böylelikle kötücül yazılımın neden olduğu kötü amaçlı saldırılardan mobil platformun ve mobil kullanıcının korunması amaçlanmıştır. Bu doğrultuda Android işletim sistemi için bir mobil uygulama geliştirilmiştir.

II. MATERYAL VE METOD

Akıllı mobil sistemlerden önce saldırıların yayılması için en etkili yollardan birinin bluetooth olduğu söylenebilir [13]. Fakat bluetooth bağlantısı olabilmesi için aradaki mesafenin kısa olması ve bluetooth sisteminin açık olması gerekmektedir. Günümüzde internet ağı yaygın olarak kullanılmakta ve hemen her mekânda mobil cihazların kullanabileceği kablosuz bir internet ağı bulunmaktadır. Bu durum kötü niyetli insanların internet ağını kullanarak zarar verme olasılığını giderek arttırmıştır. Bu nedenle bu çalışmada gerçekleştirilen uygulama da örnek teşkil etmesi açısından kablosuz internet ağı kullanılmıştır.

Çalışmada, mobil işletim sistemi olarak Android tercih edilmiştir. Bu tercihin sebebi Android işletim sisteminin açık kaynak kodlu olması, kullanılan cihaz sayısının fazla ve saldırı oranlarının yüksek olmasıdır. Android çoğu zaman işletim sisteminin tüm detaylarına ve gizlenmiş özelliklerine erişim vermemektedir. Çünkü neredeyse hiçbir mobil platform üreticisi Android tabanlı sistemlerde kullanıcılara "root" yetkisi sunmamaktadır. İşletim sisteminin tüm yeteneklerini kullanabilmek için root yetkisine ihtiyaç vardır. Bu nedenle çalışma kapsamındaki test işlemlerinde kullanılacak mobil platformun root yetkisine sahip olması gerekmektedir. Geliştiriciler için oldukça önemli avantajlar sağlayan root yetkisini kullanırken dikkat edilmelidir. Sistemi sürekli root yetkisinde kullanmak başka güvenlik sorunlarına neden olabilmektedir.

Çalışma kapsamında geliştirilen uygulama iki şekilde çalışmaktadır. Birincisi ev ve iş yerlerindeki modemler üzerinde bulunan güvenlik duvarı özelliğini kullanarak mobil platformları korumaya çalışırken gerekli olan dış İnternet Protokolü (IP: Internet Protocol) adresini tespit etmektir. Geliştirilen uygulama sayesinde dış bağlantılarda kullanılan IP adresi bilgisi rahatlıkla tespit edilebilmektedir. Bu bilgi ile kullanıcıların modem üzerinde yapacağı gerekli düzenlemeler sayesinde ev ve iş yerlerindeki bağlantılarını zararlı erişimlerden koruması planlanmaktadır. Diğer kullanım şeklinde ise kablosuz internet ağından gelen sinyaller ve erişimler sürekli kontrol edilecektir. Bu kontroller sonucunda hangi uygulamaya erişildiği ve arka planda kullanıcıdan habersiz hangi uygulamaların çalıştığı liste halinde kullanıcıya sunulmuştur. Kullanıcı, çalışmasını istemediği uygulamaları kolaylıkla kapatabilmektedir. Böylelikle zararlı ağ erişiminin engellenmesi planlanmaktadır. Uygulamanın çalışma mantığı şekil 1.'de gösterilmiştir.



Şekil 1. Geliştirilen uygulamanın işleyiş biçimi

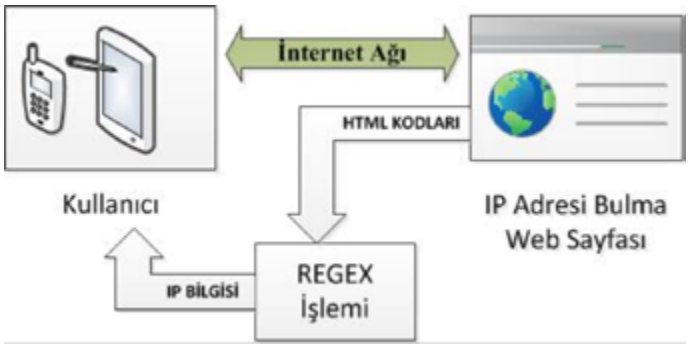
Şekil 1.'de görüldüğü gibi geliştirilen uygulama sayesinde IP bilgileri ile birlikte yerel IP adresi, MAC adresi ve Host adresi de kullanıcıya sunulmaktadır. Burada belirtilen IP adresi dış IP adresini temsil etmektedir. Gizli ağ bağlantılarını görebilmek için ise kullanıcı ağı dinlemeyi başlatmalıdır. Ağ üzerinde

yapılan dinlemeler kaydedilir, yorumlanır ve ağı kullanan uygulamalar bu işlemler sonucunda kullanıcıya listelenir. Kullanıcı listeden istediği uygulamayı seçip kapatabilecektir. Uygulamada gizli ağ bağlantılarını tespit etmek için paket koklama tekniği kullanılmaktadır. Paket koklama işlemi için Tcpdump paket analizcisi tercih edilmiştir. Tcpdump, ağ arabiriminden geçen paketleri kaydedip, pcap (packet capture) destekli herhangi bir araç kullanarak kaydedilmiş paketleri okuma işinde kullanılır [14]. Tcpdump yoğun ağ trafiğinde bile sorunsuz çalışabilmektedir. Bu nedenden dolayı ağı dinlerken Tcpdump tercih edilmiştir. Çalışma kapsamındaki mobil uygulama, Android uygulama geliştirme aracı olan Eclipse üzerinde geliştirilmiştir. Programlama dili olarak ise Java seçilmiştir.

III. UYGULAMA GELİŞTİRME

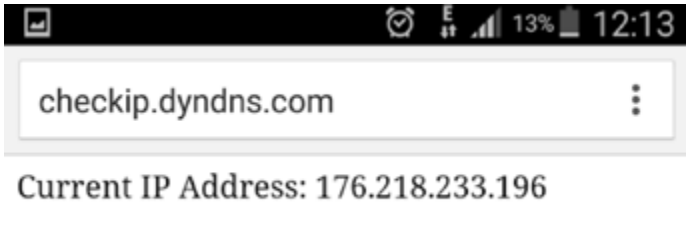
Mobil uygulama geliştirilirken ilk etapta bir emülatör tanımlamak gerekmektedir. Emülatör, mobil bir platformun donanım ve yazılım özelliklerini içeren sanal bir cihazdır [15]. Bir uygulamanın test ve modellemesinin rahat yapılabilmesi için Android Sanal Aygıt yapılandırılmasının yapılması gerekmektedir [15].

Geliştirilen uygulamanın ilk çalışma şekli olan dış IP adresini bulmak için metin içerisinde kural tabanlı anlamlı ifadeler elde etme tekniği Regex (Regular Expression: Düzenli İfade) kullanılmıştır. Regex belirli bir ifadeyi belli bir kalıba göre bulmayı, değiştirmeyi veya parçalamayı sağlayan yazılım algoritmasıdır [16]. Regex bu çalışmada dış IP bilgisini alma işlemin için kullanılmıştır. Şekil 2.'de Regex işleminin uygulamada yaptığı görev açıklanmıştır.



Şekil 2. Regex işlemi

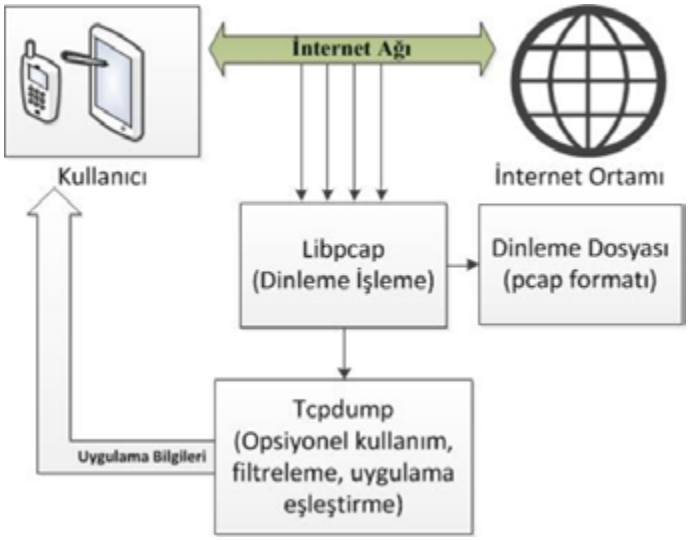
Kullanıcı mobil platformlardan uygulamayı açarak IP bilgilerini bulma işlemini başlatır. Ardından şekil 2.'de görüldüğü gibi mobil platform internet ağı yardımıyla dış IP adresini tespit etme hizmeti sunan bir web sayfasına bağlanır. Bu web sayfasının adresi http://checkip.dyndns.com/ adresidir. Android bir platform ile web sayfasına bağlantı yapıldığında alınan bilgi resim 1.'de gösterilmiştir.



Resim 1. IP adresini bulmak için kullanılan web sayfa bağlantısı

Resim 1.'de görülen web sayfasının Hiper Metin İşaretleme Dili (HTML: Hypertext Markup Language) kodları içerisinden gerekli bilgiler alınmakta ve Regex işlemine tabi tutulmaktadır. Bu işlem sonucu elde edilen dış IP bilgisi kullanıcıya sunulmuştur. Mobil uygulamada kullanıcıya gösterilmek istenilen yazı, metin ve resimler için ise ayrı ayrı hazırlanan TextView'ler oluşturulmuştur.

Gerçekleştirilen uygulamanın ikinci çalışma şekli gizli ağ bağlantılarını kullanıcıya göstermek ve engellemektir. Bunun için hazırlanan uygulamanın açılması ve kullanıcının ağı dinlemeyi başlatması gerekmektedir. Böylelikle ağ analiz verileri kaydedilebilecektir.



Şekil 3. Gizli ağ bağlantılarını dinleme işlemi

Şekil 3.'te görüldüğü gibi kullanıcı mobil platformlardan geliştirilen uygulamayı kullanarak ağı dinlemeyi başlatır. Bu adımdan sonra ağdaki tüm hareketler ve dalgalanmalar libpcap kütüphanesi yardımıyla pcap formatında kaydedilmektedir. Kaydedilen bu bilgiler tcpdump ile işlenerek çalışan uygulamalarla eşleştirilmiştir. Çalışan uygulamalar kullanıcıya liste halinde sunulmuş ve kullanıcının tüm ağ bağlantısı hakkında bilgi sahibi olması sağlanmıştır. Ayrıca kullanıcının, listelenen uygulamalardan istediğini kapatabilmesine olanak verilmiştir.

Uygulama gelişilirken bazı noktalarda Linux komutlarından faydalanılmıştır. Android işletim sistemi üzerinde geliştirilen uygulamanın root yetkisine sahip olması gerekmektedir. Kullanıcı değiştirme işlemi ve pcap uzantılı dosyanın oluşturulması işleminde Linux konutlarından yararlanılmıştır. Ağı dinleme sonucu alınan çalışan uygulamaların listesi daha önce hazırlanan ve kullanıcıya liste yapısını sunmakta yararlanılan "MyCustomBaseAdapter" ile kullanıcıya sunulmaktadır. Böylelikle ağı kullanan uygulamalar ekranda listelenip kullanıcıya bildirilmektedir. Bu işlemler sonucunda kullanıcıdan habersiz çalışan tüm uygulamaların kontrol altına alınması sağlanmıştır. Kullanıcı liste ekranından istediği uygulamayı seçip kapatabilmektedir.

IV. ANALİZ VE TEST

Sistem geliştirilirken kullanılan emülatör sayesinde gerekli test işlemleri uygulama geliştirmenin her aşamasında yapılmıştır. Görünüm dosyalarının oluşturulması vb. işlemlerde başarılı bir şekilde çalışan emülatör ağ dinleme

işlemlerinde tıkanmalara neden olmuştur. Bunun nedeni ise geliştirilen uygulamada kullanılmak istenen kablosuz ağın emülatörde çalışmamasıdır. Bu nedenle gerekli test işlemleri root yetkisine sahip ve kablosuz internet iletişimini destekleyen Android işletim sistemi yüklü bir cihazda gerçekleştirilmiştir. Resim 2.'de geliştirilen uygulamanın kullanıcıyı karşılama ara yüzü gösterilmektedir.

Resim 2. Kullanıcı karşılama ara yüzü

Geliştirilen uygulamadan faydalanarak yapılabilecek işlemler Resim 2.'de gösterilmektedir. IP bilgilerini bulmak için "Ip Bilgileri", gizli ağ bağlantılarını görmek ve ağ dinlemeyi başlatmak için "İşlem Dinleme" ve uygulamayı kapatmak için "Çıkış" butonuna basmak yeterlidir.

Resim 3.'te yer alan ekran görüntüsünde ise geliştirilen uygulamada ağ dinleme sonucu çalıştığı tespit edilen uygulamaların listesi gösterilmektedir.

Resim 3. Ağ dinleme sonucu çalışan uygulamaların listesi

Resim 3.'te görüldüğü gibi ağı dinleme sonucu ağı kullanan uygulamaların listesi kullanıcıya sunulmuştur. Geliştirilen uygulama sayesinde kullanıcı, çalışan uygulamalardan istediğini seçebilmekte ve uygulananın çalışmasına son verebilmektedir. Bu işlem için ekranda listelenen uygulamalardan birini seçip "Durdur" butonuna basması yeterlidir. Yapılan testler sonucu uygulamanın dış IP adresini öğrenme fonksiyonundan, ağı dinleme ve çalışan uygulamaları sonlandırma gibi bütün fonksiyonlarının çalıştığı net olarak görülmüştür. Ayrıca kullanıcının kullanmadığı fakat arka planda ağı kullanan uygulamaların olduğu gözlemlenmiştir.

V. SONUÇ VE ÖNERİLER

Bu çalışmada, ilk etapta hiçbir kötücül yazılım içermeyen daha sonra zararlı hale gelen bir saldırı türü ele alınmıştır. Çalışma kapsamında zaman içerisinde kendine ait kötü amaçlı yazılımı güncelleyen ve bu işlemi gizli ağ bağlantılarıyla gerçekleştiren bir saldırı çeşidi incelenmiştir. Bu tür gizli ağ saldırılarını engellemek için çalışma kapsamında bir mobil uygulama geliştirilmiştir. Uygulama sayesinde farklı durumlarda oluşan ağ trafiği kontrol edilerek kullanıcıdan habersiz çalışan uygulamalar kullanıcıya bildirilmiştir. Yapılan test ve incelemeler sonucu ağ üzerinden yapılan saldırılarda zararlı yazılım bulaşan uygulamaların ağ davranışları ve trafik desenlerinin zararlı yazılım tespitinde oldukça önemli rol oynadığı görülmüştür. Gizli ağ saldırılarında ağ trafiğini inceleyip bu incelemeye göre saldırı önlemi almanın diğer yöntemlere göre daha başarılı olduğu görülmüştür. Ağı kullanan her uygulamanın zararlı olmayacağı için ileriki çalışmalarda bu konu üzerine farklı çalışmalar yapılabilir. Çalışma kapsamında geliştirilen uygulamanın kendi güvenliğinin sağlanması ise farklı bir çalışma olarak ele alınabilecektir.

KAYNAKLAR

[1] Gelenbe, E., Gorbil, G., Tzovaras, D., Liebergeld, S., Garcia, D., Baltatu, M., & Lyberopoulos, G. (2013, June). Security for smart mobile networks: The NEMESYS approach. In Privacy and Security in Mobile Systems (PRISMS), 2013 International Conference on (pp. 1-8). IEEE.

[2] Kabakuş, A. T., Doğru, İ. A., & Çetin, A. Android kötücül yazılım tespit ve koruma sistemleri. Erciyes Üniversitesi Fen Bilimleri Dergisi, (31), 9-16.

[3] Wang, Y., Streff, K., & Raman, S. (2012). Smartphone security challenges.Computer, (12), 52-58.

[4] Wang, Y., & Alshboul, Y. (2015, February). Mobile security testing approaches and challenges. In Mobile and Secure Services (MOBISECSERV), 2015 First Conference on (pp. 1-5). IEEE.

[5] Sudin, S., Tretiakov, A., Ali, R. H. R. M., & Rusli, M. E. (2008, December). Attacks on mobile networks: An overview of new security challenge. In 2008 International Conference on Electronic Design.

[6] Oberheide, J., Veeraraghavan, K., Cooke, E., Flinn, J., & Jahanian, F. (2008, June). Virtualized in-cloud security services for mobile devices. In Proceedings of the First

Workshop on Virtualization in Mobile Computing (pp. 31-35). ACM.

[7] Yu, D., Chen, N., & Tan, C. (2009, March). Design and implementation of mobile security access system (MSAS) based on SSL VPN. In Education Technology and Computer Science, 2009. ETCS'09. First International Workshop on (Vol. 3, pp. 152-155). IEEE.

[8] Ahmad, M. S., Musa, N. E., Nadarajah, R., Hassan, R., & Othman, N. E. (2013, July). Comparison between android and iOS Operating System in terms of security. In Information Technology in Asia (CITA), 2013 8th International Conference on (pp. 1-4). IEEE.

[9] Sun, Y., Wang, Y., & Wang, X. (2014, November). Mobile Security Apps: Loyal Gaurds or Hypocritical Thieves?. In P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2014 Ninth International Conference on (pp. 568-572). IEEE.

[10] Penning, N., Hoffman, M., Nikolai, J., & Wang, Y. (2014, May). Mobile malware security challeges and cloud-based detection. In Collaboration Technologies and Systems (CTS), 2014 International Conference on (pp. 181-188). IEEE.

[11] You, W., Xu, L., & Rao, J. (2013, May). A comparison of TCP and SSL for mobile security. In Sensor Network Security Technology and Privacy Communication System (SNS & PCS), 2013 International Conference on (pp. 206-209). IEEE.

[12] Rashwan, A. M., Taha, A. E. M., & Hassanein, H. S. (2014). Characterizing the Performance of Security Functions in Mobile Computing Systems. Internet of Things Journal, IEEE, 1(5), 399-413.

[13] Willems, E. (2013). Android under attack. Computer Fraud & Security,2013(11), 13-15.

[14] Fuentes, F., & Kar, D. C. (2005). Ethereal vs. Tcpdump: a comparative study on packet sniffing tools for educational purpose. Journal of Computing Sciences in Colleges, 20(4), 169-176.

[15] Bläsing, T., Batyuk, L., Schmidt, A. D., Camtepe, S. A., & Albayrak, S. (2010, October). An android application sandbox system for suspicious software detection. In Malicious and unwanted software (MALWARE), 2010 5th international conference on (pp. 55-62). IEEE.

[16] Thompson, K. (1968). Programming techniques: Regular expression search algorithm. Communications of the ACM, 11(6), 419-422.

Saadin Oyucu, 2012 yılında Gazi Üniversitesi Teknik Eğitim Fakültesi, Bilgisayar Sistemleri Eğitimi Bölümünden mezun oldu. Lisans eğitimi boyunca çeşitli kurslarda eğitici olarak görev yaptı. Lisans eğitimini bitirdikten sonra 2 yıl özel sektörde web yazılım ve arayüz geliştiricisi olarak çalıştı. Bu dönemde birçok önemli kuruluştaki projelerde görev aldı. 2014 yılında ÖYP kapsamında Adıyaman Üniversitesi, Bilgisayar Mühendisliği Ana bilim Dalına Araştırma Görevlisi olarak atandı. Aynı yıl Gazi Üniversitesi Fen Bilimleri Enstitüsüne görevlendirilme ile geldi. 2015 haziran ayında yüksek lisansını tamamladı. Halen görevlendirme ile geldiği, Gazi Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği bölümünde Araştırma Görevlisi olarak çalışmaktadır. Çalışma ve ilgi alanları; M2M, IoT, Robotik, Web Servisler, Yazılım, NoSQL Veritabanları, Mobil sistem Güvenliği, Web uygulama geliştirme, Web arayüz geliştirme.

Hüseyin Polat, 1993 yılında Karadeniz Teknik Üniversitesi, Mühendislik Mimarlık Fakültesi Elektrik Elektronik Mühendisliği Bölümünden mezun oldu. 1995 yılında Gazi Üniversitesi Teknik Eğitim Fakültesi Elektronik Bilgisayar Eğitimi Bölümüne Uzman olarak göreve başladı. Gazi Üniversitesi Fen Bilimleri Enstitüsünde 1998 yılında yüksek lisans ve 2006 yılında doktorasını tamamladı. 2010 yılında Gazi Üniversitesi Teknik Eğitim Fakültesi Elektronik Bilgisayar Eğitimi Bölümüne Yardımcı Doçent olarak atandı. 2011 yılında da Gazi Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümüne Yardımcı Doçent olarak atandı. Halen burada görevine devam etmektedir. Çalışma ve ilgi alanları; Bilgisayar ağları, Bilgisayar donanımı, Endüstriyel otomasyon, Makinalar arası iletişim(M2M) Biomedikal sinyal işleme, Yapay sinir ağları

İ. Alper Doğru, 2004 yılında Atılım Üniversitesi Bilgisayar Mühendisliği Bölümünden mezun oldu. 2007 yılında Gazi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümünde yüksek lisansını 2012 yılında Elektronik-Bilgisayar eğitimi anabilim dalında doktorasını tamamladı. Halen Gazi Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümünde Yrd. Doçent olarak görev yapmaktadır. Çalışma ve ilgi alanları, Mobil şebeke teknolojileri, Mobil tasarsız ağlar, Mobil güvenlik, Network forensics, Mobil forensics, Mobil kötücül yazılım tespiti ve Bulut bilişim sistemleridir.

A BLIND AUTHENTICATION PURPOSE DISCRETE WAVELET WATERMARKING

Ahmet Şenol[1], Ersin Elbaşı[2], Kıvanç Dinçer[1], Hayri Sever[1]

1 Computer Engineering Department, Hacettepe University, 06800 Çankaya Ankara
2 Department of Computer Engineering, Çankaya University, 06790 Etimesgut Ankara
A. Ş. Author (phone: +90-312-4175190/2641; e-mail: asenol@ kho.edu.tr).
E .E. Author (e-mail: eelbasi@cankaya.edu.tr)
K.D. Author (e-mail: kivanc.dincer@hacettepe.edu.tr)
H.S. Author (e-mail: sever@hacettepe.edu.tr) .

Abstract — Image watermarking is used for proving ownership of images, tracking advertisement broadcasts, testing the image's authenticity etc. For watermarking types that aim proving ownership, watermark must resist image operations and watermark must remain in the image after operations. Watermark must be fragile or semi-fragile for authentication watermarking types, where testing the image's being same as original is the main purpose. There are previous authentication-purpose studies in literature but not as much as proving-ownership types. In this study, a blind discrete wavelet transform based authentication purpose watermarking method is proposed. Although the study resembles the methods in literature, it is genuine in its method of embedding in DWT transform and detecting changes in images. The method is capable of detecting changes in images on 4-pixel block base.

Index Terms — Blind, Authentication, Discrete Wavelet Transform, Image Watermarking, Semi-fragile.

I. INTRODUCTION

The Internet made it possible to spread one's copyrighted property without owner's permission. People have to put or share their digital content for various reasons on the Internet by web sites, blog sites, e-mail, social networking etc. By the time, traditional copyright protecting methods such as sticking copyright labels on digital content packs, encrypting data while transferring became insufficient to satisfy needs. Encrypting data has its own problems such as sending keys to target person or content's being defenseless after decryption. Digital watermarking evolved as a new technology to solve digital content ownership proving. The digital content to be protected is called the host or cover data and a special type of data called watermark is embedded in the data itself. The watermark itself is relatively small amount compared to host cover data and becomes part of host data throughout the digital content's life. The watermark can be a visual company logo or a biometric face or sound of owner, a pseudo random number sequence etc. It is preferable for watermark data to have a normal distribution because natural image scenes have normal distribution and embedded watermark must not seem different from original by human eye. The type of watermark for proving ownership is called robust type of watermark and there are many studies on this subject [1]–[7].

For any type of watermarking, fidelity is important. Fidelity is the extent the watermarked image looks like the original one i.e. the similarity between original image and watermarked image. Fidelity is measured by peak signal to noise ratio PSNR given in (1)

$$PSNR = 20 \log_{10}(255/RMSE) \quad (1)$$

where RMSE is square root of mean squared error between original and distorted images as in (2)

$$RMSE = sqrt((\sum_{i,j} (I^*_{ij} - I_{ij})^2) / (NxN)) \quad (2)$$

For some institutions or businesses such as military, aviation, satellite transmission, medical data, it is very important to ensure that downloaded content is same as original. For this purpose, authentication type of watermarking emerged. Authentication type of watermark is fragile or semi-fragile so that watermark disappears or become un-extractable when the watermarked content is modified to some extent. It is preferable for authentication type of watermark that semi-fragile watermark resist for not-ill-purpose image operations such as lossy compression, intensity adjustment, blur filter etc. Semi fragile watermark must deteriorate when a malign operation is applied to digital content such as changing face of a person in the image, putting a non-existent object in the image etc. In authentication type of watermarking, the algorithm should be able to decide whether the digital content in question is genuine without the original content at hand, i.e. algorithm should be blind type of watermarking.

Wolfgang and Delp developed a fragile and semi-fragile watermark for authentication purposes. In fragile watermark, they used the image hash and timestamp in watermarking phase and even a bit change in the watermarked content results in failing to authenticate the image. In their second variable-watermark two-dimensional algorithm (VW2D), the algorithm is able to categorize image as "unaltered", "slightly affected", "definitely altered but still originating from the watermarked image", "completely changed and not originating from watermarked image"[8].

Wong devised an authentication algorithm that is capable of detecting changes in pixel base [9]. He divided the image into blocks and used some blocks for calculating an MD5 hash using image size values M,N, the other block content that will hold the hash data, and private key of watermarking person. On the receiving side, image size, public key of sender, and MD5 holding block's content is used to authenticate the block. The MD5 holding block holds the MD5 hash of the other block in its least significant bits.
Chamlawi, Khan and Idris propose a method that embeds two watermarks for authentication and recovery purposes[10]. They use integer Wavelet transform (IWT) instead of discrete wavelet transform to reduce computational complexity.

II. PROPOSED METHOD

Watermark Embedding Algorithm:

1. Divide Original Image into 8x8 non overlapping parts
2. Take DWT of each block independently
3. Make pairs for 8x8 blocks
   For the second block, for each (i,j) DWT value,
   if the sum (HL_2(i,j)+ LH_2(i,j)+ HH_2(i,j)) <= Threshold
      LL_1(i,j) = floor(LL_1(i, j)/10)*10 + 2;
   else
      LL_1(i,j) = floor(LL_1(i, j)/10)*10 + 7;

4. Take the inverse DWT transform and obtain the watermarked image.

The main idea is divide the image to non-overlapping blocks so that it is possible to localize image file changes. Block pairs are formed so that in one block some values are calculated, in the other block that calculated value is embedded as the watermark. The embedding procedure can be seen more clearly in fig.1. By dividing the LL value by 10 and taking the floor of this value, and then multiplying by 10, last digit of LL value is made zero. Then by adding 2, we make the last digit of LL value 2. For the other case, last digit of LL value is made 7. Two is the middle (also average) value for the interval {0,1,2,3,4}, seven is the middle value for {5,6,7,8,9}. By choosing these two medium values, some degree of resilience to value distortions is provided.

Authenticating Algorithm

1. Divide the Image to be authenticated in 8x8 non overlapping parts
2. Take DWT of each block independently
3. Make pairs for 8x8 blocks
   a. For the second block, for each (i,j) DWT value,
   calculate sum = (HL_2(i,j)+ LH_2(i,j)+ HH_2(i,j))
   if (sum <= Threshold)
      if (LL_1(i,j) >=0 and LL_1(i,j) <=4)
         4 pixel values corresponding to this LLVal is genuine
      else
         4 pixel values corresponding to this LLVal not genuine
      end if
   else
      if (LL_1(i,j) >=5 and LL_1(i,j) <=9)
         4 pixel values corresponding to this LLVal is genuine
      else
         4 pixel values corresponding to this LLVal not genuine
      end if
end if

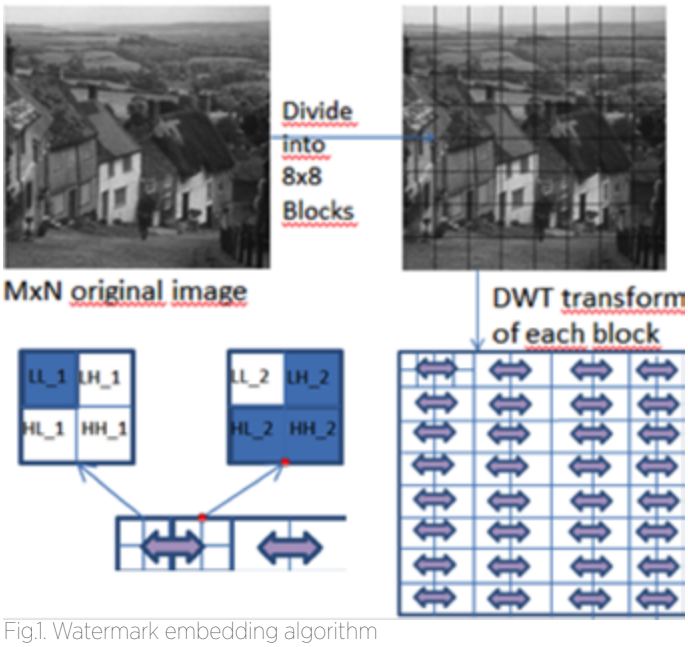


Fig.1. Watermark embedding algorithm

To decide the value to be used as threshold, the sum (LH,HL,HH) values are analyzed. Histogram of those values can be seen in Fig.2.

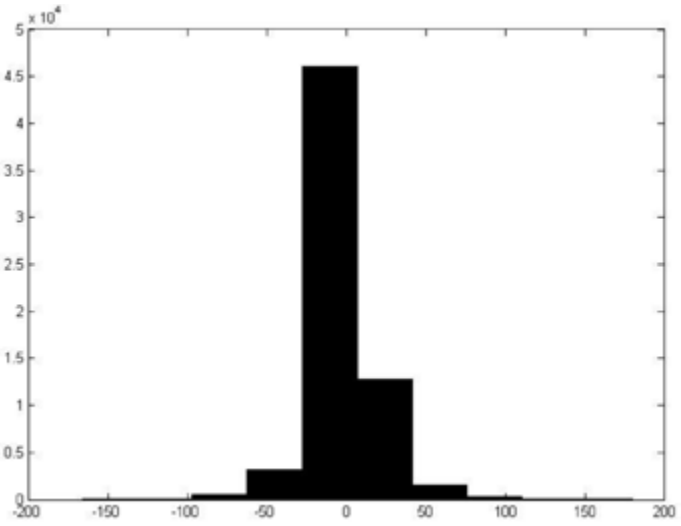


Fig.2. Histogram of sum(HL,LH,HH) values

By looking at the histogram, threshold value -8 is chosen.

III. EXPERIMENTS AND RESULTS

The host image is a grayscale image. But the algorithm can be run on color images by taking the color image to YUV format or by watermarking one or all of the three color bands. The watermarked image is seen in fig.3. Watermarked image has PSNR value of 45.738338 which can be considered a good PSNR value. Algorithm can be run by dividing the image into 512x512(whole image), 64x64, 32x32, 16x16, 8x8 image blocks. One must take into consideration that change detection sensitivity decreases while the block size increases.

The watermarked image is modified in two places where two chimneys disappeared by block copy paste operations. Modified watermarked image is seen in fig.4. Since the aim is to detect changes in the image, altered image's non-professional looking modification is not bothered. It will

not affect the algorithm's success whether the changes are made in a smooth way so that human eye cannot detect the change.

Fig.3.Watermarked image, PSNR : 45.788338

Fig.4 Modified watermarked image. Two chimneys are removed by copy paste.

The image authentication algorithm is run on modified image. The result is seen in fig.5. Two chimney regions where the image was modified is detected and marked by white (255) color value by the algorithm. The detected region is seen as blocky because the algorithm detects modified regions by 4-pixel base. By chance, for some values in modified region, sum(HL,LH,HH) are same as original resulting in a blocky appearance.

The authentication algorithm is tested for the attacks that are used for testing robust type of watermarking applications. The aim here is to test how the algorithm behaves in innocent type lossy compression or scale operations. The watermarked image is subjected to Jpeg compression by %75, %50, %25 image quality, blur filter 3x3, scale-rescale, gamma correction, Gaussian noise, histogram equalization, intensity adjustment.

Fig. 5 Image authentication result for modified image

The image authentication algorithm is run on images that are attacked by common image operations. Results of authentication can be seen in fig.8. For crop operation, authentication can detect the blocks that are original. For the other attacks, since almost all of the blocks are affected by the operation, almost all of the blocks are marked as modified. By intuition, it can be seen and decided that a common innocent operation is applied to the image.

The algorithm is tried on different images with different manipulations as seen on fig 6 and fig 7. The manipulated parts of the image are successfully detected by the algorithm.

IV. CONCLUSION

There are many studies for authentication purpose watermarking. Some of the studies are completely fragile that even a pixel value difference causes authentication to give negative result. Some studies propose semi-fragile type that tolerates some degree of innocent modifications to the image. Most of the studies do not fully describe the details of the algorithm so that the algorithm can be implemented and run.

In this study, a semi fragile blind DWT-based watermarking method is proposed. The algorithm is simple and well-presented so that it can be coded and run easily for testing purposes. The algorithm can detect changes in the detail degree of 4 pixel blocks. Algorithm embeds the watermark in LL band of DWT transform values.

According to experiments done, when a simple operation is applied to whole of the image, the authentication result gives a hint to the observer or it can be decided automatically that the modification is ill-purposed or not.

The PSNR value for watermarked image can be fairly considered as high which is value 45.788338. The proposed study can be used as an alternative to previous authentication methods.

REFERENCES

[1] I. J. Cox, S. Member, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," in Image Processing, IEEE Transactions, 1997, vol. 6, no. 12, pp. 1673–1687.

[2] R. Dugad, K. Ratakonda, and N. Ahuja, "A new wavelet-based scheme for watermarking images," Proc. 1998 Int. Conf. Image Process. ICIP98 (Cat. No.98CB36269), vol. 2, pp. 1–5, 1998.

[3] A. M. ; G. E. Eskicioglu, "Robust DWT-SVD Domain Image Watermarking : Embedding Data in All Frequencies," 2004.

[4] E. Elbasi, A. M. Eskicioglu, and I. Science, "A DWT-based robust semi-blind image watermarking algorithm using two bands," vol. 6072, pp. 1–11, 2006.

[5] O. Jane and E. Elbasi, "A new approach of nonblind watermarking methods based on DWT and SVD via LU decomposition," Turkish J. Electr. Eng. Comput. Sci. doi10.3906/elk-1212-75, pp. 1–13, 2012.

[6] O. Jane, H. Gökhan, and E. Elbaşı, "A Secure and Robust Watermarking Algorithm Based on the Combination of DWT , SVD , and LU Decomposition with Arnold ' s Cat Map Approach," no. 3, pp. 306–310, 2014.

[7] H.-C. Huang and W.-C. Fang, "Metadata-based image watermarking for copyright protection," Simul. Model. Pract. Theory, vol. 18, no. 4, pp. 436–445, Apr. 2010.

[8] R. B. Wolfgang, E. J. Delp, R. B. Wolfgang, and E. J. Delp, "Fragile Watermarking Using the VW2D Watermark," Proc. SPIE/IS&T Int. Conf. Secur. Watermarking Multimed. Contents, vol. 3657, pp. 204–213, 1999.

[9] P. W. Wong and W. Road, "A Public Key Watermark for Image Verification and Authentication," Image Process. 1998. ICIP 98. Proceedings. 1998 Int. Conf., vol. 1, pp. 455–459, 1998.

[10] R. Chamlawi, A. Khan, A. Idris, and Z. Munir, "A Secure Semi-Fragile Watermarking Scheme for Authentication and Recovery of Images based on Wavelet Transform," vol. 2, no. 2, pp. 727–731, 2008.

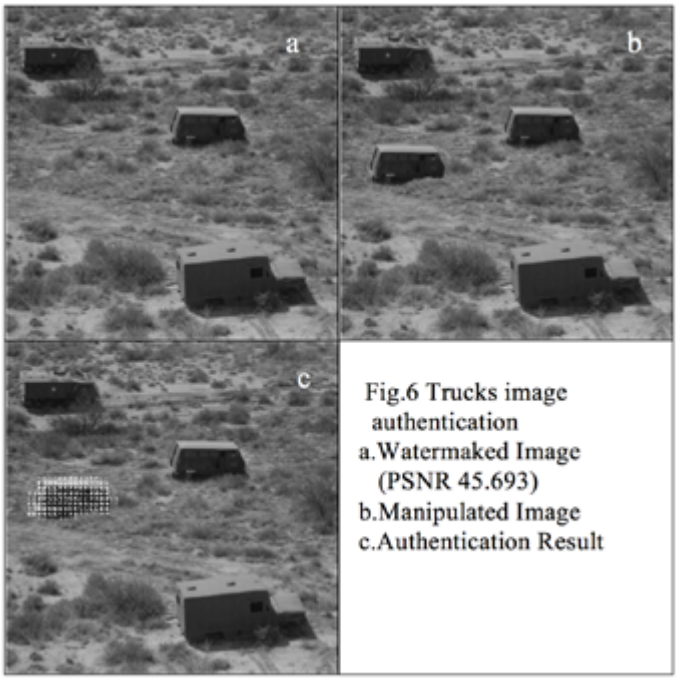


Fig.6 Trucks image authentication
a.Watermaked Image (PSNR 45.693)
b.Manipulated Image
c.Authentication Result

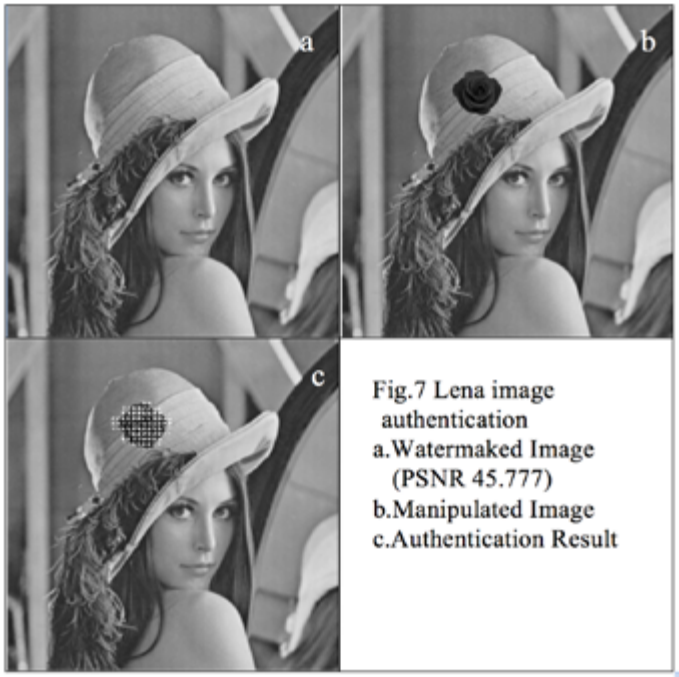


Fig.7 Lena image authentication
a.Watermaked Image (PSNR 45.777)
b.Manipulated Image
c.Authentication Result

Fig. 8. Image authentication applied to watermarked and modified images by various image operations

BRAILLE ALFABESİ TABANLI OLASILIKSAL GÖRSEL SIR PAYLAŞIMI METODU

T.Tuncer ve E. Avcı
Türker Tuncer, Fırat Ünv. Teknoloji Fak. Adli Bilişim Müh. Böl. 23119 Elazığ/TÜRKİYE (turkertuncer@firat.edu.tr)
Engin Avcı, Fırat Ünv. Teknoloji Fak. Yazılım Müh. Böl. 23119 Elazığ/TÜRKİYE (enginavci23@hotmail.com)

Özet — Naor ve Shamir imgelerin ve şifreleme anahtarlarının gizliliğinin korumak için (k,n) görsel sır paylaşımı (GSP) metodunu önermiştir. Bu metot kullanılarak, gizli veri karmaşık hesaplamalar yapılmaksızın sır parçalarına ayrılabilmektedir. GSP metotları güvenilirliği sağlar ancak gürültü benzeri imgeler saldırganların dikkatini çekmektedir. Bu makalede alt pikseller, braille alfabesindeki harflerden oluşmaktadır yani sır parçaları anlamlı parçalardan oluşmaktadır. Bu çalışmada, olasılıksal bir yaklaşım kullanılarak yeni bir GSP önerilmiştir. Yeniden yapılandırma aşamasın özel veya (XOR) operatörü kullanılmıştır. Ayrıca, önerilen braille tabanlı GSP (BGSP) kullanılarak veri gizleme uygulaması gerçekleştirilmiştir. Böylece, sır parçaları saldırganların dikkatini çekmeden alıcı tarafa iletilebilecektir.

Anahtar Kelimeler — Braille tabanlı görsel sır paylaşımı, Olasılıksal görsel sır paylaşımı, Veri gizleme, Damgalama, Bilgi güvenliği, İmge işleme.

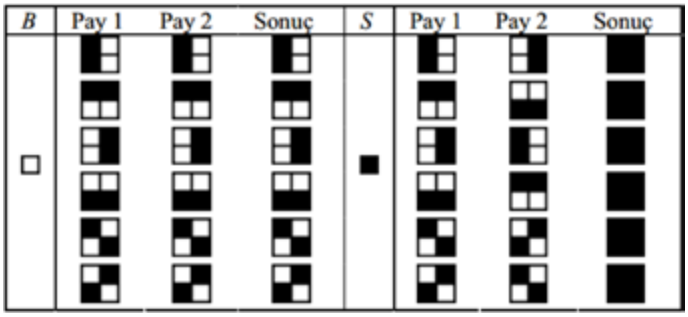
Abstract — (k,n) visual cryptography scheme is firstly proposed by Naor and Shamir to protected image contents and encryption keys. Secret data can be divided into secret shares without complex calculation by using this method. Visual cryptography methods are provided security but a lot of attack is developed for noise- like image by attacker. In this paper, subpixels are coded with braille coding. We created meaningful secret shares by using braille. In this study, we presented a new probabilistic braille based visual cryptography algorithm with XOR operator. The proposed method is used XOR for reconstruction. Also, we used data hiding for camouflage. Thus, the secret shares are sent to receiver without attracting attention of attackers.

Keywords — Braille based visual cryptography, Probabilistic visual cryptography, Data hiding, Watermarking, Information security.

I. GİRİŞ

Bulut teknolojisinin kullanımının artmasıyla birlikte, bilgi güvenliğinin de önemi artmıştır. Çünkü bulut, kullanıcılarının erişimine açık bir platformdur. Bulutta bulunan bilgilerinin güvenliğini ve gizliliğini sağlayabilmek için bilgi güvenliği yöntemlerinin kullanılması gerekmektedir. Bu güvenlik önlemlerinin başında ise şifreleme ve veri gizleme gelmektedir. Şifreleme, bir verinin içeriğini değiştirmeye yönelik kullanılırken; veri gizleme örtü nesnesinin içeriğini değiştirmez sadece gizli veriyi bir örtü nesnesine gizler. Veri gizlemedeki en temel amaç ise, gizli verinin sezilememesidir [1-3]. Kısacası, şifreleme verinin içeriğini korumayı amaçlarken, veri gizleme verinin sezilememesini amaçlamaktadır. Verilerin güvenilir olarak paylaşımı ve saklanması için sır paylaşımı algoritmaları önerilmektedir [4].

Sır paylaşımı algoritmaları, ilk olarak 1979 yılında Blakley ve Shamir tarafından önerilmiştir [5, 6]. Bu yöntemlerin temel amacı şifreleme anahtarı korumak ve güvenilir bir dağıtıcı ile sır parçalarını dağıtmaktır. Sır parçaları bir araya gelince anahtarı oluşturacaktır. GSP şemaları ise ilk olarak 1994 yılında Naor ve Shamir tarafından önerilmiştir [7]. Bu algoritmayla, gizli mesaj belirlenen kurallara göre sır parçalarına ayrılmaktadır. Gizli veriyi yeniden elde etmek için karmaşık matematiksel işlemlere gerek yoktur. Sır parçalarının üst üste gelmesiyle gizli mesaj elde edilebilmektedir. Shamir' in görsel sır paylaşımı algoritmasının kodlama tablosu Tablo 1' de verilmiştir.



Tablo 1. Shamir' in GSP şemasında piksellerin kodlanması [4].

Tablo 1' de de görüldüğü gibi, Shamir'in GSP şemasında imgenin yeniden yapılandırılması için mantıksal VEYA operatörü kullanılmıştır. Bu metodun yanı sıra olasılıksal GSP (OGSP) şemaları da mevcuttur. Wang' ın şeması XOR ve VE operatörü kullanan ve en yaygın kullanılan OGSP'lerden biridir [8]. Bu şemada sır parçalarının bir kısmı rastgele üretilmektedir. Diğer sır parçaları ise istenilen sonuca göre üretilmektedir.

İmge kimliklendirmek ve gizli verinin güvenliğini arttırmak için veri gizleme ve görsel şemalarının bir arada kullanılması önerilmiştir. Ayrıca GSP ve veri gizlemenin birlikte kullanıldığı çalışmalar şu şekilde sıralanmıştır. Lee vd. PNG imgeleri kimliklendirmek için Shamir'in (k,n) görsel sır paylaşımı şemasını kullanmışlardır [9]. Yuan sır paylaşımı algoritmalarını kullanarak çoklu örtü imgesi tabanlı uyarlamalı staganografi algoritmasını önermiştir. Önerilen algoritma Shamir'in sır paylaşımı algoritmasını kullanmaktadır ve veri gizleme fonksiyonu olarak ±1 operatörü kullanılmaktadır. Bu algoritmayla, yüksek görsel kalite elde edilmiştir [10].

Ayrıca, literatürde harf tabanlı GSP şemaları da bulunmaktadır. Takizawa vd. Japon harflerini kullanan iki adet sır paylaşımı metodu önermiştir. İlk metotta bir veritabanı oluşturulmuştur. Oluşturulan veritabanı kullanılarak harflerin morfolojik analizi gerçekleştirilmiştir. Belirlenen harfler döndürülerek, sır parçaları elde edilmiştir. Takizawa vd. Ikinci yaklaşımında ise, harfler kullanılarak anlamlı cümleler elde edilmiştir. Anlamlı cümleler sır parçaları olarak kabul edilmiştir. Birden fazla anlamlı cümlenin bir araya gelmesiyle mesaj elde edilmiştir [11].

Lin vd. çince, korece, japonca ve latince harflerini tabanlı bir görsel sır paylaşımı metodu önermiştir. Bu metot temel olarak Shamir'in (k,n) görsel sır paylaşımını algoritmasını temel almaktadır. Alt pikseller, harflerden oluşmaktadır [12]. Wang vd. görsel şifreleme için Braille adlı bir makale yayınlamıştır ve bu makalede RGB imgelerin kimliklendirilmesiyle ilgili bir çalışma yapılmıştır [13].

Bu makalede Braille alfabesinde harflere karşılık gelen kodlar analiz edilmiştir ve XOR operatörü kullanılarak yeni bir GSP şeması oluşturulması öngörülmüştür.

Bu makalenin organizasyonu aşağıdaki gibi verilmiştir. İkinci bölümde motivasyon ve tasarım, üçüncü bölümde Braille alfabesi, dördüncü bölümde önerilen algoritma, beşinci bölümde deneysel sonuçlar ve altıncı bölümde ise sonuç ve önerilerden bahsedilmiştir.

II. MOTİVASYON VE TASARIM

Bu makalede çok seviyeli bir güvenliği metodu oluşturularak, gizli verinin güvenliği sağlanmıştır. Braille alfabesi kullanılarak anlamlı sır parçalarından oluşturulmuş yeni bir GSP şeması oluşturulmuştur. Ayrıca oluşturulan sır parçaları örtü nesnesinin içerisine gizlenmiştir, böylece sır parçaları için güvenilir veri iletim hattı oluşturulmuştur. Saldırgan steganaliz yöntemlerini kullanarak sır parçalarını elde etse dahi, sır parçaları Braille alfabesinde bulunan harflerden oluştuğu için, saldırgan sır parçasında bir şifre olduğu sanıp o şifreyi çözmeye çalışacaktır.

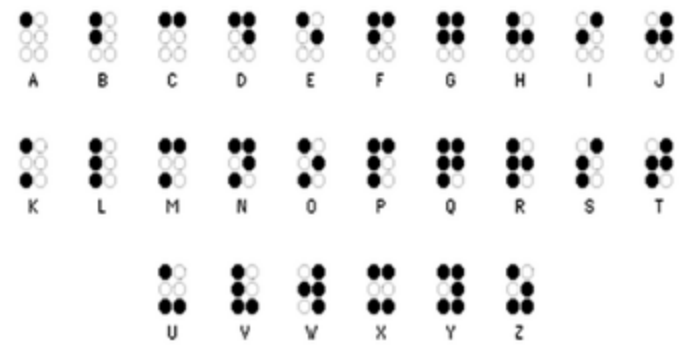
Kısacası bu makalede, saldırganın dikkatini çekmeden sır parçalarını alıcı tarafa gönderilmesi hedeflenmiştir. Çünkü gürültü benzeri sır parçaları saldırganların dikkatini çekmekte ve bu sır parçalarını elde eden saldırganlar, çeşitli saldırılar ve hileler düzenleyerek gizli veriyi değiştirebilmektedir. Eğer sır parçaları elde edilirse, Braille kodlar sayesinde saldırganın dikkati başka bir yöne doğru çekilecektir. Saldırgan ilk etapta Braille kodlarını anlamlandırmaya çalışacaktır.

Önerilen metodun motivasyonu, anlamlı alt parçalardan oluşan OGSP metodu tasarlamak ve bu metodu veri gizleme algoritmalarıyla birlikte kullanıp, yüksek seviyeli veri güvenliğini sağlamaktır.

III. BRAİLLE ALFABESİ

Görme engelli kişilerin kullanması için, 1829 yılında Louis Braille tarafından önerilmiştir. Louis Braille' in keşfettiği bu alfabe literatürde Braille alfabesi olarak adlandırılmaktadır. Braille alfabesi 6 adet noktadan oluşmaktadır ve bu noktalar 3 x 2 boyutundaki bir matrise yerleştirilmiştir. Alfabede yer alan işaretlerin tamamı bu 6 noktanın pozisyonuna göre oluşturulmuştur [14].

Harflerin Braille alfabesine göre kodlanması şekil 1' de verilmiştir.



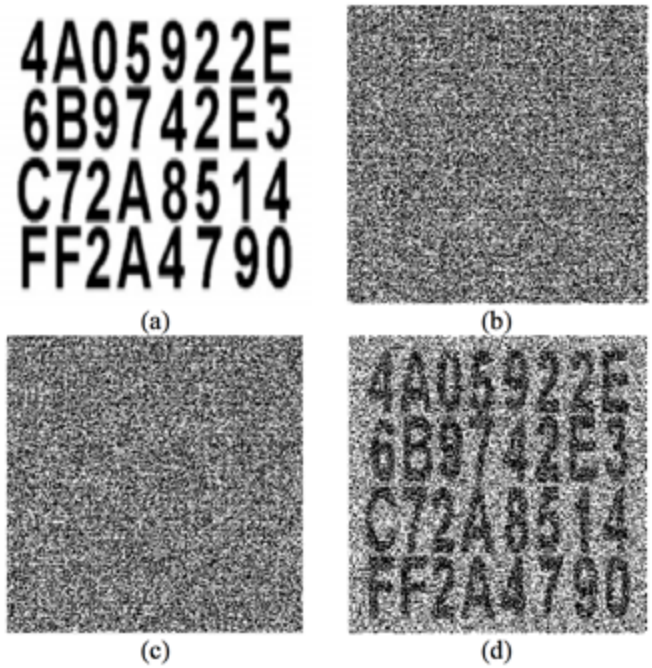
Şekil 1. Harflere Ait Braille kodları [15].

IV. ÖNERİLEN METOT

Bu makalede sır parçaları olarak, şekil 1' de gösterilen Braille harfleri kullanılmıştır. Önerilen yöntem Shamir' in (k,n) GSP şeması ve Wang' ın OGSP şemasından esinlenerek ileri sürülmüştür [7,8]. Olasılıksal BGSP oluşturulduğu için rastgele sayı üreteçleri kullanılmıştır. Sır paylaşımı gerçekleştirmek için XOR operatörü kullanılmıştır. Bu operatörün kullanılmasının temel sebebi ise 0 ve 1' in oluşmasından 0.5' e en yakın olasılığın oluşmasıdır. (2,2) BGSP şemasında, 26 adet harf kullanıldığı için toplam olasılık sayısı 262=676'dır. Bir bloğun siyah piksele eşit olması için en az 4 adet siyah pikselin olması gerekmektedir. Diğer durumlarda o blok beyaz (1) olarak ifade edilecektir. Bu koşullar altında 362 çift Braille koda XOR işlemi uygulanması sonucu 1, 314 çift Braille kodun XOR işlemi uygulanması sonucu 0 elde edilecektir. Bu makalede önerilen olasılıksal yöntemin, Yang' ın [16] sunduğu OGSP' ye göre en temel farkı alt piksellerin anlamlı olmasıdır. Yang' ın şemasında imgenin kontrastı söz konusuyken, önerilen metotta harflerin gelme olasılığı hesaplanmalıdır ve bu harflerin biraraya geldiğinde 0 ve 1' i elde etme olasılıklarının hesaplanması gerekmektedir. Ayrıca önerilen algoritmada kullanılan kural tablosu kontrast olasılığının gerçekleşmesi için de modifiye edilebilmektedir. Sözde rastgele sayı üreteçleri kullanılarak harflerin gelme olasılıkları uniform olarak ayarlanabilir. Kullanılan sözde rastgele sayı üreteçlerinin büyük bir kısmı üniform özellik gösterdiği için, bu yöntemde kullanılan rastgele sayı üretecinin türünün pek bir önemi yoktur. Uniform dağılım gösteren herhangi bir rastgele sayı üreteci kullanılabilmektedir. P(0)=314/676=0.4645 ve P(1)=362/676=0.5355 olacaktır. Olasılıkların 0.5' e yakın olmasından dolayı XOR operatörü kullanılmıştır. Önerilen olasılıksal BGSP' nin algoritması aşağıdaki gibidir. Ayrıca sır parçalarını gizlemek için veri gizleme algoritmalarından faydalanılmıştır.

Adım 1: 0 ve 1 kombinasyonlarını iki ayrı listeye kaydet.
Adım 2: İkili imgeyi gir.
Adım 3: İkili imgenin piksel değeri 0 ise sıfırlar listesinden rastgele Braille kodları seç.
Adım 4: İkili imgenin piksel değeri 1 ise birler listesinden rastgele Braille kodları seç.
Adım 5: Braille kodlarını sır parçası olan imgelere yerleştir.
Adım 6: İkili imgenin boyutu kadar adım 3-5' i tekrarla.
Adım 7: Elde edilen sır parçalarını örtü imgelerine veri gizleme fonksiyonunu kullanarak gizle.

Şekil 2' de (2,2) BGSP kullanılarak yapılan sır paylaşımı işlemi gösterilmiştir.



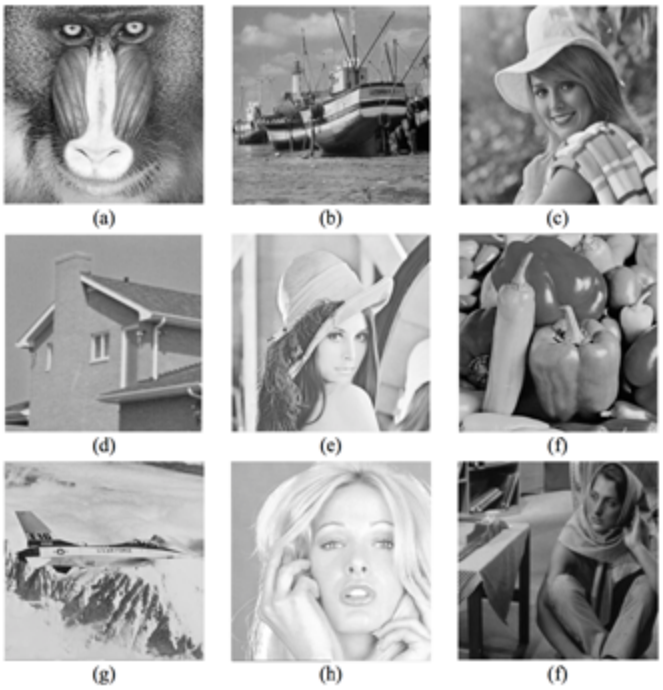
Şekil 2. (2,2) BGSP (a) Gizli veri (b) 1. Sır parçası (c) 2. Sır parçası (d) Yeniden elde edilmiş imge

İmgeyi yeniden elde etmek için, XOR operatörü kullanılmaktadır.

Horng vd. [17]' nin sunduğu makalede, GSP şemalarında meydana gelebilecek sahtekârlıktan bahsedilmiştir. Eğer herhangi bir sır parçası saldırganın eline geçerse ve saldırgan logonun ne olduğu bilirse, elde ettiği sır parçasını modifiye ederek farklı bir logonun oluşturulmasını sağlayabilmektedir. Bu tip sahtekârlıklardan korunabilmek için logonun saldırgandan gizlenmesi gerekmektedir. Horng vd. makalesinde kimlik doğrulama sistemi geliştirilerek bu tip sahtekârlıkların önüne geçilmeye çalışılmıştır. Bu makalede, sır parçalarını, bu tip saldırılardan koruyabilmek için veri gizleme uygulaması gerçekleştirilmiştir. Sır parçaları örtü nesnelerin içerisine gizlenerek, sır parçalarının sezilememesi ve elde edilememesi sağlanmıştır ve şemayı daha güçlü bir hale getirmek için veri gizleme uygulamaları kullanılmıştır. Sır parçalarını kamufle etmek için kullanılan veri gizleme algoritması ise 2LSBs (Least significant bits – En anlamsız bite gömme) algoritmasıdır. Bu algoritma kullanılarak, sır parçaları örtü nesnesinin en anlamsız iki bitine gömülmüştür.

V. DENEYSEL SONUÇLAR

Önerilen BSGP metoduyla sır parçalarına ayrılmış verileri veri gizleme uygulamasını test edebilmek için SIPI [18] imge veritabanı kullanılmıştır. Kamuflaj safhasında veri gizleme algoritmaları kullanılmıştır ve BGSP tabanlı veri gizleme algoritmasının görsel kalitesi test edilmiştir. Kullanılan imgeler 512 x 512 boyutundadır ve şekil 3' te gösterilmiştir.



Şekil 3' te gösterilen test imgelerinin görsel kalitesini test edebilmek için PSNR (peak signal-to noise ratio) ve MSE (mean square error) metrikleri kullanılmıştır. MSE ve PSNR' nin formülleri formül 1 ve 2'de verilmiştir.

$$MSE = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(CI_{i,j} - SI_{i,j})^2 \quad (1)$$

$$PSNR = 10\log\frac{Max(CI_{i,j}^2)}{MSE} \quad (2)$$

Şekil 3' te gösterilen imgeler 512 x 512 boyutundadır ve bu imgelere 524,288 bit veri gömülmüştür. Elde edilen PSNR sonuçları Tablo 2' de verilmiştir.

Örtü İmgesi	PSNR (dB)	
	Sır 1	Sır 2
Baboon	44.65	44.93
Boat	44.99	44.37
Elaine	45.52	45.11
House	45.36	45.54
Lena	44.01	44.75
Peppers	44.88	44.16
Airplane	43.97	44.61
Tiffany	45.32	45.88
Barbara	45.36	44.93

VI. SONUÇ

Bu makalede Braille alfabesinde bulunana harfler kullanılarak, yeni bir anlamlı sır paylaşımı metodu önerilmiştir. Önerilen metot kullanılarak gizli veri sır parçalarına ayrılmış ve her bir sır parçası bir örtü nesnesinin içerisine gizlenerek veri gizleme uygulaması gerçekleştirilmiştir. Önerilen BGSP algoritması olasılıksal metot ve XOR operatörünü kullanmıştır. Bu metotta XOR kullanılarak olasılıklar 0.5' e yaklaştırılmıştır. Rastgele sayı üreteci kullanılarak, Braille

kodların sır parçası üzerinde uniform dağılımı sağlanmıştır. Kamuflaj aşamasında ise veri gizleme algoritmalarından faydalanılmış ve başarılı sonuçlar elde edilmiştir.

Gelecekteki çalışmalarda, biyometrik bilgi güvenliğini sağlayabilmek için ve yüksek görsel kaliteye sahip veri gizleme tabanlı imge kimlik doğrulama algoritmaları oluşturabilmek için önerilen algoritma kullanılacaktır.

KAYNAKLAR

[1] C. Deng, X. Gao, X. Li, D. Tao, A local Tchebichef moments-based robust image watermarking, Signal Process. 89 (8) (2009) 1531–1539.

[2] J. Fridrich, D. Soukal, Matrix embedding for large payloads, IEEE Trans. Inf. Forensics Secur. 1 (3) (2006) 390–395.

[3] X. Gao, C. Deng, X. Li, D. Tao, Geometric distortion insensitive image watermarking in affine covariant regions, IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. 40 (3) (2010) 278–286.

[4] V.V. Nabiyev, M. Ulutas, G. Ulutas, Doğruluk oranı iyileştirilmiş (2,n) olasılıklı görsel sır paylaşımı şeması, 3. Information Security & Cryptology Conference with International Participation, 2008.

[5] G.R. Blakley, Safeguarding Cryptographic Keys, Proceedings of the National Computer Conference, American Federation of Information Processing Societies Procedings, New York, USA, pp. 313-317, June 1979.

[6] A. Shamir, How to Share a Secret, Communications of ACM, vol. 22, no 11, pp. 612-613, 1979.

[7] M. Naor, A. Shamir, Visual cryptography, in: A. DeSantis (Ed.), Advances in Cryptology – EUROCRYPT'94, Lecture Notes in Computer Science, Perugia, Italy, vol. 950, 1994, pp. 1–12.

[8] D. Wang, L. Zhang, N. Ma, and X. Li, Two secret sharing schemes based on Boolean operations. Pattern Recognition 40(10), 2776–2785, 2006.

[9] C. Lee, W. Tsai, A data hiding method based on information sharing via PNG images for applications of color image authentication and metadata embedding, Signal Processing, pp. 2010-2025, (93), 2013.

[10] H. Yuan, Secret sharing with multi-cover adaptive steganography, Information Sciences, pp. 197-212, (254), 2014.

[11] O. Takizawa, A. Yamamura, A proposal of secret sharing using natural language text, in: IPSJ Computer Security Symposium, 2001, pp. 343–348.

[12] H. Lin, C. Yang, C. Laih, H. Lin, Natural language based visual cryptography scheme, J. Vis. Commun. Image R. , pp. 318-331, (24), 2013.

[13] G. Wang, F. Liu, W. Q. Yan, Braille for visual cryptography, IEEE International Symposium on Multimedia, 2014.

[14] MEB, Özel eğitim okulları için Braille kabartma yazı kılavuzu, MEB devlet kitapları, pp. 4, 1991. (URL: http://orgm.meb.gov.tr/alt_sayfalar/yayimlar/ozelegitim/blair/blair.pdf)

[15] http://sanlitarihim.blogcu.com/braille-alfabesini-dunyada-ilk-kez-osmanli-kullandi/6593257 (Son Erişim Tarihi: 16/08/2015)

[16] C. -N. Yang, New visual secret sharing schemes using Probabilistic method, Pattern Recognition Letters, 25, pp. 481-494, 2004

[17] G. Horng, T. Chen, D., -S. Tsai, Cheating in visual cryptography, Designs, Codes and Cryptography, 25, pp. 219-236, 2006.

[18] SIPI Image Database, http://sipi.usc.edu/database/ (Access Date: 26/08/2015)

TÜRKİYE'DE E-DÖNÜŞÜM HİZMETLERİNDE KİŞİSEL BİLGİLERİN GİZLİLİĞİNİN KORUNMASI

Gülsüm KAPANOĞLU, Yiğit HACIEFENDİOĞLU, Mertcan ÜNAL ve H. İbrahim BÜLBÜL

Gazi Üniversitesi Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, ANKARA
gulsumkapanoglu@gmail.com, yigith1@gmail.com, mertronot@gmail.com, ibrahmhalil@gmail.com

Özet — Sürekli gelişim içerisinde olan bilgisayar ve internet teknolojileri; ihtiyacımız olan bilgiye daha hızlı ve daha kolay bir şekilde ulaşmamıza olanak sağlamaktadır. Bilişim teknolojilerini kullanan kullanıcı sayısının hızlı bir artış göstermesi ve veri kaynaklarının çeşitlenmesi, veri hacminin artmasına neden olmakta ve büyük bir potansiyel doğurmaktadır. Bu da, kişisel verilerin de içerisinde bulunduğu büyük verinin güvenlik açıklarını ve güvenliği sağlama yöntemlerini araştırma ihtiyacını beraberinde getirmiştir. Ülkemizde de e-Dönüşüm başlığı altında birçok kurum ve kuruluş, kişisel verileri dijital ortama taşımaya başlamıştır. Böylece kişisel verilerin güvenliğine verilen önem de bir kat daha artmıştır. Bu çalışmada da; e-Dönüşüm kavramının ne denli önemli olduğunu açıklamak, verilen hizmetlerin kişisel bilgilerin güvenliği açısından karşılaşabilecek tehditleri ve bu tehditlere karşı geliştirilebilecek önlemler hakkında güncel tespitleri ortaya koymak amaçlanmıştır.

Anahtar kelimeler — e-dönüşüm, gizlilik, kritik altyapı, kişisel bilgiler, tehditler

Abstract — Continuously developing computer and internet technologies provide us means and tools to reach information that we need faster and more effectively. The increasing number of information technology users and the massive variety of big data sources constitutes a huge potential risk in terms of personal privacy, and this brings about the need to investigate and find the ways of keeping the gigantic mass of data secure and protected. In Turkey, many private/government constitutions and companies started to transfer individual data and public services from paper documents to the digital environment. As a result of this e-transformation, the personal information security gained more and more significance. This study aims to explain the importance of e-transportation, threats related to personal information security in e-transformation services and to offer solutions for minimizing and eliminating these threats.

Keywords — e-transformation, privacy, critical infrastructure, threats

I. GİRİŞ

Bilginin, nitelik ve nicelik bakımından değerinin her geçen gün artmasıyla birlikte; içerisinde bulunduğumuz çağ bilgi çağı olarak adlandırılmaktadır. Sürekli bir gelişim içerisinde olan bilgisayar ve internet teknolojileri de, ihtiyacımız olan bilgiye daha hızlı ve daha kolay bir şekilde ulaşmamıza imkân sağlamaktadır. Türkiye İstatistik Kurumu'nun 2015 yılı nisan ayı verilerine göre; Türkiye'de 16-74 yaş grubundaki bireylerin bilgisayar kullanım oranı %54,8 iken internet kullanım oranı %55,9'dur. Aynı zamanda genişbant internet erişim imkânına sahip olan hanelerin oranı %67,8 olarak açıklanmıştır. Cep telefonu veya akıllı telefona sahip olma oranı da %96,8 olarak araştırma sonuçları içerisinde kendine yer bulmuştur [1].

Hem kullanıcı sayısının artması hem de veri kaynaklarının çeşitlenmesi, veri hacminin artmasına neden olmakta ve büyük bir potansiyel doğurmaktadır. Teknolojinin uygunsuz kullanımı ve bireylerin bilgi güvenliği tehditlerine yönelik farkındalık seviyelerinin düşük olması, telafisi güç bilgi güvenliği risklerini ve siber suçlarını ortaya çıkarmaya başlamıştır [2]. Bu da, kişisel verilerin de içerisinde bulunduğu büyük verinin güvenlik açıklarını araştırmayı ve güvenliği sağlama yöntemlerini beraberinde getirmiştir.

Kişisel verilerin hukuki bir çerçevede güvenliğinin sağlanması için, bu verilerin nelerden oluştuğunun bilinmesi gerekmektedir. Keser vd. göre kişisel veriler şu şekilde tanımlanmaktadır [3]:

- Kişiler tarafından oluşturulan bloglar, yorumlar, fotoğraflar, videolar,
- Kişinin internet aktivitesine dair bilgiler, yaptığı aramalar,
- Sosyal platformlardaki veriler, kişinin arkadaşları ve çevreleri,
- Kişinin konum bilgisi,
- Kişinin demografik bilgileri,
- Resmi niteliğe sahip ve kişiyi tanımlamak için kullanılabilecek finansal veriler, hesap bilgileri, sağlık kayıtları ve emniyet kayıtları.

Bilgiye erişim yöntemlerinin değişmesi ve gelişmesi ile birlikte e-Dönüşüm başlığı altında kişisel veriler, birçok kurum ve kuruluşça internet ortamına aktarılmaya başlanmıştır. MERNIS, e-okul, e-devlet ve e-ticaret gibi dijital ortamlar, kullanıcılarının; sayısız alanda bilgiye ulaşmalarına olanak sağlamaktadır. e-Dönüşüm uygulamalarının getirdiği kolaylıklar aynı zamanda kişisel verilerin güvenliği ve gizliliği konusunun da önemini bir kat artırmıştır. Bu tip sistemlere erişim için kullanılan kullanıcı adı ve şifre gibi araçların güvenliği bireylere düşerken, sistemlerin yetkisiz kullanıcılarca erişimine engel olmak da e-Dönüşüm uygulamalarının yöneticilerine düşmektedir.

ISO 27001 Bilgi Güvenliği Yönetimi Sistemi gereksinimlerini tanımlayan tek uluslararası denetlenebilir standarttır. Yeterli ve orantılı güvenlik denetimleri seçilmesini sağlamak için tasarlanmıştır. e-Dönüşüm uygulamaları da bu standardın gerekliliklerini yerine getirmek için çalışmalar yapmak durumundandır. ISO 27001'in getirdiği standartları şu şekilde sıralayabiliriz [4]:

- Gizlilik: Önemli ve hassas bilgilerin istenmeyen biçimde yetkisiz kişilerin eline geçmesi önlenmelidir ve sadece erişim yetkisi verilmiş kişilerce erişilebilir olduğu garanti altına alınmalıdır.
- Bütünlük: Bilginin sahibi dışındaki kişilerce değiştirilmesinin ve silinmesinin önlenmesi gerekmektedir.
- Kullanılabilirlik: Bilgi veya bilgi sistemleri sürekli kullanıma hazır ve kesintisiz çalışır durumda olmalıdır.
- Doğrulama: Kullanıcı kimliğinin doğrulanması gerekmektedir.

En son standartlarda, en yeni teknolojiyle donatılan bir sistemde dahi yarın için farklı tehditler vardır. Hiçbir zaman için bir sistem yüzde yüz etkin başarı sağlar demek mümkün değildir. Yarının ihtiyaçlarını karşılayabilmek için devamlı gelişen bir süreç içerisinde hareket edilmesi gerekmektedir[5]. Bu nedenle e-Dönüşüm hizmetlerini veren kurum ve kuruluşların, gelişen teknolojiyi ve güvenlik yöntemlerini dinamik bir şekilde takip etmesi ve kendi sistemine entegre etmesi gerekmektedir.

Bu çalışmanın amacı; e-Dönüşüm kavramının ne denli önemli olduğunu açıklamak, verilen hizmetlerin kişisel bilgilerin güvenliği açısından karşılaşabilecek tehditleri ve bu tehditlere karşı geliştirilebilecek önlemler hakkında güncel tespitleri ortaya koymaktır.

II. E-DÖNÜŞÜM HİZMETLERİ

E-dönüşüm, insanların yaşamını her yönüyle etkileyen ve kolaylaştıran, devletin vatandaşlarına sunduğu hizmeti farklı boyutlara taşıyan bir süreçtir. Bilgi toplumunun olağan getirisi olan bilgi üretimi, toplumların yapısını yeniden düzenlemekte ve sistemler oluşturmaktadır. Üretilen bilgilerin yönetimi e-hizmetler sistemleri ile yeni boyutlara taşınmaktadır. Gelişen bilişim teknolojileri giderek kamu yönetimini de kuşatmış ve kamu alanlarının dönüşümüne de destek olmuştur. Kamu yönetimi alanı, kurumların e-devlet sistemleri ve e-dönüşüm projelerinde şekil değiştirmektedir [8].

Bireylerin borç ödemek veya alışveriş yapmak amacıyla kullandıkları internetten, başvuru ve arama işlemlerinin tamamlanmasının ardından sonuç alınabilmesi e-dönüşüm sistemi ile sağlanabilmekte ve e-dönüşüm kavramı, kazandığı boyutlarıyla birlikte görülmektedir [9]. Devletin etkili ve şeffaf biçimde çalışmasını sağlayacak olan yönetsel reform çalışmalarının hız kazanması ve az kaynakla çok iş yapma gereksinimi e-dönüşüm hizmetlerinin ortaya çıkış sürecini hızlandırmıştır. e-dönüşüm sistemiyle beraber kamu kuruluşları arasında yeni ortaklıklar ve işbirliklerine ve yeni iş yapma yol ve yöntemlerine ihtiyaç artmaya başlamıştır.

Bilişim ve iletişim teknolojilerinin etkin kullanımıyla kurumların ürün ve hizmetlerinin değiştirilmesi süreci e-dönüşümle olmaktadır. Bu sebepten ülkeler e-dönüşüm süreci için interneti yaygınlaştırmak ve bilgisayar okuryazarlığını arttırmak için çeşitli politikalar geliştirmektedirler [7].

E-dönüşüm süreci içerisinde iş, eğitim, sağlık, temel hizmetler ve ulaşımı barındırmaktadır. Gelişen internet ortamıyla, toplumlar bu hizmetlere ulaşımı sağlamıştır. Dünya'daki e-dönüşüm süreçlerinin ardından Türkiye'de e-dönüşüm sürecinde atılan ilk adım e-dönüşüm Türkiye projesi olmuştur. Bilişim teknolojilerinin her alana yayılması, e-dönüşüm sistemlerini zorunlu kılarak yeni kavramların doğmasına sağlamıştır. Bu kavramlar; e-devlet, e-okul, e-ticaret vb.'dır.

Bilişim teknolojilerindeki bu baş döndürücü gelişmelere ek olarak kişiler, kurumlar ve işletmelerin sahip oldukları veriler, bilgi hırsızlıkları, hackerlar, elektronik saldırılar, bilgi sızdırma, sosyal mühendislikler, kuruşların kendi çalışanlarınca oluşturulabilecek potansiyel iç saldırılar gibi çok geniş bir alana yayılan kaynaklardan gelen tehdit ve tehlikelerle karşı karşıya kalmaktadır. Casus yazılımlar, sosyal mühendislerin kullandığı yöntemler, kişisel veya kurumsal bilgilerin izinsiz olarak elde edilmesi ya da değiştirilmesi konusundaki tehditler artarak sürmektedir. Kurumlar ve kişiler bu tehlikelerle karşı karşıya kalmak durumundadır.

Kurumların ve kişilerin işleyişini kolaylaştıracak hizmetlerin internet ortamında sağlanması, açık ve özel ağlar arasındaki geçişler, bilgilerin halka açık sistemlerde paylaşılması gibi uygulamaların artmasıyla bilgilere erişimin sınırlandırılması ve denetlenmesi zorlaşmakta ve bilgi güvenliği zafiyetlerine sebep olmaktadır [6].

III. KİŞİSEL BİLGİLERİN GİZLİLİĞİNE DAİR TEHDİTLER

Yasal çerçeve ve bilgi güvenliği tanımı gereği kişilere ait özel bilgilere sadece ve sadece ilgili kişinin kendisi ve onay verdiği kişilerin ulaşması gerekmektedir. E-dönüşüm sistemlerindeki sunulan hizmetlerde ise bu konuda gerekli hassasiyet gösterilmemekle beraber saldırganlar tarafından geçilmesi çok da zor olmayan güvenlik kontrolleri bulunmaktadır. Bunların en yaygın olarak kullanılanı TC kimlik numarasıdır. Birçok otantisite sistemine bakılınca, TC kimlik numarasının sadece o kişinin bilebileceği bir numara gibi sahte bir güvenlik algısı insanlara aşılanmakta ve hatta bundan dolayı birçok insan TC kimlik numarasını diğer insanlardan bir sır gibi saklamakta ve nüfus cüzdanı kopyasının başka insanların elinde olmasından büyük çekince göstermektedir. Oysaki TC kimlik numarası ve diğer nüfus cüzdanı bilgilerinin bir başka insanın eline geçmesinin pek çok yolu bulunmaktadır ve zaten bu tür bilgiler kısmen de olsa birçok yerde isteğimiz dışı sergilenmektedir.

Nüfus cüzdanı kimlik bilgilerimize nasıl erişilebileceğine dair örneklere geçmeden önce, sadece TC kimlik numarası gibi bir bilgiyle neler yapılabileceğine birkaç örnekle değinelim. Örneğin (şu an bu çalışmanın yapıldığı an itibariyle), T.C. Yüksek Öğrenim Kredi ve Yurtlar Kurumu'ndan [15] katkı kredisi almış bir öğrencinin tüm borçlandırılma bilgilerine, ne zaman ödeme yapması gerektiğine, tarihleriyle birlikte geçmiş ödemelerine, ne kadar borcu kaldığına ve -ilginç bir şekilde- baba adına erişmek için sadece TC kimlik numarası yeterli olmaktadır. Bu örnek tek başına bile dikkate alındığında; bir kişinin sadece TC kimlik numarası bilen bir sosyal mühendis, artık onun ismini ve soy ismini, mezun olduğu lisans türünü, babasının adını, borç ve ödeme bilgilerini öğrenmiş bulunmaktadır. Bu kişinin böyle bir kaydına ulaşılması bile tek başına birçok gizil bilgi içermektedir. Böyle bir sorgulama sonucunun sorgulama ekranında başarıyla getirilmesi, o kişinin eğitiminin bir şekilde sonlandığını ve bu eğitim sürecinde bir öğrenim kredisine ihtiyaç duyduğu gerçeğini gösterir. Borç miktarı ve ödeme durumu, bireyin geçmişine ve sosyo-ekonomik durumuna dair küçük de olsa çeşitli ipuçları sunar. Yuvarlandıkça büyüyen kartopu misali, bir sosyal mühendis bu şekilde kurbanı hakkındaki topladığı bilgileri parça parça büyüterek ve her elde edilen bilgiyi başka yeni bir bilgiyi elde etme yolunda kullanarak amacına adım adım yaklaşabilir.

Sosyal mühendis kavramından söz etmişken konuyla ilgili bu önemli kavrama da kısaca değinmek gerekiyor. Sosyal mühendislik, basitçe, insanlara tanımadığı bir yabancı

için normalde yapmayacağı şeyleri yaptırmaktır. Sosyal mühendisler, etkileme ve ikna yeteneklerini başka insanları kandırmak ve onları gerçekte olmadığı bir insan olduğuna ikna etmek için kullanır. Bu şekilde, sosyal mühendis teknolojinin yardımı olarak ya da olmaksızın kişilerin zaaflarından yararlanıp bilgi elde eder [13]. Bilişim korsanları, sosyal mühendislik ataklarını gerçekleştirebilmek için de sosyal medya aracılığıyla ve kimlik hırsızlığı yoluyla elde ettiği bilgilerden faydalanırlar [14]. E-dönüşüm hizmetlerindeki güvenlik zafiyetleri ise sosyal mühendisler için oldukça kullanışlı bilgiler sağlayabilir ve sosyal mühendisler bu sistemlerden kurbanları hakkında elde ettikleri bilgileri, onları çeşitli psikolojik yöntemler aracılığıyla kandırma ve manipüle etmede kullanabilirler.

Bir başka e-dönüşüm sistemi olan T.C. Sağlık Bakanlığı uygulaması olan Merkezi Hekim Randevu Sistemi'ne [16] de kayıt olmak ve randevu kaydı oluşturmak için nüfus kimliği bilgileri yeterli olmaktadır. E-mail adresi, telefon gibi iletişim bilgileri de o kaydı oluşturan kullanıcının seçimine bağlı olduğu için kimlik bilgileri girilen kişiye ait olup olmadığının kontrolü mümkün olmamaktadır. Bu da, bir vatandaşın kimlik bilgilerini kullanarak onun adına sahte bir randevu kaydı oluşturabileceğimiz anlamına gelmektedir.

T.C. kimlik numarasının sadece kişinin kendisinin bildiği varsayımı üzerine bazı kurumlar, öğrenci ya da personellerinin kayıt sistemlerine ilk girişlerini yaparken kullanıcı adı olarak T.C. kimlik numarası, şifre olarak ise bu numaranın son/ilk 6 karakteri olacağı duyururlar. Bu yolla, gerek bilgi işlem departmanının gerekse de kullanıcının işini kolaylaştırmayı ve hızlılaştırmayı hedeflerler. İlk kullanıcı girişi yaptıktan sonra kullanıcıdan yeni bir parola belirlemesi istenir ve bu kullanıcının kendi belirlediği ilk parolası olur. Ancak bu tür sistemler, başka insanların izinsiz olarak hızlı davranarak ilk girişi yapmasını ve o sisteme erişerek bireyle ilgili bilgilere ve o sistemle ilgili işlemlere -fark edilip müdahale edilinceye kadar- erişim hakkı kazanmasını sağlar. Bazen de sadece kimlik bilgileriyle oluşturulan sahte bir giriş hesabıyla yapılan haksız işlemlerle gerçek kişi hak kaybına ya da yasal zarara uğratılabilir. Mağdur ise belki uzun vadede bu işlemlerin kendisine ait olmadığını ve sorumlu olamayacağını ispat edebilir ancak yine de belli bir süre için çeşitli zararlara uğrayacağı kesindir.

Kimlik bilgilerinin sahte kişilerin kullanımı ile kişilerin bilgilerin gizliliğinin ve kişiye erişim ihlali yoluyla verilebilecek zararların örnekleri çoğaltılabilir. Bu örneklerin bir kısmı daha lokal uygulamalar olup bazıları ise ulusal arenada yer almaktadır. Ancak hepsinin ortak noktası; e-devlet [17], ÖSYM [18] sistemlerinde olduğu gibi sadece kişinin kendisinin bizzat ilgili resmi ofislere müracaat ederek oluşturduğu bir parola ile giriş yapılmayıp başkalarının da erişebileceği bilgilerle giriş yapılmasına olanak tanınmasıdır.

E-dönüşüm sistemlerinde, kişisel bilgilerin gizliliği ve çevrimiçi bireysel yetki ve erişim ihlalleri açısından büyük güvenliği açığı oluşturan kimlik bilgisi temelli otantisite sistemlerini eleştirmemizin en büyük nedeni bu kimlik bilgilerine kısmen ya da tümüyle başkaları tarafından erişilip elde edilebilme durumudur. Hürriyet gazetesinin bir haberinde, İngiltere merkezli yaşam destek firması olan CPP Türkiye şirketinin GfK araştırma şirketine yaptırdığı araştırmaya göre, Türk halkının yüzde 86'sının kimlik hırsızlığına maruz kalmaktan korktuğunu, kadınların ise bu konuda erkeklerden daha fazla endişe duyduğunu ortaya çıkarmıştır. Kimlik hırsızlığı denilince, araştırma katılımcılarının, yüzde 42'sinin aklına ilk olarak nüfus cüzdanı, yüzde 26'sının aklına TC kimlik numarası ve yüzde 19'unun aklına ise ad-soyad bilgileri geldiğini görülmektedir. Araştırmayla ilgili bir başka dikkat çekici bulgu ise, günlük hayatlarında katılımcıların yüzde 46'sı kimlik hırsızlığı suçlarının başlarına gelmesinden, yüzde 20'si işlemedikleri suçlardan dolayı gözaltına alınmaktan, yüzde 6'sı dolandırılmaktan, yüzde 20'si ise itibar kaybına uğramaktan büyük endişe ve korku duymakta olduğudur.[19]. Sosyal mühendislikle ilgili literatür ve haber gündemlerine taşınmış yaşanmış vakalar göz önüne alındığında bu korkuların yersiz ve esassız olmadığı açıkça görülebilir.

Kimlik ve şahsi bilgilerimizin başkalarının eline geçebilmesinin pek çok farklı yolu bulunmaktadır. Sayısız kuruma, kişilere verdiğimiz nüfus cüzdanı, pasaport, ehliyet vb. fotokopiler, bu fotokopileri elde ederken ki süreçte kimliğin aslının ve fotokopisinin temas ettiği sayısız dijital cihaz ve maruz kaldığı insan görüşü, çeşitli kurumların gerek matbu gerek elektronik yolla TC kimlik numarası başta olmak üzere çeşitli kimlik bilgileriyle yaptıkları umumi ilanlar ve duyurular, illegal toplu kimlik bilgisi satışları, yetersiz kimlik doğrulama sistemlerine sahip e-dönüşüm servisleri, vatandaşların sosyal mühendislik konusundaki eğitim ve farkındalık eksikliği ilk akla gelenler olarak sıralanabilir. Kimlik bilgileri de işleyiş ve doğası gereği (soyadı, medeni hali, veriliş tarihi gibi bilgiler istisna olmak üzere) kolayca değişmeyeceği için bu bilgilerin bir kere yanlış ellere ulaşması kalıcı bir potansiyel tehlike oluşturacaktır. Kimlik kartlarında, kredi ve banka kartları gibi kaybolduğunda ya da gizliliğinin ihlal edildiğinin anlaşılınca iptal edilip yenilenerek bilgilerinin değiştirilmesi ve tehlikenin bertarafı gibi bir durum söz konusu değildir. Çoğu önemli kimlik bilgisi içeriği ise, vatandaşın doğumundan ölümüne kadar aynı kalmaktadır ya da çok nadir değişmektedir. Bu da, kimlik bilgilerinin bir kere bile ifşa edilmesi ya da elde edilmesinin geri alınamaz kalıcı bir problem yaratacağı ve bu bilgilerin güvenlik kontrolleri ve otantisite sistemleri için güvenilir bir başvuru kaynağı olamayacağı anlamına gelir.

İletişim bilgilerimiz de kişisel bilgilerimizin önemli bir parçasıdır ve sosyal mühendislerin de kullandığı önemli bir saldırı aracıdır. İletişim bilgileri kullanılarak, sahte e-posta, phishing, mesaj sağanakları, elektronik dolandırıcılık ve e-posta aldatmacaları yapılmaktadır [2]. Saldırganın iletişim bilgileriyle birlikte bazı kimlik bilgilerine de sahip olması bu tehlikelerin boyutunu arttırmaktadır. İletişim bilgilerimizin gizliliği konusunda ise durum kimlik bilgilerimizin gizliliğinden çok daha vahimdir. Sayısız kişi ve kuruma verdiğimiz bu bilgilerin de benzer şekilde gizliliğinin titizlikle korunması, sadece ilgili sorumlularca yalnızca gerektiği zamanlarda kullanılması esas olmalıdır.

E-dönüşüm hizmetleri kapsamında değerlendirilebilecek bir diğer servis ise fatura bilgilendirme ve ödeme sistemleridir. Bu sistemler, özelleştirme politikaları sonucu olarak genelde özel şirketler, kurumlar ve bankalar tarafından geliştirilmekte ve yönetilmektedir. Özellikle internet bankacılığında tüm kişilerin faturaları -ilgili fatura numarası bilinmek kaydıyla- herkesçe kontrol edilebilmektedir. Bu fatura çeşitlerinin başlıcalarını elektrik, su, doğal gaz, telefon, internet, televizyon, eğitim, sınav, SGK, vergi, ceza, icra ve diğer

kurum ödemeleri olarak sıralayabiliriz. Bazı kurumsal sitelerde ve internet bankacılığı sistemlerinde fatura sahibi olmaksızın her müşterinin fatura kaydına sadece ilgili numaranın bilinmesiyle ulaşılması ise yine kişisel bilginin ihlaline sebebiyet vermektedir. Çünkü fatura numarası da gizli ve karmaşık bir parola niteliği taşımamaktadır ve bir parola gibi kişi tarafından belirlenip değiştirilememektedir. Bazı fatura sorgulamalarında fatura ödeme tutarı ve tarihleriyle birlikte fatura sahibinin adı ve soyadı da kodlu veya kodsuz olarak karşımıza gelmektedir. Örneğin fatura sahibinin ismi, Me*** Yı*** olarak kısmen kapatılmış olarak doğrulama amaçlı gösterilmektedir. Ancak, isimlerinin baş harfleri bile sosyal mühendisler için diğer toplanılan bilgiler ışığında anlamlandırılabilir ve amaçları doğrultusunda kötüye kullanılabilir.

Fatura ve ödeme sistemlerinde, bankalar ve ilgili bazı kurumlar gerek sistemi basit tutmak gerekse de ödemeyi kolaylaştırmak adına herkese açık, ID temelli bir sorgulama sistemi oluşturmuşlardır. Bu sisteme göre, bir bireye ait ödeme tutarını içeren kaydına gerekli ID bilinerek erişilebilmekte ve istenirse de ödeme yapılabilmektedir. Bu ID'ler sadece ilgili kişinin bildiği ve gerektiğinde değiştirebileceği şifre güvenliğini sağlayamaz. ID'ler genelde rakamlardan oluşmakta olup yeni müşteriler sisteme eklendikçe belirli bir ardışıklıkta arttırılarak verilmekte olması da tahmin edilebilirliğini kolaylaştırmaktadır. Rastgele bir borçlunun değil de hedef şahsın ödeme bilgisinin ID'sine ve dolayısıyla ödeme bilgilerine nasıl erişilebileceğine gelirsek bunun için de çeşitli senaryolar bulunabilir. Elektrik/su gibi herkesçe kullanılan fatura türlerini baz alırsak, ilgili kurum görevlilerinin meskenleri kapı kapı dolaşarak her ayın belli bir gününde posta kutularına bıraktığı bilgilendirme kağıtlarından gerekli ID'ye fatura sahibinden önce bir kere erişilmesi ve saldırganca bu ID'nin saklanması gelecek tüm fatura bilgilerine ay ay erişimine olanak sunacaktır. Bu gibi çeşitli ihtiyaçlara yönelik fatura tutarları bilgisiyle bireye ne gibi zararlar verilebileceği saldırganın amaç ve tutumlarına göre şekillense de, izinsiz erişim tek başına bile gizlilik hakkının bir çeşit ihlalidir. Bir bireyin elektrik faturaları düzenli olarak incelense ve birey hakkında önceden toplanmış diğer başka bilgiler ışığında zekice analiz edilerek yorumlansa birçok örtülü anlam çıkartılabilir. En basitinden, aylık elektrik ve su faturalarındaki dikkat çekici bir düşüş meskenin kullanım ve doluluk oranıyla ilgili ipuçları sağlayabilir. Sosyal mühendislerin sıklıkla kullandığı başka bir yöntem ise, çeşitli fatura borçlarına ait bilgileri kullanarak alacaklı kurumu taklit etmeleridir ve elde edilen bilgi ve ipuçları da bu yöntemin uygulanması doğrultusunda saldırgana olanaklar ve kolaylıklar sağlar.

Bireylerineğitimbilgilerini,akademikvekariyerözgeçmişlerini, yaşamsal ihtiyaçlarına dair harcamalarıyla ilgili bilgilerini (elektrik, su, doğal gaz, telefon, internet, televizyon, vb. faturaları), sınav ve mülakat sonuçlarını sadece kendilerinin ve onay verdiği insanların görüntüleyebilmelerini sağlamak kişilerin bilgilerinin gizliliğinin korunması konusunda önemli ve zorunlu bir gerekliliktir. Bunu sağlamak için de bu tür bilgilere sadece bazı ID ve numaralarla sorgulama yapılarak erişilmesinin engellenmesi ve kişiye özel parolalar ve geçici SMS şifresi gibi kimlik doğrulama tekniklerinden yararlanılması önerilebilir. ÖSYM gibi bazı kurumlar da bu doğrultuda sadece kimlik bilgisi temelli sonuç sorgulama yöntemini bırakıp bireysel olarak oluşturulan parola kontrollü otantisite sistemlerine geçmiştir.

IV. SONUÇ

En son standartlarda, en yeni teknolojiyle donatılan bir sistemde dahi yarın için farklı tehditler vardır. Hiçbir zaman için, oluşturulan sistemlerin yüzde yüz etkin başarı sağlaması mümkün değildir. Yarının ihtiyaçlarını karşılayabilmek için devamlı gelişen bir süreç içerisinde hareket edilmesi gerekmektedir [5]. e-Dönüşüm hizmetlerini veren kurum ve kuruluşların da gelişen teknolojiyi ve güvenlik yöntemlerini dinamik bir şekilde takip etmesi ve kendi sistemine entegre etmesi gerekmektedir. Aynı zamanda kullanıcıların da dikkat etmesi ve önlemler alması gereken bir takım konular mevcuttur. Bu nedenle yapılan bu çalışmada gerek e-dönüşüm hizmetlerinin yöneticileri için gerekse kullanıcıları için önemli görülen noktalar açıklanmaya çalışılmış ve geliştirilebilecek noktalar üzerinde durulmuştur. E-dönüşüm hizmetlerinde kişisel verilerin gizliliğine dair değindiğimiz tehditler göz önüne alındığında ve incelendiğinde bu sistemlerde kişilerin verilerinin gizliliğinin korunmasının her zaman ön planda tutulmadığı, yazılım geliştirme-uygulama konusunda sadelik, kullanım kolaylığı, ulaşılabilirlik ve anlaşılabilirlik kavramlarının öne çıktığı görülmektedir. Ancak gerek kurum gerekse müşteri için sağlanacak herhangi bir kolaylık kişisel bilgilerin gizliliğini arka plana atacak kadar önemli olmamalı ve gizlilik hakkından asla ödün verilmemelidir.

Yukarıda değinilen tehditler ve çıkış noktaları düşünüldüğünde bu tehditleri azaltmak için yasa, idare, kurum ve eğitim düzeylerinde atılacak önemli adımlar bulunmaktadır. Bu adımlar sahtekarlık, dolandırıcılık ve kişisel bilgilerin istenmeyen kişilerce erişimini en aza indirecek önlemlerden oluşacaktır. İlk olarak sadece kişinin belirlediği ve bildiği şifre dışında hiçbir kişisel bilgi otantisite yöntemi olarak kullanılmamalıdır. Bunun sağlanması için gerekli yasalar spesifik şekilde düzenlenmeli ve titizlikle uygulanarak denetlenmelidir. Telefon fatura borcundan bir sınav sonucu sorgulamasına kadar tüm elektronik sorgular sadece kişinin devlet onaylı bir otantisite sisteminden geçilerek öğrenilmelidir. Bu önerilen sistem de e-devlet sisteminin tüm kurumlarla uyumlu çalışabilecek bir otantisite sistemi geliştirmesine bağlı olmaktadır. Nasıl üçüncü parti birçok web sitesi facebook sosyal medya hesabıyla bir eklenti ile giriş yapmaya ve servislerinden yararlanılmaya olanak sağlıyorsa bu kurumlar da devlet tarafından ortak bir kimlik kontrol sistemi ile kişiye ait verinin sadece ilgili kişinin tek ve eşsiz e-devlet hesabından ulaşıldığını kontrol edebilir. Eğer bu dev ve kompleks sistem uygulamaya geçirilirse e-devlet hesabının güvenliği ve korunması çok daha fazla önem kazanacağından e-devlet sistemindeki otantisite sitemi de bankacılık sistemlerindeki gibi birkaç aşamalı hale getirilmesi gerekecektir. İnternet bankacılığında uygulanan hem vatandaş tarafından belirlenecek bir şifre veya e-imza hem de telefon SMS şifresiyle doğrulama çifte bir koruma ve güvenlik sağlayabilir. Ve e-devlet hesabına yapılan girişlerin günlükleri istendiğinde IP bazlı olarak inceletilebilir ve hesaba erişildiğinde yine kişi tarafından resmi yollarla belirlenecek olan e-mail adresine oturum açma bilgilendirme e-mailleri gönderilebilir.

Saldırganların, sms, posta, e-mail, telefon görüşmeleri vs. yollarla resmi kurumları taklit etmelerinin önüne geçilmesi

için kurumların e-devlet tarafından resmi olarak onaylanıp onaylanmadığı ve gerekli güvenlik politikalarına ve yönergelerine uyulup uyulmadığı birey/kullanıcı tarafından resmi bir kanalla doğrulanabilmelidir. Bunu sağlamak için de, her yazışmaya/görüşmeye münhasır hükümet uzantılı bir portal aracılığıyla geçerliliğini ve güvenilirliğini kontrol etmesine olanak sağlayan bir özel kod kullanıcıya sağlanabilir. Tüm yasal, teknik ve idari önemler kadar önemli olan ve güvenliğin en zayıf halkalarından birisi olan insan faktörü konusunda ise bireylerin konuyla ilişkili tüm tehditler ve önlemleri konusunda da eğitilmeleri ve uyarılmaları da bir o kadar önem ve zorunluluk teşkil etmektedir. Bu eğitim tüm örgün eğitim basamaklarında verilebileceği gibi televizyondaki kamu spot yayınları gibi yollarla da aktarılabilir.

KAYNAKÇA

[1] "Hanehalkı Bilişim Teknolojileri Kullanım Araştırması", 2015 [Online], Available: http://www.tuik.gov.tr/PreHaberBultenleri.do?id=18660

[2] Öğütçü, G. "e-Dönüşüm Sürecinde Kişisel Bilişim Güvenliği Davranışı ve Farkındalığının Analizi", Başkent Üniversitesi, Fen Bilimleri Enstitüsü, 2010

[3] Keser, L., Kaya, M. B., Kınıkoğlu B., Şahbaz, U., Alpaslan, İ. B. ve Sökmen, A. "Türkiye'de Kişisel Verilerin Korunmasının Hukuki ve Ekonomik Analizi", İstanbul Bilgi Üniversitesi, Bilişim ve Teknoloji Hukuku Enstitüsü ve TEPAV, 2014

[4] ISO/IEC 27001: 2013, "Information technology — Security techniques — Information security management systems — Requirements"

[5] Akay, İ. G. "Bilgi Güvenliği Yönetim Sistemleri: Bilgi Güvenliği Uygulama Mülakatları", Bilecik Şeyh Edebali Üniversitesi, Sosyal Bilimler Enstitüsü, 2014

[6] Ersoy, E. "Gizlilik, Bireysel Haklar, Kişisel Verilerin Korunması", Telekomünikasyon Kurumu, 2015 [Online], Available: http://ab.org.tr/ab06/bildiri/6.doc

[7] Öz, E. ve Bozdoğan, D. "Türk Vergi Sisteminde e-maliye Uygulamaları", Süleyman Demirel Üniversitesi, İİBF Fakülte Dergisi, 2012

[8] Bensghir, K. T. "E-dönüşüm ve E-imza Uygulamaları", TODAİE e-Devlet Merkezi Uygulamaları e-imza Semineri, 2011

[9] Çetiner, T. "E-dönüşümde Türkiye Nerede?", Uluslararası Ekonomik Sorunlar, 2014

[10] Ketizmen, M. ve Ülküderner, Ç. "e-Devlet Uygulamalarında Kişisel Verilerin Korunmaması"

[11] Karaaslan, E., Koç ve S., Akın, G." Vatandaşlık Numarası Bazlı e-Devlet Sistemlerinde Kişisel Veri Mahremiyeti Durum Saptaması", 2010

[12] "Başbakanlık Kanunlar ve Kararlar Genel Müdürlüğü, Kişisel Veri Güvenliği Kanun Tasarısı", 2014 [Online], Available: http://www.kgm.adalet.gov.tr/Tasariasamalari/Tbmmkms/Tbmmkom/ki%C5%9Fisel%20veriler.pdf

[13] Mitnick, K., ve Simon, W. L. "The Art of Deception: Controlling the Human Element of Security". New York: John Wiley & Sons, 2002.

[14] U. Yavanoğlu, Ş. Sağıroğlu ve İ. Çolak, "Sosyal Ağlarda Bilgi Güvenliği Tehditleri ve Alınması Gereken Önlemler", Politeknik Dergisi, cilt 15, no.1, pp. 15 -27, 2012.

[15] "Kredi Geri Ödeme Sorgu Ekranı", 2015 [Online], Available: https://www.kyk.gov.tr/web/YENI_GRODEME/geriOdemeParametreGiris.do

[16] "T.C. Sağlık Bakanlığı Merkezi Hekim Randevu Sistemi", 2015 [Online], Available: http://www.hastanerandevu.gov.tr/Randevu/

[17] "e-Devlet Kapısı", 2015 [Online], Available: https://www.turkiye.gov.tr/

[18] "ÖSYM Aday İşlemleri Sistemi", 2015 [Online], Available: https://ais.osym.gov.tr/

[19] "Kimlik hırsızlığı Türk halkının korkulu rüyası", 2011 [Online], Available: http://www.hurriyet.com.tr/planet/19169690.asp.

PARMAK İZİNDEN CİNSİYET TANIMA: YENİ BİR VERİTABANI İLE TEST

Eyüp Burak CEYHAN, Şeref SAĞIROĞLU

Gazi Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü
Maltepe, ANKARA
eyupburak@gmail.com, ss@gazi.edu.tr

Özet — Bu çalışmada, adli vakalarda ve güvenlik kontrollerinde en çok tercih edilen biyometrik özelliklerden biri olan parmak izi kullanılmıştır. Daha önce geliştirmiş olduğumuz parmak izinden cinsiyet tanıyan zeki sistemimizin yeni bir parmak izi veritabanıyla testi gerçekleştirilmiştir. Sunulan çalışmada Biosecure şirketine ait Multimodal veritabanındaki parmak izleri kullanılmıştır. Aynı veritabanındaki yüz verileri ile parmak izleri ilişkilendirilerek cinsiyet bilgileri de verisetine eklenmiştir. Parmak izi tepe yoğunluğu bilgileri kullanılarak sistemin cinsiyet tahmin başarısı hesaplanmıştır. Yapay sinir ağları ve 10-kat çapraz doğrulama kullanılarak elde edilen başarı %80 olarak tespit edilmiştir. Ayrıca bayların bayanlara göre daha düşük parmak izi tepe yoğunluğuna sahip olduğu görülmüştür. Sunulan çalışma Biosecure veritabanıyla ileride yapılacak çalışmalara referans olacak, ayrıca önerdiğimiz sistemin başarısını artırmak için yapılacak çalışmalara da ışık tutacaktır.

Anahtar Kelimeler — Biyometri, parmak izi, cinsiyet, tepe yoğunluğu, Biosecure veritabanı.

Abstract — In this study fingerprint, one of the most preferred biometric feature in criminal cases and in security controls, was used. Test of our intelligent system that recognizes gender from fingerprint which we developed before was done with a new fingerprint database. In presented work, fingerprints in Multimodal database of Biosecure company was used. Gender information is added to the dataset with associating the face data and the fingerprint data in the same database. Gender prediction success was evaluated using fingerprint ridge density. Accuracy was obtained as 80% using Artificial Neural Network algorithm and 10-fold cross validation technique. Also, it was found that fingerprint ridge density of men is lower than women. The presented study will be a reference to future studies using Biosecure database, also will shed light on the work to improve our proposed system's accuracy.

Index Terms — Biometrics, fingerprint, gender, ridge density, Biosecure database.

I. GİRİŞ

Günümüzdeki birçok uygulamada parmak izi görüntüsü verisini paylaşma ihtiyacı gizliliğin korunması ihtiyacını artırmıştır. Parmak izi görüntülerinden yumuşak biyometriklerin tahmini için hazır algoritmaların kullanımı mümkündür. Bir kişiyi var olan bir parmak izi ile birebir eşleştirmek mümkün olmasa da, yaş ve cinsiyet bilgisinin elde edilmesi istenmeyen sonuçlara yol açabilir. Yazarlar bu çalışmada yumuşak biyometrikleri gizlemek için parmak izi görüntülerini kimliksizleştirme üzerine yoğunlaşmışlardır. Çalışmada, öncelikle parmak izindeki yumuşak biyometriklerin kimliksizleştirilmesi için genel bir çerçeve sunulmuştur. Bu çerçeve, geçici görüntü filtreleme kullanarak parmak izi görüntülerinden başarılı bir cinsiyet tahmini riskini azaltmak için geliştirilmiştir. Sunulan yaklaşım West Virginia Üniversitesi'nde toplanan yuvarlanmış parmak izi veriseti kullanılarak yapılan deneylerle değerlendirilmiştir. Deneyde KNN algoritması (k=1) kullanılarak 10-kat çapraz doğrulama yöntemiyle model eğitilip test edilmiştir. Cinsiyet tahmin etme başarısı normalde %88,7 iken oluşturulan model ile tekrar veriseti test edildiğinde sistemin cinsiyet tahmin etme başarısı %50,5'e düşmüştür. Sonuçlar sunulan metodun parmak izi görüntülerinden cinsiyet tahminini önlemede başarılı olduğunu göstermektedir [1].

Parmak izleri morfolojik, biyolojik, antropolojik ve adli çalışmalarda kullanılan önemli bir değerdir. Olay yerinden ve olayda kullanılan eşyaların üzerinden toplanan parmak izleri, suçluları, kurbanları veya yüzeye dokunan diğer kişileri tespit etmede başarıyla kullanılmaktadır. Üst derideki tepelerin kalınlığı kişiden kişiye değişmektedir. Bayanlar baylardan daha yüksek tepe yoğunluğuna sahiptir. Sunulan çalışmada, Kuzey Hindistan toplumuna ait parmak izlerinin sol üst, sağ üst ve alt alanlarındaki parmak izi tepe yoğunluğundan cinsiyet ayrımı yapılmaya çalışılmıştır. Yaşları 18 ile 25 arasında değişen 97 bay 97 bayan toplam 194 kişi çalışmaya dâhil edilmiş ve katılanların tüm parmak izleri toplanmıştır. Böylelikle toplamda 1940 parmak izi elde edilmiş ve parmak izinin sol üst, sağ üst ve alt alanlarındaki üst derideki tepeler her parmak için sayılmıştır. Her üç alandaki ve cinsiyetler arasındaki parmak izi tepe yoğunluğu t-test kullanılarak istatistiksel olarak karşılaştırılmıştır. Sonuçlar her üç alan için de bayanların baylardan daha fazla tepe yoğunluğuna sahip olmaya eğilimli olduğunu göstermektedir. Parmak izinin sağ üst ve sol üst alanlarındaki parmak izi tepe yoğunluğu alt alandan önemli ölçüde daha yüksektir. Sunulan çalışma, parmak izi tepe yoğunluğunun olay yerinden alınan kime ait olduğu bilinmeyen parmak izinin sahibinin cinsiyetini ayırt etmek için uygun ve kullanışlı bir parametre olarak kullanılabileceğini önermektedir [2].

Parmak izinden cinsiyet tespiti için Kızılderili toplumu üzerinde yapılan bir çalışmada, parmak izi tepe yoğunluğunun topolojik ve cinsiyete göre farklılıkları tespit edilmeye çalışılmıştır. Çalışmada kullanılmak için Arjantin'in kuzeybatısında (Jujuy ili) bulunan farklı rakımlardaki iki farklı bölgeden kişiler seçilmiştir. Sonuçlar İspanyol toplumundan elde edilen sonuçlar ile karşılaştırılmıştır. 393 genç Arjantinli bay ve bayanın tüm parmaklarının verileri kullanılmıştır. Bunlardan 193'ü Puna-Quebrada (deniz seviyesinden 2500 metreden daha fazla yükseklikte) bölgesinden 200'ü ise Ramal (deniz seviyesinden 500 metre yükseklikte) bölgesinden seçilmiştir. Parmak izinin üç farklı bölgesinden elde edilen tepe yoğunluğu değerleri her kişinin 10 parmağı için de elde edilmiştir. Her iki grup için de farklı alanlarda belirgin farklılıklar elde edilmiştir. Sağ üst tepe yoğunluğu > sol üst tepe yoğunluğu > alt tepe yoğunluğu olduğu tespit edilmiştir. Gruplardaki baylar arasında farklılık gözlenmemiştir fakat bayanlar sağ üst ve alt bölgelerde birbirinden belirgin bir şekilde farklılık göstermektedir. Tüm alanlarda, her parmak için, bayanlar baylardan daha yüksek parmak izi tepe yoğunluğuna sahiptir. Bayes teoremi kullanılarak tepe yoğunluğu eşik değeri elde edilmiş, Arjantinli ve İspanyol toplumu arasındaki ayrım için de bu eşik değer kullanılmıştır [3].

Yapılan çalışmada, Türk bireylerden elde edilen parmak izi tepe yoğunluğu değerleriyle parmak izinden cinsiyet tespiti yapılmaya çalışılmıştır. Veriseti 17-28 yaşları arasında değişen 118 bayan 88 bay toplam 206 öğrenciye ait basit mürekkepleme metoduyla elde edilen parmak izlerinden oluşmaktadır. Her kişiden 10'ar parmak izi alınarak sol üst, sağ üst ve alt bölgelerden 5 mm x 5 mm'lik kare kesitin köşegeni üzerindeki tepeler sayılmıştır. Üç farklı bölge ve cinsiyetler arasındaki parmak izi tepe yoğunluğu Mann Whitney U testi ve Friedman testi kullanılarak karşılaştırılmıştır. Elde edilen sonuçlara göre, çalışılan tüm bölgelerde ve tüm parmaklarda bayanlar baylardan daha fazla tepe yoğunluğuna sahiptir ve parmak izinin sol üst ve sağ üst bölgelerindeki tepe yoğunluğu alt bölgeden belirgin bir şekilde daha fazladır [4]. Li ve arkadaşları tarafından yapılan çalışmada, yüz ve parmak izi bilgilerinden cinsiyet tanıma üzerine çalışılmıştır. Cinsiyet tanıma için güvenilir ve ayırt edici bir performansı sağlayabilmek için kişinin parmak izi ve yüzündeki görsel gözlemler birleştirilerek sonuç alınmaya çalışılmıştır. Kendi veritabanlarını kullanarak gerçekleştirdikleri deneylerde başarılı sonuçlar almışlardır [5].

Parmak izleri kullanılarak cinsiyet sınıflandırma problemini ele alan bir başka çalışmada makine öğrenmesi kullanılarak parmak izleri arasındaki farklar belirlenmeye çalışılmıştır. Veritabanındaki her görüntü, tepe kalınlığının vadi kalınlığına oranı (RTVTR) ve tepe yoğunluğu değerlerinden oluşan bir öznitelik vektörü kullanılarak kaydedilmiştir. Destek vektör makineleri (SVM) kullanılarak 150 bay ve 150 bayana ait parmak izlerinden oluşan set ile eğitim işlemi yapılarak bay ve bayanların öznitelik vektörleri örüntüleri için başarılı bir sınıflandırma işlemi fonksiyonu elde edilmiştir. Deney sonuçları sonucunda geliştirilen sistemin adli antropolojide kullanılabileceği ve %96 oranında başarılı sınıflandırma gerçekleştirdiği tespit edilmiştir [6].

Rajan ve arkadaşlarının yaptığı bir çalışmada, cinsiyeti sınıflandırmak için çeşitli biyometrik özelliklerin kullanıldığı fakat tek bir biyometrik özelliğin kullanılarak cinsiyet sınıflandırmanın düşük bir başarı oranı sunduğu vurgulanmaktadır. Bu sebeple yazarlar iris ve parmak izi biyometrik özelliklerinin birleşimini kullanarak cinsiyeti sınıflandırmaya çalışmışlardır. İris görüntüsünden çıkarılan öznitelikler ortalama ve standart sapma, parmak izi görüntüsünden çıkarılan öznitelikler ise tepe kalınlığının vadi kalınlığına oranıdır (RTVTR). 50 kişiden alınan 100 iris görüntüleri ve 120 kişiden alınan parmak izi görüntülerinin her ikisinden çıkarılan öznitelikler sinir ağlarını eğitmek için kullanılmıştır. Fakat çalışmada sinir ağlarının yapısından hiç bahsedilmemiştir. Sonuç olarak cinsiyeti sınıflandırmak için uygun bir öznitelik vektörü oluşturduklarını iddia etmektedirler [7].

Rajesh ve Punithavalli parmak izlerinin çok çözünürlüklü analizi ile cinsiyet tahmin etmek için yeni bir yaklaşım sunmuşlardır. Frekans bölgesinde parmak izlerini analiz etmek için ayrık dalgacık dönüşümü (DWT) kullanılmıştır. Sınıflandırma işlemi için gauss karışımı modellenmiştir. Öznitelikler olarak DWT katsayıları kullanılmış ve sıralama ile sadece baskın öznitelikler seçilerek sınıflandırma için Gauss Karışım Modeli'ne (GMM) verilmiştir. Geliştirilen sistem 80 bayan 100 bay toplam 180 kişilik bir veritabanı ile oluşturulmuştur. Test sonuçları sunulan sistemin 16 gauss yoğunluğuyla 3. seviye DWT ayrışımında %92,67 başarıya ulaştığını göstermektedir [8].

Parmak izi kanıtı bugüne kadar mahkemelerde en güvenilir ve kabul edilebilir kanıttır. Parmak izleri olay yerinden, eski anıtlardan ve kazıyla ortaya çıkan eserlerden elde edilir. Parmak izlerinin kimlik saptamanın etkin bir metodu olarak büyük bir potansiyeli olduğundan dolayı, sunulan çalışmada kişinin parmak izi ile cinsiyeti arasında korelasyon olup olmadığını analiz etmek için ayrık dalgacık dönüşümü (DWT) ve tekil değer ayrışımı (SVD) kullanılarak araştırılmıştır. Sınıflandırma için K en yakın komşu algoritması (KNN) kullanılmıştır. Yazarlar geliştirdikleri bu metodu 500 bay ve 500 bayandan alınan toplam 1000 parmak izinden oluşan bir veritabanı ile denemişlerdir. Yaptıkları analizde bayanların sol serçe parmaklarında parmak izinden cinsiyet sınıflandırma başarısı %82,90 olarak tespit edilmiştir. Ayrıca bay ve bayanların diğer parmaklarındaki sınıflandırma başarısı da baylarda %80,40 ve bayanlarda %76,84 olarak tespit edilmiştir. Baylarda en yüksek başarı sol el işaret parmağı ile, bayanlarda en yüksek başarı ise sol el serçe parmağı ile elde edilmiştir. Yine en yüksek başarı bay ve bayanların sol elleri ile altıncı seviyede elde edilmiştir [9].

Parmak izlerinin diğer biyometriklere göre uygulanabilirliği, birbirlerinden belirgin bir şekilde farklı olması, sabit olması, güvenilir olması ve kabul edilebilir olması gibi bazı avantajları olduğundan dolayı dünya genelinde güvenlik ve kişi saptamada kullanılmaktadır. Ayrıca dünya genelindeki mahkemelerde yasal bir kanıt olarak da kullanılmaktadır. Sunulan çalışmada 100 bay 100 bayan toplam 200 kişinin parmak izinden oluşan bir veritabanı kullanılmıştır. Sınıflandırmada öklid uzaklığını kullanarak test edilen parmak izlerini bay veya bayan olarak sınıflandıran k en yakın komşu sınıflandırıcı (KNN) kullanılmıştır. Baylarda %50 civarı bayanlarda ise %70 civarı bir sınıflandırma başarısı elde edilmiştir [10].

Sunulan bir başka çalışmada Marathi toplumundan 100 bay ve 100 bayan toplam 200 kişinin her iki başparmakları alınarak oluşturulan toplamda 400 parmak izinin tepe yoğunluklarında cinsiyetler arasında belirgin bir fark olup olmadığı araştırılmıştır. Çalışmaya dahil edilen kişilerin yaş aralığı 18-30'dur. Sağ ve sol başparmak izlerinin tepe yoğunluğu yeni tasarlanan bir model ile belirlenip istatistiksel olarak analiz edilmiştir. Bay ve bayanlar için frekans dağılımından (LoC, RoC ve birleşimi) elde edilen olasılık yoğunlukları, Bayes teoremi kullanılarak kişilerin bilinen tepe sayıları için cinsiyet atamasının olabilirlik oranı ve sonsal olasılığı hesaplanması için kullanılmıştır. Sonuç olarak belirlenen alanlar tek tek ve birleştirilmiş olarak analiz edildiğinde bayanların baylara göre her iki durumda da daha fazla başparmak izi tepe yoğunluğuna sahip olduğu görülmüştür. T-test uygulanarak LoC (merkezin solu), RoC (merkezin sağı) ve birleştirilmiş (LoC+RoC) alanlarında bay ve bayan parmak izlerinin tepe yoğunluklarındaki farklar $p<0,01$ seviyesinde istatistiksel olarak farklı bulunmuştur. Ayrıca bayların LoC değeri 11,58, RoC değeri 11,82 ve LoC+RoC değeri 23,40 iken bayanların LoC değeri 14,6, RoC değeri 14,56 ve LoC+RoC değeri 29,16 olarak tespit edilmiştir. Bu sonuçlar parmak izi tepe yoğunluğu ile cinsiyet arasında ilişki olduğunu göstermektedir [11].

II. KULLANILAN MATERYAL VE ELDE EDİLEN SONUÇLAR

Sunulan çalışmada cinsiyet sınıflandırması için parmak izi tepe yoğunluğu değerlerini baz alarak sınıflandırma yapan 2012 yılında önerdiğimiz sistem [12] kullanılmıştır. Veritabanı olarak ise Biosecure Multimodal veritabanındaki parmak izleri arasından 50 bay ve 50 bayana ait parmak izleri seçilerek kullanılmış ve geliştirilen sistemin cinsiyet sınıflandırma başarısı ölçülmüştür. Veritabanındaki parmak izlerinin bay veya bayana ait olduğu, parmak izlerine denk gelen yüz görüntülerine bakılarak verisetine eklenmiştir.

DS2 veriseti PC-tabanlı, çevrimdışı, masaüstü ortamında, denetimli koşullarda 2 oturum halinde toplanan bir verisetidir. Bu verisetinde ses, yüz, imza, parmak izi, el ve iris biyometrikleri işlenmemiş halde bulunmaktadır. Toplam donör sayısı 667'dir. Fakat bu 667 kişinin yukarıda bahsedilen tüm biyometrikleri verisetinde yoktur. Veriseti toplanırken sağlanan ortamda, veri toplama donanımı için geniş bir masa ve denetmen ile katılımcı için hazırlanan iki rahat sandalye bulunmaktadır. Veri elde etme donanımı standart bir bilgisayar ve bilgisayara USB veya Bluetooth arayüzü ile bağlanmış birkaç veri toplama sensöründen oluşmaktadır. Multimodal veritabanında DS3 verisetinden de parmak izleri bulunmaktadır. DS3 verisetinde mobil cihaz tabanlı, bina içi ve dışında kontrolsüz durumlarda çekilmiş parmak izleri bulunmaktadır. 713 donörden 2 oturum dahilinde veriler toplanmıştır. Bu verisetinde ses, yüz, imza ve parmak izi biyometrikleri işlenmemiş halde bulunmaktadır. Verisetindeki parmak izi verileri PDA HP iPAQ hx2790 ile elde edilmiştir [13].

Çalışmada bayların bayanlara göre daha düşük parmak izi tepe yoğunluğuna sahip olmaları bilgisinden yola çıkılarak geliştirilen sistem kullanılmıştır. Sistem parmak izinin en üst boğumundaki merkez noktasını baz alarak sol elden alınan parmak izlerinde merkezden sağ üst köşeye doğru olan 80 piksellik kesiti, sağ elden alınan parmak izlerinde ise sol üst köşeye doğru olan 80 piksellik kesiti dikkate alarak işlem yapmaktadır. Alınan kesitin merkez noktasından karşı köşeye kadar olan köşegen üzerindeki piksel değerleri ve bu köşegen üzerindeki tepe sayısı zeki sisteme giriş olarak alınmakta, geliştirilen zeki model ile eğitilerek test edilmektedir. Geliştirilen YSA yapısında tan, log, pur transfer fonksiyonları kullanılarak 3 ağ katmanında sırasıyla 50, 30, 1 nöron kullanılmıştır. Biosecure veritabanındaki rastgele 50 bay ve 50 bayana ait parmak izleri sisteme verilerek 10-kat çapraz doğrulama yöntemiyle eğitim ve test işlemleri gerçekleştirilmiştir.

Biosecure veritabanıyla elde edilen başarı [12] numaralı çalışmamızla karşılaştırılmış, farklı toplumlardan toplanan parmak izleriyle oluşturulmuş olan Biosecure Multimodal veritabanının da [12] numaralı çalışmada sunulan aynı yöntem kullanılarak benzer bir sınıflandırma başarısına sahip olduğu gözlenmiştir. [12]'de sistemin başarısı Türkiye vatandaşları veritabanı kullanılarak %78 olarak elde edilmiş, sunulan çalışmada ise Biosecure veritabanı kullanılarak sistemin sınıflandırma başarısı %80 olarak elde edilmiştir. Sonuçlar Tablo 1'de diğer çalışmalarla karşılaştırılmıştır

Kaynak	Sınıflandırma Algoritması	Örnek Sayısı	Veritabanı	Sınıflandırma Başarısı
[9]	KNN	500 Bay 500 Bayan	Kendi veritabanları	%78
[10]	KNN	100 Bay 100 Bayan	Kendi veritabanları	~%60
[12]	YSA	375 Bay 375 Bayan	Kendi veritabanımız	%78
Sunulan Çalışma	YSA	50 Bay 50 Bayan	BIOSECURE Multimodal veritabanı	%80

Tablo 1. Farklı çalışmalardaki parmak izinden cinsiyet sınıflandırma başarılarının karşılaştırılması.

III. SONUÇ VE TARTIŞMA

Adli vakalarda parmak izi en çok tercih edilen kanıtlardan biridir. Parmak izleri bu özelliklerinden dolayı biyometri çalışanlar tarafından en fazla araştırılan konulardandır. Adli bir vakayı açığa çıkarmada olay yerinde bulunan parmak izinin kime ait olduğu suçlu veritabanında kayıtlı değilse, bu parmak izini kullanarak zanlı hakkında en fazla bilgiye ulaşma en mantıklı yol olarak görülmektedir. Bu sebeple parmak izinden cinsiyet tahmin edilerek zanlı sayısı %50 oranında azaltılmaktadır. Bu da çok büyük veritabanlarında güvenlik birimleri için hem zamandan hem de emekten tasarruf etmeyi sağlamaktadır. Sunulan çalışmada daha önce yayınladığımız çalışmalarda önerdiğimiz parmak izinden cinsiyet tanıyan zeki sistem ile Biosecure firmasına ait Multimodal veritabanı kullanılarak cinsiyet tahmini başarısı ölçülmüştür. Elde edilen sonuçlara göre cinsiyet sınıflandırma başarısının %80 olduğu ve bayların bayanlara göre daha düşük parmak izi tepe yoğunluğuna sahip olduğu tespit edilmiştir. Sunulan çalışma Biosecure veritabanıyla ileride yapılacak çalışmalara referans olacak, önerilen sistemin başarısını artırmak için yapılacak çalışmalara ışık tutacaktır.

TEŞEKKÜR

Biosecure veritabanının alınmasında maddi desteklerinden dolayı Gazi Üniversitesi Bilimsel Araştırmalar Projeleri Birimi'ne teşekkür ederiz.

KAYNAKÇA

[1] Lugini, L., Marasco, E., Cukic, B., Dawson, J., "Removing gender signature from fingerprints," 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 1283-1287, (2014).

[2] Krishan, K., Kanchan, T., Ngangom, C., "A study of sex differences in fingerprint ridge density in a North Indian young adult population", Journal of Forensic and Legal Medicine, 20: 217-222, (2013).

[3] Gutiérrez-Redomero, E., Sánchez-Andrés, A., Rivaldería, N., Alonso-Rodríguez, C., Dipierri, J. E., Martín, L. M., "A comparative study of topological and sex differences in fingerprint ridge density in Argentinian and Spanish population samples", Journal of Forensic and Legal Medicine, 20: 419-429, (2013).

[4] Oktem, H., Kurkcuoglu, A., Pelin, İ. C., Yazici, A. C., Aktas, G., Altunay, F., "Sex differences in fingerprint ridge density in a Turkish young adult population: A sample of Baskent University", Journal of Forensic and Legal Medicine, 32: 34-38, (2015).

[5] Li, X., Zhao, X., Fu, Y., Liu, Y., "Bimodal gender recognition from face and fingerprint", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2590-2597, (2010).

[6] Arun, K. S., Sarath, K. S., "A machine learning approach for fingerprint based gender identification", IEEE Recent Advances in Intelligent Computational Systems (RAICS), 163-167, (2011).

[7] Rajan, B. K., Anto, N., Jose, S., "Fusion of iris & fingerprint biometrics for gender classification using neural network", 2nd International Conference on Current Trends in Engineering and Technology (ICCTET), 216-221, (2014).

[8] Rajesh, D. G., Punithavalli, M., "Wavelets and Gaussian mixture model approach for gender classification using fingerprints", 2nd International Conference on Current Trends in Engineering and Technology (ICCTET), 522-525, (2014).

[9] Shinde, M. K., Annadate, S. A., "Analysis of Fingerprint Image for Gender Classification or Identification: Using Wavelet Transform and Singular Value Decomposition", International Conference on Computing Communication Control and Automation (ICCUBEA), 650-654, (2015).

[10] Tarare, S., Anjikar, A., Turkar, H., "Fingerprint Based Gender Classification Using DWT Transform", International Conference on Computing Communication Control and Automation (ICCUBEA), 689-693, (2015).

[11] Kapoor, N., Badiye, A., "Sex differences in the thumbprint ridge density in a central Indian population", Egyptian Journal of Forensic Sciences, 5: 23–29, (2015).

[12] Ceyhan, E. B., "Parmak izinden cinsiyet tanıyan zeki sistem", Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2012.

[13] Internet: Biosecure Veritabanı, http://biosecure.it-sudparis.eu/AB/index.php?option=com_content&view=article&id=25&Itemid=31 , Erişim Tarihi: 30.08.2015.

Şeref SAĞIROĞLU, Gazi Üniversitesi Fen Bilimleri Enstitüsü Müdürü ve Bilgisayar Mühendisliği Bölüm Başkanıdır.

Eyüp Burak CEYHAN, Gazi Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümünde araştırma görevlisidir.

CYBER SECURITY AWARENESS OF ENGINEERING STUDENTS: A QUALITATIVE ANALYSIS ON COMPUTER & MECHATRONIC DEPARTMENTS

Hasan TINMAZ and Mehmet Ali BARIŞKAN

Assist. Prof. Dr. Hasan TINMAZ. Istanbul Gelisim University, Faculty of Engineering and Architecture, Department of Computer Engineering. Cihangir mah. Şehit Jandarma Komando Er Hakan Öner Sk. No:1 Avcılar / İstanbul / Turkey (corresponding author to provide phone: +90 5324159940; fax: +90 2124227401; e-mail: htinmaz@gelisim.edu.tr).
Res. Assist. Mehmet Ali BARIŞKAN. Istanbul Gelisim University, Faculty of Engineering and Architecture, Department of Computer Engineering. Cihangir mah. Şehit Jandarma Komando Er Hakan Öner Sk. No:1 Avcılar / İstanbul / Turkey (e-mail: mabariskan@gelisim.edu.tr).

Abstract — Furnishing university students with the necessary Information and Communication Technologies (ICTs) has been an indispensable instructional activity for the last decade. When it comes to the engineers, who are the technological stakeholders and decision makers, these instructional activities become even more important. This article discusses how computer (n=6) and mechatronic (n=6) engineering students in their fourth year perceive cyber security phenomenon after finishing a cyber security course. With a qualitative interviewing technique and analysis, it was revealed that students were satisfied with the course in terms of cognitive and affective acquisitions. As a result, students seemed not highly aware of adverse effects of lacking cyber security precautions. Moreover, even though mechatronic engineering students are in ICTs field, they seemed to have less concerns and interest related to cyber security. The study enlightens how cyber security courses should be instructionally designed and implemented.

Index Terms — Cyber Security, Higher Education, Security Teaching, Security Awareness

I. INTRODUCTION

Information and Communication Technologies (ICTs) are central elements of our daily lives. ICTs (especially the Internet) have brought many advantages to humanity while creating new concerns such as security in macro (countrywide national security) and micro (personal use) levels. It was predicted that 70% of world population would be using the Internet in 2015 [1] which also means that 70% of the entire world population might face these concerns in any second. Cyber security is one of the major concerns of this century. According Merriam Webster dictionary, cyber security means "measures taken to protect a computer or computer system (as on the Internet) against unauthorized access or attack" [2]. In that sense, cybersecurity covers different tools, processes and actions deliberately developed for shielding computers, software, data and networks from attack, illegal access and damage. Generally, when the ICTs stakeholders talk about security, they highly refer to cyber security. As being that significant in personal and professional lives, knowing and taking actions about cyber security are obligatory for all ICT users.

According to Norton, as being a computer security firm, the cost of security investment was $100 billion in 2012 [1]. This cost increased $400 billion in 2014 where it was also estimated that cybercrime costs reflecting on the global economy would increase that cost every single year [3]. Moreover, in parallel to the increase in number of mobile devices and smart device users are generally unaware that they are under threat from these devices. As a result, the investments, need of cyber security knowledge and awareness toward cyber security are getting higher and higher. For instance; while downloading and/or installing apps from unknown providers, users make themselves open to attacks which deactivate security procedures setup by manufacturers [4].

A well known cyber security company, Cyren, released "2015 Cyberthreat Yearbook Report" in early March 2015, which compared threats with year of 2014. According to the report use of malicious software increased by 50%, "phishing" e-mails 233%, malware affecting Android-based mobile systems increased by 61%, in comparison to 2014 [5]. Moreover, Turkish Data Security Association [Bilgi Güvenliği Derneği] added that cyber security threats are not only resulting from technological issues, but also from human aspects requiring awareness toward cyber security related issues [6].

As a consequence of its significance, cyber security and its training are getting widespread both in education and business sectors. It is expected that engineers who are the planners, adapters, users and multipliers of ICTs, play an important role regarding with cyber security issues. Therefore, engineers should attend cyber security courses to develop themselves in both knowledge and awareness level. In these courses, the difficult part is to make students understand how important cyber security is so that it reflects their overall perception of cyber security in knowledge and implementation levels. Affective domain of cyber security (aka psychological approach of students on cyber security) is the most important part of it, as they will decide for security procedure in their prospective jobs.

Literature also recommends that cyber security teaching/learning is important not only for engineering students but also for non-engineering students such as psychology and economics [7]. While social media tools (especially, Facebook, Instagram and Twitter) are getting more users who are predominantly from the youngster and youth, cyber security gains are more important for students. Therefore, cyber security awareness must be built into all people's minds [8].

To sum up, awareness of cyber security issues is an initial and vital step for taking precautions against cybercrime. Therefore, this study focuses on computer and mechatronic engineering students' awareness issues regarding with the cyber security after they fulfilled the cyber security course in their faculty.

II. METHOD

A. Research Context

The "TSD411 Cyber Security" elective course was delivered in the "2014-2015 Fall Semester" for computer (n=17) and mechatronic (n=20) departments at a faculty of engineering and architecture at a private university for the senior (fourth year) students (total 37 students). The researchers collected qualitative data via interviews with purposeful sampling at the end of the semester, after the students learnt their overall grades from the course.

B. Research Methods & Analysis

This is a descriptive case study concentrating on revealing perceptions of university students on cyber security as being one of the most important concerns for current ICT world. The students were prospective computer or mechatronic engineers who will work in a key position for giving decisions on ICT work and who took a cyber security course just before this study. A qualitative method was used during the data collection, to obtain comprehensive information. By purposeful sampling technique (which is common in qualitative research [9]), six students from each department were selected for the study (35% of entire computer engineering students, 30% of entire mechatronic engineering students, 33% of entire students registered to course). Students were asked about their course grade; AA (n=1), BB (n=3), CB (n=2), CC (n=2), DC (n=1), and FF (n=2). According to student course grade distribution, it seems that sample represents the class population.

The researchers organized an interview schedule with the help of literature and their professional experiences. After scrutiny from their colleagues and Turkish language experts, the final draft was piloted on a student who was excluded from results of this study. The final version of the interview schedule was applied to the selected students who were informed that their grades would not be affected by their answers. Moreover, none of the researchers were the instructor or any relationship with the course. Therefore, the students were assumed to provide information with less pressure. Interviews were conducted individually at the researchers' office on an appointment based plan.

The data was collected by means of semi-structured interviews with the researchers. The interviews were transcribed and coded. Then, the transcriptions were analyzed by means of content analysis. The researchers conducted an analysis of single transcriptions to create a set of categories and subcategories. Themes derived from each participant's responses were shared and discussed between the researchers.

III. RESULTS

A. Change in knowledge

The first questions were about to reveal how students perceive change in their knowledge on CS (Cyber Security) after they had finished a course on CS. From Table 1, it seems that course did not change their CS knowledge much.

Computer Engineering	n	Mechatronic Engineering	n	Total	n
Basic to Intermediate	3	Zero to Zero	1	Zero to Zero	1
Zero to Basic	2	Zero to Basic	3	Zero to Basic	5
Basic to Basic	1	Zero to Intermediate	1	Zero to Intermediate	1
		Basic to Basic	1	Basic to Basic	2
				Basic to Intermediate	3

Table I - Change in knowledge

B. General evaluation of the Cyber Security course

The students were asked evaluate CS course. Nearly all students did not want to make a general comment on the course. 1 student remarked that it was an essential course for his professional development. The students were asked the difficulty level of the course. Table 2 shows that students perceived CS course as a difficult one.

Computer Engineering	n	Mechatronic Engineering	n	Total	n
No Comment	1	Easy	2	No Comment	1
Hard	3	Hard	2	Easy	2
Moderate	1	Very Hard	2	Moderate	1
				Hard	5
				Very Hard	2

Table II - Course difficulty level

Afterwards, the students pointed the underlying reasons of this difficulty. It seems that students had problems with lack of preliminary knowledge, lecturer, and lack of motivation (Table 3).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
No Comment	2	Not necessary	1	No Comment	2
Introductory topics	1	Requirement of prerequisite knowledge	1	Introductory topics	1
Lecturer could not transfer the knowledge	1	Lecturer's communication skills should be better	1	Lecturer could not transfer the knowledge	1
Lecturer was not dominant in the course	1	Requirement of preparation before the course	1	Lecturer was not dominant in the course	1
Similar with other lessons	1	Students' willingness	1	Similar with other lessons	1
		Use of many terms in English	1	Not necessary	1
				Requirement of prerequisite knowledge	1
				Lecturer's communication skills should be better	1
				Requirement of preparation before the course	1
				Students' willingness	1
				Use of many terms in English	1

Table III - Reasons of difficulties

C. Necessity of Cyber Security course

When the necessity of having such a course was asked, computer-engineering students strongly agreed whereas mechatronic students had suspicious thoughts on the issue (Table 4).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
Yes	4	Yes	2	Yes	6
Absolutely	2	No	2	No	2
		No Comment	1	Absolutely	2
		A little	1	A little	1
				No Comment	1

Table IV - Necessity of Course

Furthermore, students were asked about why they feel such a necessity of attending such a course. Some students argued that using a computer or being computer engineering cannot be realized without CS awareness. Mechatronic engineering students did not feel such a necessity to join a course on CS (Table 5).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
Computer security is in all areas of life	1	No Comment	4	Computer security is in all areas of life	1
Necessary	1	Not Necessary	1	Necessary	1
Cannot think Computer Engineering without security	1	It is a necessity to store personal information safely	1	Cannot think Computer Engineering without security	1
No Comment	3			No Comment	7
				Not Necessary	1
				It is a necessity to store personal information safely	1

Table V - Reason of Necessity

Five computer engineering students declared that the CS course must be mandatory where only 1 computer-engineering student noted that CS course should stay as an elective course as it used to be. On the other hand, the situation is totally opposite for mechatronic engineering students where 5 students support elective course idea and 1 only agreed to have CS course as a mandatory in department curriculum.

D. Learning about Cyber Security

Students were questioned to what extent they believe that they learned fundamentals of CS via this course. Unfortunately, half of students did not believe that they had furnished themselves with basics of CS (Table 6).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
Not Really	1	Very Little	1	Very Little	1
No	2	Not completely	1	A little	1
Basics, yes	3	Basics, yes	2	No	2
A little	1	Yes	2	Not Really	1
				Not completely	1
				Basics, yes	4
				Yes	2

Table VI - Knowledge level at the end of course

For detailed information, students were asked to unfold their reasons for not-comprehending the fundamentals. Most students did not want to make any comment on that issue. Nonetheless, some students noted problems regarding to technological problems, disliking the course and necessity of having extracurricular study on CS (Table 7).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
No Comment	5	No Comment	2	No Comment	6
Software & Hardware Problems	1	Course apathy	1	Course apathy	1
		Students required to make additional research & self-development	2	Software & Hardware Problems	1
				Students required to make additional research & self-development	2

Table VII - Reasons of this knowledge levels

Furthermore, the students were asked if their expected topics in CS course were fulfilled or not. Most of the students pointed that many topics they expected to be covered in the CS course were presented during the semester (Table 8).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
Yes	3	Yes	3	Yes	6
Not Completely	2	Nearly all of them	1	Nearly all of them	1
Not very much	1	Generally	1	Generally	1
		No Comment	1	Not Completely	2
				Not very much	1
				No Comment	1

Table VIII - Expected topics fulfillment

The students explained why they thought that their expectations in relation to CS course topics were not fulfilled. Students were complaining about departmental differences and pointless topics within the course (Table 9).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
No Comment	4	No Comment	3	No Comment	7
Requires too much research beforehand	1	It is not correct to offer this course to Mechatronic Engineering Department	1	It is not correct to offer this course to Mechatronic Engineering Department	1
Too much unrelated subjects	1	Lecturer helped enough	1	Requires too much research beforehand	1
		More than enough	1	Too much unrelated subjects	1
				More than enough	1
				Lecturer helped enough	1

Table IX - Reasons of expected topics fulfillment

As a follow-up question, students listed the topics that they wanted to see within CS course. Only one student added that “attacks” should be in CS course curriculum whereas 11 students did not have any comment.

E. Teaching about Cyber Security

The students stated whether or not the CS course should be taught as theory or as implementation based on instruction. Eight of the students (five computer engineering and three mechatronic engineering) remarked that CS course should be implementation based. Four students (one computer engineering and three mechatronic engineering) emphasized that CS course should be both theory and implementation based together in balance.

Additionally, students added their comments on what an implementation based CS course should include. Students want to apply attack and counter attack simulation based on software (Table 10).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
Attack examples	3	Attack examples	3	Attack examples	6
Security organization against an attack	1	Attack & counter attack programs	3	Attack & counter attack programs	3
Maintaining a Firewall	1			Security organization against an attack	1
Simulative attacks	1			Maintaining a Firewall	1
				Simulative attacks	1

Table X - Prospective topics students want to see in the course

Moreover, the students made comments on instructional materials used in CS course. Less than half of the students were satisfied with the instructional materials used (Table 11).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
No Comment	1	No Comment	1	No Comment	2
No	3	No	1	No	4
Not Bad	1	Yes	4	Not Bad	1
Yes	1			Yes	5

Table XI - Level of materials satisfaction

Students mostly complained about lack of implementation sessions in a computer lab and the language of the materials, which was mostly in English. Similarly, students adversely expressed their attitudes towards the instructor and the course which affected their standpoint toward instructional materials (Table 12).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
Computer Labs had problems including administration rights	3	Computer Labs had problems including administration rights	1	Computer Labs had problems including administration rights	4
Materials were mostly in English (must be Turkish)	2	Materials were mostly in English (must be Turkish)	1	Materials were mostly in English (must be Turkish)	3
Lecturer could not use materials properly	1	I liked the exercises	2	I liked the exercises	2
Not much documents	1	Apathy toward the course	1	Not much documents	1
				Lecturer could not use materials properly	1
				Apathy toward the course	1

Table XII - Reasons of material satisfaction

In addition to the instructional materials, the students reflected on "assessment criteria" of the CS course. Initially, they did not want to make any comments, although they were already graded at the end of the semester. Some students doubted the exam procedures and instructor's objectivity (Table 13).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
No Comment	4	No Comment	3	No Comment	8
I doubt about the quality of exams	1	Students were affecting instructor's grades.	1	I doubt about the quality of exams	1
Assessment criteria were not equal for every student	1	Very Good	1	Assessment criteria were not equal for every student	1
		Correct at both teaching procedure & assessment.	1	Students were affecting instructor's grades.	1
				Very Good	1
				Correct at both teaching procedure & assessment.	1

Table XIII - Assessment criteria

Students were asked to assume the role of instructor and to tell what kind of assessment activities they would realize in the CS course. Even though minority of the students still would like to have a paper-based exam, the majority notes that they would apply computers based exams in the CS course (Table 14).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
I would care how much students learned from this course than exam results	1	I would grade students based on lab activities	2	I would grade students based on lab activities	2
Mainly, I would make exams in multiple choice formats	1	No Exam at all	2	No Exam at all	2
I would make application based activities and grade them, and additionally I would make an exam	1	No Comment	3	I would care how much students learned from this course than exam results	1
I would give better grades who finalize the activities/attacks first	1			Mainly, I would make exams in multiple choice formats	1
Paper based exam	1			I would make application based activities and grade them, and additionally I would make an exam	1
No Comment	1			I would give better grades who finalize the activities/ attacks first	1
				Paper based exam	1
				No Comment	4

Table XIV - How students would do assessments

The students stated their opinions on how to offer such a course, if they were the course instructor. This is one of the questions that students shared many different instructional ideas (Table 15).

Computer Engineering	n	Mechatronic Engineering	n	Total	n
Firstly, I would give fundamental terms	2	Same as the professor we got the course	2	Firstly, I would give fundamental terms	2
First theory and then application / implementation	1	More theory	1	I'd divide class into sections	2
Firstly, I would measure the level of students	1	First theory and then application / implementation	1	Same as the professor we got the course	2
I would show them how to attack & defense	1	I never want to deliver this course	1	First theory and then application/ implementation	1
I would show them how to install & use the related software	1	I'd divide class to sections	1	Firstly, I would measure the level of students	1
I would select one type attack & focus only on that one	1			I would show them how to attack & defense	1
I'd divide class into sections	1			I would show them how to install & use the related software	1
I'd give research subjects to students	1			I would select one type attack & focus only on that one	1
				More theory	1
				I never want to deliver this course	1
				I'd learn myself first and then offer	1

Table XV - How would students offer the lesson

Lastly, the students (n=10) pointed that universities could establish departments within a faculty or vocational school just focusing on cyber security.

IV. CONCLUSION AND RECOMMENDATIONS

This study reflects on how prospective engineers, who will be the technology stakeholders, perceive cyber security concepts and how a cyber security course could contribute to that perception change. Computer and mechatronic departments were deliberately selected to show that security is an essential element for their work. The results showed that mechatronic students underestimate importance of cyber security by calling it "not an interest for their future work". The cyber security course was pointed as a "not beneficial or not interesting course" for their departmental curriculum. Therefore, there needs to be an awareness movement for mechatronic departments which is directly a part ICTs. Additionally, the cyber security course should have a different section for them emphasizing overtly related cases or examples from implementation of mechatronic engineering. Computer engineering students gave the impression that they had realized the importance of the cyber security for their personal (especially on social media) and professional lives by even highlighting that a computer engineer cannot survive without cyber security knowledge in his/her life.

The results showed that students do not feel comfortable about their cyber security knowledge as a result of lack of attack/counterattack based implementations. It was strongly highlighted that cyber security courses must have more implementation than theory which would be realized in computer labs than regular classes. Moreover, the students noted that cyber security knowledge evaluation must also stem from real case applications, not paper based exams.

Students urged that learning or applying cyber security, students must bring prerequisite knowledge to the class. They even added that the instructors should check that prerequisite knowledge at the beginning of the semester. Filling the gap with what the students should know at the beginning of the semester will increase the motivation and willingness toward the cyber security topics.

Students complained about the course materials, since they are dominantly in English. Therefore, academicians or implementers of cyber security must create more instructional materials (including books and lecture notes) in Turkish.

This study includes small sample of students contributed and qualitative as a method. Qualitative method based studies frequently investigate the research problem in-depth. Therefore, the study results might not generalize to other cases. Based on these findings, future research should be conducted either with a quantitative or mixed research method approaches.

REFERENCES

[1] K. W. Brenda, "The role of psychology in enhancing cybersecurity" Cyberpsychology, Behavior, And Social Networking, vol. 17, no.3, pp. 131-132, 2014.

[2] Merriam Webster Dictionary. (2015, August 10). [Online]. Available: http://www.merriam-webster.com/dictionary/cybersecurity

[3] D. Zureich and W. Graebe, "Cybersecurity: The continuing evolution of insurance and ethics" Defense Counsel J., vol. 82, no.2, pp. 192-198, April 2015.

[4] J. Imgraben, A.Engelbrecht and K. R. Choo, "Always connected, but are smart mobile users getting more security savvy? A survey of smart mobile device users", Behaviour & Information Technology, vol. 33, no. 12, pp. 1347-1360, June 2014.

[5] Cyren, Inc. 2015 Cyber threat Yearbook., (2015, August 10). [Online]. Available: https://www.cyren.com/tl_files/downloads/CYREN_2015_CyberThreat_Yearbook.pdf

[6] Turkish Data Security Association [Bilgi Güvenliği Derneği]. "Cyber Security Report: First Quarter, 2015". (2015, August 10). [Online]. Available: http://www.bilgiguvenligi.org.tr

[7] J. Cano, R. Hernández, and S. Ros "Bringing an engineering lab into social sciences: didactic approach and an experiential evaluation", IEEE Communications Magazine, vol. 52, no. 12, pp. 101-107, December 2014.

[8] B. Simpson and M. Murphy, "Cyber-privacy or cyber-surveillance? Legal responses to fear in cyberspace", Journal Information and Communications Technology Law, vol. 23, no.3, pp. 189-191, October 2014.

[9] M.Q. Patton, Qualitative Evaluation and Research Methods, Newbury Park, CA: Sage, 1990.

Dr. Hasan Tınmaz, Assist. Prof. received his bachelor's degree from the Department of Computer Education, Faculty of Education, from Middle East Technical University in 2001. He completed his M.Sc. degree in Curriculum and Instruction Program, from the Department of Educational Sciences at METU (2004). He received his Ph.D. from Computer Education and Instructional Technology at METU (2011). He is now is an assistant professor, in the Faculty of Engineering and Architecture, Department of Computer Engineering, at Istanbul Gelisim University. His research focuses on Web 2.0/Web 3.0 technologies, social media, instructional design, and human & computer interaction.

Mehmet Ali Barışkan, Res. Assist. received his bachelor's degree from the Department of Computer Engineering in English, Faculty of Engineering and Architecture, Istanbul Aydin University in 2013. He is currently a graduate student in the Computer Engineering Master Program at the Science Institute of Istanbul University. He is also a research assistant in the Computer Engineering Department, in the Faculty of Engineering and Architecture, Istanbul Gelisim University. His research focuses on computer security, cyber security, data recovery and reverse engineering.

Improved Contract Signing Protocol Based on Certificateless Hybrid Verifiably Encrypted Signature Scheme

Ömer Sever, Ersan Akyıldız
Institute of Applied Mathematics, Middle East Technical University, Ankara, Turkey
e-mail: severomer@gmail.com, ersan@metu.edu.tr

***Abstract*—Contract signing protocols are being widely used over digital environment and treated as an application of non-repudiation protocols. As a kind of non-repudiation protocols, the most important property of contract signing protocols is fairness. In this paper we analyze a recent contract signing protocol based on hybrid verifiably encrypted signature scheme (HVESS) and show a common attack. Further we propose improvement to the protocol, adaptation of certificateless public key cryptography to HVESS (CL-HVESS) and expansion of CL-HVESS to Type-III pairings.**

***Keywords*—Contract Signing protocols, Non-Repudiation protocols, Pairing Based Cryptography, Certificateless Cryptography, verifiably encrypted signature scheme**

I. Introduction

Non-repudiation protocols are used for exchange of information with evidence of non-repudiation. Application of non-repudiation protocols are spreaded over Certfied E-mail, Electronic Contract Signing, e-commerce and electronic payment. E-contract is any kind of contract formed in the course of e-commerce by the interaction of two or more individuals using electronic means, such as e-mail, the interaction of an individual with an electronic agent, such as a computer program, or the interaction of at least two electronic agents that are programmed to recognize the existence of a contract [7]. There are many examples of e-contract platforms over the Internet, some of them serve for general purpose contracts [8], [9] and some of them serve for specific purposes like real estates [10] or like telecommunication suppliers [11].

II. General Description

A. Non-Repudiation and Contract Signing Protocols

Non-repudiation is defined as a security service by which the entities involved in a communication can not deny having participated, specifically, the sender can not deny having sent a message and the receiver can not deny having received a message [1].
Non-repudiation is primarily depending on asymmetric cryptography specifically to signatures which are accepted as evidences. Regarding how used in a protocol, evidence of origin supplies Non-Repudiation of Origin (NRO) and evidence of receipt supplies Non-Repudiation of Receipt(NRR).

Non-repudiation protocols can satisfy various properties in different ways like:

- Fairness: Strong, weak, light
- Non-Repudiation: NRO, NRR, NRS, NRD
- State storage: Statefull, stateless
- Timeliness: Synchronous, Asynchronous
- TTP Inclusion: In-line, On-line, Off-line, Probabilistic

These properties and non-repudiation protocols have been studied in [2], [3], [21] [25] and [5].

As a kind of non-repudiation protocol, contract signing share similar properties with other protocols. The goal of contract signing protocols is exchange of evidence of non-repudiation not the message itself. Differing from certified e-mail or fair exchange is that obtaining message content is not important but exchanging signed message/contract fairly is the main goal of the contract signing protocol.

B. Pairing Based Cryptography

Public key cryptography (PKC) is generally based on certificates binding identities with public keys which are approved by Certificate Authorities. Differing from classical PKC, in ID-Based Cryptography public keys are dependant on user identities and/or identifiers. This difference brings advantages and disadvantages together as discussed in [14]. The advantages of ID-Based Cryptography are mainly achieving different encryption and signature schemes like ID-Based encryption [15], blind [16], short [17], ring [18] and verifiably encrypted [19], [26] signatures which are summarized in [4]. The disadvantage of ID-Based cryptography is if the public key is dependant only on identity of a user, key generator knows the private keys of users when generation. In this paper we adapted the certificateless public key cryptography described in [24] to the hybrid verifiably encrypted signature scheme [26].

1) Bilinear Pairings: Pairings in elliptic curve cryptography are functions which map a pair of elliptic curve points to an element of the multiplicative group of a finite field. Below is the simple definition of a bilinear pairing , more information on pairings like Weil or Tate pairings, divisiors

and curve selection can be found in [6] as a summary and in [27] in more details.

Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be additive abelian group of order q and $\mathbb{G}_3$ be multiplicative group of order q, a pairing is a function

$$e : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_3 \quad (1)$$

e is suitable for cryptographic schemes when it is an efficiently computable bilinear pairing which satisfies the following properties:

a) e is bilinear: For all $P, S \in \mathbb{G}_1$ and $Q, T \in \mathbb{G}_2$ we have $e(P+S, Q) = e(P, Q)e(S, Q)$ and $e(P, Q+T) = e(P, Q)e(P, T)$
b) e is non-degenerate: For all $P \in \mathbb{G}_1$, with $P \neq 0$ there is some $Q \in \mathbb{G}_2$ such that $e(P, Q) \neq 1$ and for all $Q \in \mathbb{G}_2$, with $Q \neq 0$ there is some $P \in \mathbb{G}_1$ such that $e(P, Q) \neq 1$

Consequtive properties of bilinearity are:

- $e(P, 0) = e(0, Q) = 1$
- $e(-P, Q) = e(P, Q)^{-1} = e(P, -Q)$
- $e([a]P, Q) = e(P, Q)^a = e(P, [a]Q)$ for all $a \in \mathbb{Z}$

In Section IV we will use this notation to expand HVESS to Type-III pairings.

2) Modified Pairings: In Section III we will use Type I [28] supersingular curves for pairing instantiation in which $\mathbb{G}_1 = \mathbb{G}_2$, to show how we adapted certificateless ID-Based PKC [24] to HVESS. In this type $\mathbb{G}_1$ is a subgroup of $E(\mathbb{F}_q)$. There is a distortion map ψ which maps $\mathbb{G}_1$ into $E(\mathbb{F}_{q^k})$ and the modified pairing $\hat{e}(P, Q) : \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_3$ for $P, Q \in \mathbb{G}_1$ is defined by:
$\hat{e}(P, Q) = e(P, \psi(Q))$ as shown in section X in [27].

III. Adaptation of Certificateless Public Key Cryptography to HVESS

ID-Based signature verification and encryption schemes use publicly known variable such as identity or e-mail of a user to derive public key without any key distribution for public keys. For signing and decrypting user contacts to a Private Key Generator (PKG, CA etc.) to derive the private key which is dependant on the identity and master key of the PKG.

This scheme has some disadvantages stated in [4]

- The PKG can calculate users private keys which is a problem for confidentiality in non-rep protocols
- User has to authenticate himself to PKG
- PKG needs a secure channel to send users private key
- User has to publish PKG's public parameters

Chen and Gu have developed and used HVESS [26] which is a pure ID based scheme. To eliminate some of the above mentioned disadvantages, we adapted Riyami and Paterson's [24] certificateless public cryptography scheme to HVESS and call the adapted scheme shortly as CL-HVESS. Most of the parts of our scheme is similar to the original work [26] naturally.

Setup : Let $\mathbb{G}_1$ be additive group of prime order q and $\mathbb{G}_3$ be multiplicative group of prime order q. Choose an arbitrary generator $P \in \mathbb{G}_1$, a random secret PKG master key $s \in \mathbb{Z}_q^*$ and a random secret adjudicator master key $s_T \in \mathbb{Z}_q^*$. Set $P_{pub} = [s]P$ and $P_{adj} = [s_T]P$ choose cryptographic hash functions $H_1 : \{0,1\}^* \to \mathbb{G}_1$ and $H_2 : \{0,1\}^*$ X $\mathbb{G}_1 \to \mathbb{Z}_q^*$. Publish the system parameters $(\mathbb{G}_1, \mathbb{G}_3, q, \hat{e}, P, P_{pub}, P_{adj}, H_1, H_2)$

Extract : Public and private key pair for user ID is computed as follows:

- TTP or PKG computes $P_{pub} = [s]P$ and $[s]H_1(ID)$ as the partial private key then send to user ID.
- User ID computes $P_{pub\_ID} = [X_{ID}][s]P$ and $R_{pub\_ID} = [X_{ID}]P$ as public keys then computes $d_{ID} = [X_{ID}][s]H_1(ID)$ as private key.

Sign : Given a private key d_{ID} and a message m, pick a random $r \in \mathbb{Z}_q^*$, compute $U = [r]P, h = H_2(m, U), V = [r]H_1(ID) + [h]d_{ID}$ and output a signature (U, V).

Verify : Given a signature (U, V), of an identity ID and public keys P_{pub_ID}, R_{pub_ID} first check certificateless public keys as $\hat{e}(R_{pub\_ID}, P_pub) \stackrel{?}{=} \hat{e}(P_{pub\_ID}, P)$ then compute $h = H_2(m, U)$, and accept the signature and return 1 if and only if $\hat{e}(P, V) = \hat{e}(U + [h]P_pub\_ID, H_1(ID)$. The proof of verification for a valid signature (U, V) is as follows;
$\hat{e}(P, V) = \hat{e}(P, [r]H_1(ID) + [h]d_{ID})$
$= \hat{e}(P, [r]H_1(ID) + [h][X_{ID}][s]H_1(ID))$
$= \hat{e}(P, ([r] + [h][X_{ID}][s])H_1(ID))$
$= \hat{e}(([r] + [h][X_{ID}][s])P, H_1(ID))$
$= \hat{e}([r]P + [h][X_{ID}][s]P, H_1(ID))$
$= \hat{e}(U + [h]P_{pub\_ID}, H_1(ID))$

SignVE : Given a private key d_{ID} and a message m, pick randomly $r_1, r_2 \in \mathbb{Z}_q^*$, compute $U_1 = [r_1]P, U_2 = [r_2]P, h = H_2(m, U_1), V = [r_1]H_1(ID) + [h]d_{ID} + [r_2]P_{adj}$, and output a verifiably encrypted signature (U_1, U_2, V)

VerifyVE : Given a verifiably encrypted signature (U_1, U_2, V) of a user ID for a message m, compute $h = H_2(m, U_1)$, accept the signature if and only if $\hat{e}(P, V) = \hat{e}(U_1 + [h]P_pub\_ID, H_1(ID).\hat{e}(U_2, P_{adj})$. The proof of verification for a valid verfiably encrypted signature (U_1, U_2, V) is as follows;
$\hat{e}(P, V) = \hat{e}(P, [r_1]H_1(ID) + [h]d_{ID} + [r_2]P_{adj})$
$= \hat{e}(P, [r_1]H_1(ID) + [h][X_{ID}][s]H_1(ID)).\hat{e}(P, [r_2][s_T]P)$
$= \hat{e}(P, ([r_1] + [h][X_{ID}][s])H_1(ID)).\hat{e}(U_2, P_{adj})$
$= \hat{e}(([r_1] + [h][X_{ID}][s])P, H_1(ID)).\hat{e}(U_2, P_{adj})$
$= \hat{e}([r_1]P + [h][X_{ID}][s]P, H_1(ID)).\hat{e}(U_2, P_{adj})$
$= \hat{e}(U_1 + [h]P_pub\_ID, H_1(ID)).\hat{e}(U_2, P_{adj})$

Adjudication : Given the adjudicator's private key s_T and a valid verifiably encrypted signature (U_1, U_2, V) for a message m, compute $V_1 = V - [s_T]U_2$ and output the original signature (U_1, V_1) Validation requires first verification of verifiably encrypted signature (U_1, U_2, V) and then verification of adjudicated verifiably encrypted signature (U_1, V_1) as an original signature. First part is same procedure as $VerifyVE(U_1, U_2, V)$, for the validation of second part: $V_1 = V - [s_T]U_2 = V - [s_T][r_2]P = V - [r_2]P_{adj}$ $= [r_1]H_1(ID) + [h]d_{ID} + [r_2]P_{adj} - [r_2]P_{adj} = [r_1]H_1(ID) + [h]d_{ID}$ so $\hat{e}(P, V_1) = \hat{e}(P, [r_1]H_1(ID) + [h]d_{ID}) = \hat{e}(U_1 + [h]P_pub\_ID, H_1(ID)$

IV. Expansion of CL-HVESS to Type-III Pairings

In the previous section we have adapted Certificateless PKC to HVESS on Type-I pairings in which $\mathbb{G}_1 = \mathbb{G}_2$. Since Type-I pairings are susceptible to recent quasi-polynomial attacks [30], [31], here we expanded CL-HVESS to Type-III pairings. Type-II pairings are not suitable for CL-HVESS because there is not efficiently computable hash function to $\mathbb{G}_2$.

Setup : Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be additive abelian group of order q and $\mathbb{G}_3$ be multiplicative group of order q. Choose arbitrary generators $P \in \mathbb{G}_1$, $Q \in \mathbb{G}_2$ a random secret PKG master key $s \in \mathbb{Z}_q^*$ and a random secret adjudicator master key $s_T \in \mathbb{Z}_q^*$. Set $P_{pub} = [s]P$, $Q_{pub} = [s]Q$, $P_{adj} = [s_T]P$ and $Q_{adj} = [s_T]Q$ choose cryptographic hash functions $H_1 : \{0,1\}^* \rightarrow \mathbb{G}_1$, $H_2 : \{0,1\}^* \rightarrow \mathbb{G}_2$ and $H_3 : \{0,1\}^*$ X $\mathbb{G}_1 \rightarrow \mathbb{Z}_q^*$. Publish the system parameters $(\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3, q, e, P, Q, P_{pub}, Q_{pub}, P_{adj}, Q_{adj}, H_1, H_2, H_3)$

Extract : Public and private key pair for user ID is computed as follows:

- TTP or PKG computes $P_{pub} = [s]P, Q_{pub} = [s]Q$ and $[s]H_1(ID), [s]H_2(ID)$ as the partial private keys then send to user ID.
- User ID computes $P_{pub\_ID} = [X_{ID}][s]P, Q_{pub\_ID} = [X_{ID}][s]Q$ and $R\_P_{pub\_ID} = [X_{ID}]P, R\_Q_{pub\_ID} = [X_{ID}]Q$ as public keys then computes $d\_P_{ID} = [X_{ID}][s]H_1(ID), d\_Q_{ID} = [X_{ID}][s]H_2(ID)$ as private keys.

Sign : Given a private key d_Q_{ID} and a message m, pick a random $r \in \mathbb{Z}_q^*$, compute $U = [r]P, h = H_3(m, U), V = [r]H_2(ID)+[h]d\_Q_{ID}$ and output a signature (U, V).

Verify : Given a signature (U, V), of an identity ID and public keys $P_{pub\_ID}, R\_P_{pub\_ID}$ first check certificateless public keys as $\hat{e}(R\_P_{pub\_ID}, Q_p ub) \stackrel{?}{=} \hat{e}(P_{pub\_ID}, Q)$ then compute $h = H_3(m, U)$, and accept the signature and return 1 if and only if $\hat{e}(P, V) = \hat{e}(U + [h]P_p ub\_ID, H_2(ID)$. The proof of verification for a valid signature (U, V) is as follows;
$\hat{e}(P, V) = \hat{e}(P, [r]H_2(ID) + [h]d\_Q_{ID})$
$= \hat{e}(P, [r]H_2(ID) + [h][X_{ID}][s]H_2(ID))$
$= \hat{e}(P, ([r] + [h][X_{ID}][s])H_2(ID))$
$= \hat{e}(([r] + [h][X_{ID}][s])P, H_2(ID))$
$= \hat{e}([r]P + [h][X_{ID}][s]P, H_2(ID))$
$= \hat{e}(U + [h]P_{pub\_ID}, H_2(ID))$

SignVE : Given a private key d_Q_{ID} and a message m, pick randomly $r_1, r_2 \in \mathbb{Z}_q^*$, compute $U_1 = [r_1]P, U_2 = [r_2]Q, h = H_3(m, U_1), V = [r_1]H_2(ID) + [h]d\_Q_{ID} + [r_2]Q_{adj}$, and output a verifiably encrypted signature (U_1, U_2, V)

VerifyVE : Given a verifiably encrypted signature (U_1, U_2, V) of a user ID for a message m, compute $h = H_3(m, U_1)$, accept the signature if and only if $\hat{e}(P, V) = \hat{e}(U_1 + [h]P_p ub\_ID, H_2(ID).\hat{e}(P_{adj}, U_2)$. The proof of verification for a valid verfiably encrypted signature (U_1, U_2, V) is as follows;
$\hat{e}(P, V) = \hat{e}(P, [r_1]H_2(ID) + [h]d\_Q_{ID} + [r_2]Q_{adj})$
$= \hat{e}(P, [r_1]H_2(ID) + [h][X_{ID}][s]H_2(ID)).\hat{e}(P, [r_2][s_T]Q)$
$= \hat{e}(P, ([r_1] + [h][X_{ID}][s])H_2(ID)).\hat{e}([s_T]P, [r_2]Q)$
$= \hat{e}(([r_1] + [h][X_{ID}][s])P, H_2(ID)).\hat{e}(P_{adj}, U_2)$
$= \hat{e}([r_1]P + [h][X_{ID}][s]P, H_2(ID)).\hat{e}(P_{adj}, U_2)$
$= \hat{e}(U_1 + [h]P_p ub\_ID, H_2(ID)).\hat{e}(P_{adj}, U_2)$

Adjudication : Given the adjudicator's private key s_T and a valid verifiably encrypted signature (U_1, U_2, V) for a message m, compute $V_1 = V - [s_T]U_2$ and output the original signature (U_1, V_1) Validation requires first verification of verifiably encrypted signature (U_1, U_2, V) and then verification of adjudicated verifiably encrypted signature (U_1, V_1) as an original signature. First part is same procedure as $VerifyVE(U_1, U_2, V)$, for the validation of second part: $V_1 = V - [s_T]U_2 = V - [s_T][r_2]Q = V - [r_2]Q_{adj} = [r_1]H_2(ID)+[h]d\_Q_{ID}+[r_2]Q_{adj}-[r_2]Q_{adj} = [r_1]H_2(ID)+[h]d\_Q_{ID}$ so $\hat{e}(P, V_1) = \hat{e}(P, [r_1]H_2(ID) + [h]d\_Q_{ID}) = \hat{e}(U_1 + [h]P_p ub\_ID, H_2(ID)$

V. Attack and Improvement to Fair Contract Signing Protocol

A. Attack to Contract Signing Protocol

Here we show a replay attack to Chen and Gu protocol [26], in which the responder site could get the adjudicated contract but the initiator A, can not get the contract signed by the intended responder B, instead get the contract signed by a colluder C. The attack of the scenerio is figured in Fig.1 and then explained further below.

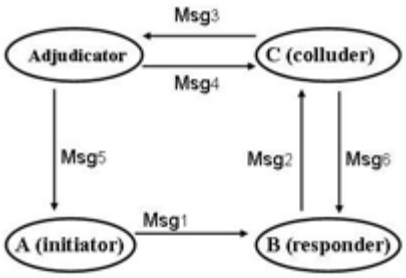


Fig 1. replay attack on protocol

Msg 1 $A \rightarrow B : ID_A, C, SignVE\{d_A, ID_A, C, P_{Adj}\}$

Msg 2 B colludes with C and sends verifiably encrypted signed contract to C
$B \rightarrow C : ID_A, C, SignVE\{d_A, ID_A, C, P_{Adj}\}$

Msg 3 C signs the contract by his private key and request from the adjudicator to resolve the dispute.
$C \rightarrow Adj : (ID_A, C, SignVE\{d_A, ID_A, C, P_{Adj}\})$
and $Sign(d_C, ID_C, C)$

Msg 4 After the adjudicator verifies the signed and verifiably encrypted signed contract, delivers the adjudicated contract to C.
$Adj \rightarrow C : Adjudication(SignVE\{d_A, ID_A, C, P_{Adj}\})$

Msg 5 After the adjudicator verifies the signed and verifiably encrypted signed contract, delivers the signed contract to A.
$Adj \rightarrow A : Sign(d_C, ID_C, C)$

Msg 6 Colluder C returns the adjudicated contract to B.
$C \rightarrow B : Adjudication(SignVE\{d_A, ID_A, C, P_{Adj}\})$

This contract signing protocol is based on HVESS as cryptographic signature scheme, which was proven as secure in [26]. Besides the cryptographic security, information sent in the protocol / signature scheme and control checks are also very important for a security protocol. In this attack we exploited a security flaw of missing information, namely identifier of responder, in the signature scheme. It may be claimed as the contract is suited for A and B but this would not be a formal security check for a protocol.

B. Improvement to Contract Signing Protocol

The improvement to the protocol is very easy as to include the identifier of responder to the signed message and check this before any response. For adding the CL-HVESS, we include the public keys of the sender to the message. The improved protocol is shown below; note that Signed or verifiably signed messages also include the original messages.

Msg 1 $A \rightarrow B : SignVE\{d_A, ID_A, ID_B, C, P_{pub\_A}, R\_P_{pub\_A}, P_{Adj}, Q_{Adj}\}$

Msg 2 $B \rightarrow A : Sign\{d_B, ID_B, ID_A, C, P_{pub\_B}, R\_P_{pub\_B}, P_{Adj}, Q_{Adj}\}$

Msg 3 $A \rightarrow B : Sign\{d_A, ID_A, ID_B, P_{pub\_A}, R\_P_{pub\_A}, Msg2\}$

C. Analysis of Protocol

Although there is not a formal security proof for CL-HVESS, we can make an informal comparison between original protocol and our work. When you use traditional ID-Based encryption and signature methods, as done in the original scheme, TTP can generate and escrow private keys of all users. But in certificateless scheme of [24] users can generate their own private keys. Also revocating a disclosed or lost private key in pure ID-Based crypto systems is difficult because you have to change the corresponding public key and so the ID of that user depends on. Using schemes of [24] TTP can not escrow keys but can revocate keys easily which is important for contract signing protocols depending on pairings. Addition to security analysis we can say the improved protocol is resistant to replay attacks. When we compare our adapted protocols with original version in view of efficiency, there is not so much difference between them. Both Type I and Type III versions of CL-HVESS have same calculations except setup phase which is done for only once. Below is the comparison of efficiency:

- **Sign** same as original; 3 scalar multiplication.
- **Verify** extra two pairings to check certificateless public keys; in total 4 pairings, 1 scalar multiplication.
- **SignVE** same as original; 5 scalar multiplication.
- **VerifyVE** extra two pairings to check certificateless public keys; in total 5 pairings, 1 scalar multiplication.
- **Adjudication** same as original; 1 scalar multiplication.

VI. Conclusion

We proposed adaptation of certificateless public key cryptography to hybrid verifiably encrypted signature scheme [26] which we call CL-HVESS. Adaptation of certificateless PKC prevents some problems of pure ID based schemes especially generation of user private keys by PKG. Then we expanded CL-HVESS to Type-III pairings to mitigate the risks of recent attacks on Type-I pairings. We also presented a replay attack to Chen and Gu protocol [26], in which the responder site could get the adjudicated contract but the initiator A, can not get the contract signed by the intended responder B, instead get the contract signed by a colluder C. Then propose an improvement to the protocol which is resistant to replay attacks and also included the CL-HVESS to the improved protocol. But notice that this attack and CL-HVESS are independent. Formal security proof of CL-HVESS remains as a future work.

References

[1] NIST *Glossary of Key Information Security Terms, FIPS 191.*
[2] S. Kremer, O. Markowitch, J. Zhou*An Intensive Survey of Non-repudiation Protocols*, Computer Communications 25 (2002)1606-1621, 2002.
[3] J.L.F. Gomilla, J.A. Onieva, M. Payeras*Certified Electronic Mail: Properties Revisited*, Computer & Security (2009) 1-13, 2009.
[4] R. Dutta, P. Barua, P.Sarkar *Pairing Based Cryptography: A Survey*, 2004.
[5] C. Calik, O. Sever, H.M. Yildirim, Z. Yuce *A Survey of Certified Electronic Mail Protocols* 4^{th} ISC Turkey, 2010.
[6] S. Akleylek, B.B. Kirlar, O. Sever, Z. Yuce, *Pairing Based Cryptography: A Survey* 3^{rd} ISC Turkey, 2008.
[7] US Legal Definition *http://definitions.uslegal.com/e/e-contract/*
[8] *https://www.e-contract.be/*
[9] *http://www.signable.co.uk/legal*
[10] *https://www.ctmecontracts.com/eContracts/wp/index.htm*
[11] *http://www.telefonica.com/en/about_telefonica/html/suppliers/soluciones/econtracts.shtml*
[12] C. Galdi, R. Giordano *Certified email with temporal authentication: An improved optimistic protocol* Proceedings of International Conference on Trust and Privacy in Digital Business (TrustBus04), LNCS, vol.3184, Springer, Berlin, 2004, pp.181-190.
[13] R. Oppliger, P. Stadlin *A certified mail system (CMS) for the Internet*, Comput.Commun.27 2004 12291235.
[14] M. Franklin, G. Price *A comparison between traditional Public Key Infrastructures and Identity-Based Cryptography*, 2002.
[15] D. Boneh, M. Franklin *Identity Based Encryption from Weil Pairing.* SIAM J.of Computing Vol.32 No.3, 2003, Extended Abstract in Crypto 2001.
[16] A. Boldyreva, *Efficient Threshold Signature, Multisignature and Blind Signature Schemes Based on the Gap-Diffie-Hellman-Group Signature Scheme.* PKC 2003,LNCS 2139, pp.31-46 Springer-Verlag 2003.
[17] D. Boneh, B. Lynn, H. Shacham. *Short Signatures from the Weil Pairing.* in Proceedings of Asiacrypt 2001.
[18] F. Zhang, K. Kim. *ID-Based Blind Signature and Ring Signature from Pairings.* Advances in Cryptology in AsiaCrypt 2002, LNCS Vol.2510, Springer-Verlag, 2002.
[19] F. Zhang, R. Safavi-Naini, W. Susilo. *Efficient Verifiably Encrypted Signature and Partially Blind Signature from Bilinear Pairings.* In Proceedings of IndoCrypt 2003, Springer-Verlag, 2003.
[20] F. Hess, *Efficient Identity Based Signature Schemes Based on Pairings*, SAC 2002, LNCS 2595 Springer Verlag, 2000.
[21] J.A. Onieva, J. Zhou and J. Lopez *Multi-Party Non-Repudiation: A Survey* ACM Computing Surveys, 2008.

[22] C. Galdi, R. Giordano *Certified E-mail with temporal authentication: An improved optimistic protocol* LNCS Vol.3184, 2004.
[23] A. Joux *One Round Protocol for Tripartite Diffie Hellman* LNCS Vol.1838, 2000.
[24] S.S. Al-Riyami, K.G. Paterson *Certificateless Public Key Cryptography* AsiaCrypt 2003.
[25] C. Bamboriya, S.R. Yadav *A Survey of Different Contract Signing Protocols*, Ijetae V.1, I:4, January 2014.
[26] L. Chen, C. Gu *Optimistic Contract Signing Protocol Based on Hybrid Verifiably Encrypted Signature* Advances in Information Sciences and Service Sciences(AISS) V.4, N:12, July 2012.
[27] I. Blake, G. Seroussi, N. Smart *Advances in Elliptic Curves in Cryptography* Number 317 in London Mathematical Society Lecture Note Series. Cambridge University Press. ISBN 0-521-60415-X, 2005.
[28] S.D. Galbraith, K.G. Paterson, N.P. Smart *Pairings for Cryptographers* Elsevier 2008, Cryptology ePrint Archive, Report 2006/165.
[29] E.R. Verheul *Evidence that XTR is more secure than supersingular elliptic curve cryptosystems.* in EuroCrypt 2001, 195-210.
[30] R. Barbulescu, P. Gaudry, A. Joux, E. Tomme *A Quasi-polynomial Algorithm for Discrete Logarithm in Finite Fields of Small Characteristic* in EuroCrypt 2014.
[31] R. Granger, T, Kleinjung, J. Zumbragel *Breaking 128 bit Secure Binary Curves*

Data Storage of Electronic Exams

Lütfü Tarkan Ölçüoğlu, Sedat Akleylek

***Abstract*—Electronic learning is one of the most popular topics of today's technology and education. The development in technology forces universities, colleges and other education institutes to transfer their materials into digitized environment. Moreover, this new education system allows students to attend lessons from their computers located out of campus. In fact, some of the universities and colleges have started to hold their examinations in electronic environment. Electronic exam is one of the hardest problems in electronic learning subject, since it needs authenticity, anonymity, robustness and secrecy for all parts. The secrecy of electronic exam depends on the secrecy of both questions and their answers. In this work, we propose a data storage model for electronic exam which provides a long term confidentiality on the sensitive data such as questions and their corresponding answers by using verifiable secret sharing scheme.**

***Index Terms*—Long term confidentiality, verifiable secret sharing scheme, threshold cryptography, e-learning, e-exam.**

I. Introduction

THE development in technology affects every part of our life. People need to reach any data in a quick way, so that almost every data in any field have started to become digital and stored in database servers. For instant access to data, lots of institutions started to store their data in cloud systems. Some of this data in cloud has secret information and should be stored in encrypted form. Unfortunately, the encrypted form of secret information in cloud does not provide full secrecy since the encryption algorithm is not resistant forever. Therefore, new storage techniques are needed to provide full secrecy.

Education is one of the crucial parts of our life and there have been enormous technological developments in this area. Over two decades, lots of institutes, universities, colleges have transferred their documents to computerized environment especially to cloud which provides instant access to materials for both students and teachers. These developments in education have revealed a new definition to literature: electronic learning (e-learning). The first definition of e-learning was done in 1999 during CBT system seminar in Los Angeles. However, the development in e-learning field enabled lots of universities, colleges and institutes to hold their exams in a computerized environment. Electronic examination (e-exam) is one of the difficult parts of e-learning since there have been great amount of sensitive data behind it. Basically, e-exam can be defined as the computerized version of paper based exam on the other hand, for holding a secure electronic exam, there are cryptographic problems to be solved. Furthermore, an e-exam consists of registration, question preparation, exam preparation, evaluation and archiving parts. In registration part, students are registered for the exam and some relevant information are taken. Questions and answers are prepared by question makers and sent to authority in question preparation part. The authority prepares exam and after examination both evaluation and archiving processes are performed. These phases can be achieved by using cryptographic techniques due to satisfying information security concepts. The security of e-exam relies on the secrecy of questions and answers. The most popular example of electronic exam is the test of English as a foreign language (TOEFL) by educational testing service (ETS) [12]. More than 30 million people [11] have taken the test all over the world. The structure of TOEFL is mainly based on creating and administering the test questions, analysing the results, rejecting or revising the questions and releasing the test questions in a test form phases. They use the Internet security protocols which is used by major financial companies for transmission of the exam questions. The exam questions are downloaded to client's computer in encrypted form. Not only TOEFL, but also other popular electronic examinations like graduate record examination (GRE) and graduate management admission test (GMAT) use the same security protocols but the details of them are not revealed. One of the detailed work in this area was done by Jordi et al., [6] who proposed an e-exam scheme divided into stages: setting up an exam, beginning, holding and submitting of the exam, grading of exam, obtaining the score of the exam answer and revising of exam. They identified the security requirements for electronic exam as authenticity, privacy, correction, secrecy, receipt and copy detection. In this scheme, the privacy was achieved by the maximum impartiality, i.e., the teachers should not know the identity of the students while grading the exam. On the other hand, exam questions were kept in secret and the secrecy of the questions and answers was achieved by the encryption with manager's public key since all participants have digital certificates. Exam questions were prepared by teachers and they were encrypted with manager's public key. The storage of sensitive data was done just only encryption with manager's public key. In [2] Huszti and Pethö's proposed electronic examination scheme, they emphasized the secrecy of student's identity. The exam scheme consists of registration, exam and grading phases. The anonymity of students' identities provided by timed-release service containing n-servers. The secret, i.e., the identity of students shared into other servers by Shamir's secret sharing system. They proposed to use Mixnet for data storage. The questions were created by a committee and they were encrypted with Mixnet's public key. The authenticity of the questions was assured by committee's signatures. Both [6] and [2] emphasize the secrecy of the identity of students and in both scheme, the exam questions are prepared, encrypted

Lütfü Tarkan Ölçüoğlu is with the Institute of Applied Mathematics, Middle East Technical University, Ankara, Turkey (e156792@metu.edu.tr)

Sedat Akleylek is with the Department of Computer Engineering, Ondokuz Mayis University, Samsun, Turkey (sedat.akleylek@bil.omu.edu.tr).

and used in the exam, so that there is no long term storage for exam questions.

A. Our Contribution

Our aim in this work is to propose a data storage protocol for secure electronic exam consisting of secret sharing and long term confidentiality on sensitive data. The long term confidentiality of sensitive data is one the biggest challenges and practical long term confidentiality is an open problem. Since the secrecy of e-exam depends on the secrecy of questions and answers, the data storage of such sensitive data is important. Other related works [6] and [2] emphasized the secrecy of the students' identity and both of them encrypted questions and answers with authority's or Mixnet's public key. Single encryption of such sensitive data is not enough, so that long term confidentiality issues should be thought. Unlike [6] and [2], our model emphasizes the long term confidentiality of the sensitive data.

B. Roadmap

The outline of the paper is as follows: In section II, we give some brief information about cryptographic needs and definitions. In section III, the data storage model is defined in details. The security analysis of the model is done in section IV. Finally, in section V, the conclusion and future work are given.

II. PRELIMINARIES

In this section we give some definitions and brief information about cryptographic requirements. The algorithms and protocols defined in this section will be used in the model. The security requirement of the given model based on symmetric cipher for the encryption of the question and answers, asymmetric key encryption for authentication issues and threshold cryptography for long term confidentiality to provide confidentiality.

A. Symmetric Key Encryption: AES

AES is used for confidentiality of our questions' main parts. The AES [7] also known as Rijndael was developed by Vincent Rijmen and Joan Daemen in 2001. It was the winner of AES competition established by NIST. AES is a substitution permutation network with the block size of 128 bits and the key size of 128,192 and 256 bits. For security, until now, there is no known practical attack to AES to get the plaintext from the ciphertext. We will use AES for encryption of questions and answers.

B. Asymmetric Key Encryption: RSA

Asymmetric key encryption is a cryptographic protocol which uses different key pairs (private and public) in encryption and decryption. RSA is one of the most preferred asymmetric key encryption method introduced by Ron Rivest, Adi Shamir and Leonard Adleman in 1977 [10]. The RSA public key cryptosystem based on the integer factorization problem. We will use 2048-bit RSA keys in our model. Since the security of RSA is based on the integer factorization problem, there is no active attack to 2048-bit RSA keys.

C. Threshold Cryptography

Threshold cryptography is the distribution of a secret to a group in a multi-sender, multi-receiver systems. In cryptography, the first definition of threshold cryptography was given by Adi Shamir in 1979 [3]. The basic idea of threshold cryptography is sharing a secret. The secret is divided into pieces by the dealer and every piece of it sent to participants. A number of participants should come together in order to get the secret. This is the main idea of $(t, n) - threshold\ cryptosystems$. In $(t, n)-$threshold cryptosystems at least t participants are required for decrypting the secret. In our model, we will use $(t, n)-$threshold cryptosystem where the authority acts as *dealer* and databases are *participants*.

1) Shamir's Secret Sharing: The secret sharing of our model based on *Shamir's Secret Sharing* scheme [3]. Shamir's secret sharing scheme based on polynomial interpolation. Let $s \in \mathbb{Z}_q$ be the secret to be shared where q is prime, t be the number of threshold for reconstructing the secret. The dealer chooses a polynomial $p(x)$ of degree t over $\mathbb{Z}_q$ such that $p(0) = s$. Each participant's secret piece s_i is computed by $p(i)$. Those $p(i)$'s are transmitted into each participant P_i in a secure channel. For reconstructing the secret, at least t participants provide their shares to get s by using polynomial interpolation.

2) Feldman's Verifiable Secret Sharing: If there is a malicious participant in secret sharing scheme, he can deal inconsistent share and reconstructing the secret will be failed. Verifiable secret sharing provides us to compute a procedure where consistent dealings can be verified. Verifiable secret sharing was firstly introduced by Chor et.al., in 1985 [4]. The most common use of verifiable secret share was introduced by Feldman [8]. Feldman's verifiable secret sharing based on Shamir's secret sharing scheme combined with homomorphic encryption scheme. They used the similar idea, trapdoor function, given in RSA. Feldman's verifiable secret sharing scheme works as follows: Let p and q be prime numbers such that $q \mid p-1$. Let $g \in \mathbb{Z}_p$ of order q. The polynomial $p(x)$ over $\mathbb{Z}_p$ with coefficients $p_0, p_1, ..., p_k$ be chosen by the dealer. Then the dealer broadcasts the values $g^{p_0}, g^{p_1}, ..., g^{p_k}$ and secretly transmits the value $s_i = p(i)$ *(mod q)* to each participants P_i. The participants verify their own share by the following equation $g^{s_i} \stackrel{?}{=} (g^{p_0})(g^{p_1})^i(g^{p_2})^{i^2}...(g^{p_k})^{i^k}$ *(mod p)* where one can also consider this $i - adic$ representation of secret. If each participant's share is proper than the equation holds. For completing the dealing of the secret, all participants' share must be proper. We will use Feldman's verifiable secret sharing in our model. The authority will act as dealer and the database servers will act as participants.

III. DATA STORAGE MODEL

In this section, we will propose a new model to store exam materials in a secure way. To achieve the security requirements, we prefer to use verifiable secret sharing scheme to assure long term confidentiality on sensitive data. Our model has four parts: question preparation, question confirmation, storage to database, retrieving data from database. Great amount of data is used in electronic examination. Basically

this data can be personal information of users or questions and answers to be used in exam. The secure storage of this sensitive data is one of the biggest challenge for security. The following questions should be considered for a secure data storage of an electronic exam:

- How the data will be kept in the database such that no one will be able get them without secret key?
- What will happen if the secure algorithm is broken down with today's technology?
- How the integrity of the data is provided?

The answers of above questions can be the change of algorithm when it is broken down. This should be efficient solution to such problems; however, it is not enough if an adversary had retrieved encrypted data before the algorithm was broken down. The other challenge is the integrity of the sensitive data which can be provided by a verifiable mechanism. Therefore, a long term confidentiality with verification of such kind of sensitive data should be provided.

Mainly, the sensitive data for electronic exam is questions and answers. At the first sight, one can define the sensitive data for electronic exam as questions. However, the answers are sensitive data like questions since any adversary can guess the question from the answers with reverse engineering techniques or there is no need to try to solve the questions. So that, both questions and answers should be thought together as sensitive data for electronic exam. These questions and their answers are prepared by question makers for exam and both of them are sent to exam authority. In our model we assume that the channel between authority and question makers is secure. First, we give some notations:

- (P_{QMk}, S_{QMk}), (P_{Aut}, S_{Aut}) are public and private keys for question makers (QMk) and the authority (Aut) relatively,
- $\mathcal{E}$: Encryption function, $\overline{\mathcal{D}}$: Decryption function,
- The package $\mathcal{P} = \{\mathcal{E}_{P_{Aut}}(Q), \sigma_Q\}_{S_{QMk}}$ is a question package where $\mathcal{E}_{P_{Aut}}(Q)$ is the encrypted form of question. The encryption is done by public key of authority and σ is the signature of question signed by question maker.

A. Question Preparation

The security of an electronic exam relies on the security of the questions and answers which are prepared by the question makers. After preparation, they sign their questions, encrypt them with authority's public key and send to authority. Question makers decide the blocks of the question consisting of question part, choices, right answer and tags. The tags consist of question subject, category, subcategory and hardness. The sensitive part of the question blocks are question part, choices and right answer. The other part does not need to be encrypted, because they will be reference for the encrypted questions. A question maker prepares his question and sends to system given in (Fig. 1).

1) Question $Q = \{$*question part, choices, right answer, tags*$\}$
2) Encryption with authority's public key: $\mathcal{E}_{P_{Aut}}(Q)$
3) Signature of question:σ_Q
4) Package to be sent $\mathcal{P} = \{\mathcal{E}_{P_{Aut}}(Q), \sigma_Q\}_{S_{QMk}}$

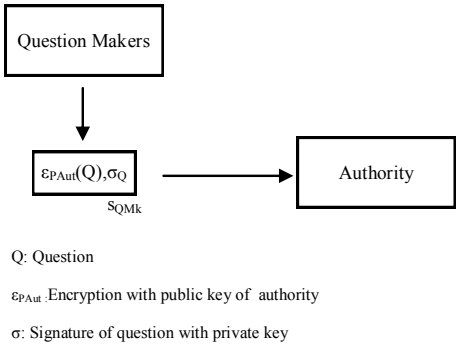


Fig. 1. Question Preparation

B. Question Confirmation

The authority should confirm the valid questions from valid users. The validation is done by the verification of the signature. If the signature is verified, the authority decrypts the question with his private key and edits the question if needed. The edition of the question is done by the editors supplied by the authority. Here, they control the blocks of the questions, revise them if needed and an automatic ID is given. The edited question is separated into two parts: Main and reference.

TABLE I
MAIN PART OF QUESTION

ID	Question Part	Choices	Right Answer
1	.....	.....	.....

TABLE II
REFERENCE PART OF QUESTION

ID	Subject	Category	Subcategory	Hardness
1	.....	.....	.....	.....

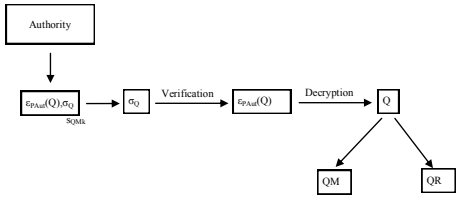


Fig. 2. Question Confirmation

The main part of the question consists of question part, choices and right answer in Table I. The reference part of the question consists of subject, category, subcategory, hardness level in Table II.

1) The authority Aut decrypts the encrypted question $\mathcal{E}_{P_{Aut}}(Q)$ with his private key.
2) Aut verifies the signature of the question σ_Q.

3) After the verification, the question Q is divided into main part QM and reference part QR shown in (Fig. 2).

C. Storage

The storage of the questions is done in two parts: main part of question and reference part of the question. The reference part of the question is not needed to be encrypted since you cannot get the question from the tags of it. They are just kept in the main database of authority.

In our model, the storage protocol consists of n databases where located in different places. The authority acts as *dealer* and sends the questions to other databases. The reference part of questions is stored in the dealer's main database without any encryption. In the storage part we prefer to use AES [7] for

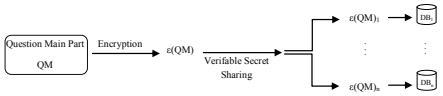


Fig. 3. Storage To Database

symmetric key encryption and verifiable secret sharing scheme [8] for secret sharing. One can also use other cryptographic protocols to provide confidentiality. The storage of the main part of the questions (Fig. 3) is done as follows:

1) The authority i.e., the dealer encrypts the main part of the question with symmetric key encryption: $QM \rightarrow \mathcal{E}(QM)$.
2) Encrypted question $\mathcal{E}(QM)$ is the secret. This secret will be shared with other databases by verifiable secret sharing scheme: $\mathcal{E}(QM) \rightarrow \mathcal{E}(QM)_1, \mathcal{E}(QM)_2, ..., \mathcal{E}(QM)_n$.
3) Every piece of the secret will be transferred into related databases: $DB_1, DB_2, ..., DBn$.

In the storage protocol we use Feldmans's verifiable secret sharing scheme [8] which provides protection of inconsistent dealings of misbehaving dealers. For long term confidentiality, the authority should renew shares periodically. In [9], Ostrovksy and Yung proposed $randomized\ secret$ verified by all dealers and updated by a polynomial. This can be used for renewal process but for consistent sharing and correct dealing of the secret we use Feldmans's verifiable secret sharing. Herzberg et.al., [1] used Feldmans's verifiable secret sharing for share renewal process to get consistent shares and correct dealing of the secret. In our model, we use the same methodology in [1] for periodically update of the share i.e., encrypted form of question pieces for long term confidentiality. Now we will show how verification of and renewal of the share are done?

1) Verification Of The Secret Share: Here we will show that how an encrypted question will be shared into the databases by Feldman's verifiable secret sharing scheme. Let p be a prime, q be a prime such that $q \mid p-1$. Let $g \in \mathbb{Z}_p$ of order q. Then the authority;

1) chooses a polynomial $p(x)$ over $\mathbb{Z}_p$ with coefficients $p_0, p_1, ..., p_k$ explained in Section II,
2) determines the secret is the encrypted form of the main part of question i.e., $\mathcal{E}(QM)$. (Here we assume that $\mathcal{E}(QM) \in \mathbb{Z}_p$),
3) broadcasts the values $g^{p_0}, g^{p_1}, ..., g^{p_k}$,
4) transmits the value of $\mathcal{E}(QM)_i = p(i)$ *(mod q)* to each database servers DB_i

After that process each DB_i should check the equation

$$g^{s_i} \stackrel{?}{=} (g^{p_0})(g^{p_1})^i(g^{p_2})^{i^2}...(g^{p_k})^{i^k} \ \textit{(mod p)}.$$

The equation holds if and only if the share of DB_i is proper. The secret sharing will complete when all DB_i's complete the verification of the equation. The scheme explained above can be done with quantum resistant schemes which is a future work in our proposed model.

2) Renewal Of The Secret Share: For long term confidentiality of any data, the encryption scheme should be updated in some periods. The major concern of this phase is what will happen if an adversary provides inconsistent share updates during the share renewal phase. To solve such scenarios [1] and [5] proposed to use verifiable secret sharing schemes. In our model, we use the same method used in [1]. We perform our share renewal with verifiable secret sharing in order to detect the wrong dealt shares by the database servers. The renewal of the share is done in database servers P_i's as follows:

1) Each P_i defines a polynomial $\delta_i(z) = \delta_{i1}z^1 + \delta_{i2}z^2 + ... + \delta_{ik}z^k$ such that k is random numbers $\{\delta_{im}\}_m$ from $\mathbb{Z}_q$ where $m \in \{1, ..., k\}$.
2) P_i computes $\epsilon_{im} = g^{\delta_{im}}$ *(mod p)* where $m \in \{1, ..., k\}$.
3) P_i computes $u_{ij} = \delta_i(j)$ *(mod q)* where $j \in \{1, ..., n\}$ and $e_{ij} = \mathcal{E}_j(u_{ij}), \forall i \neq j$
4) The message $M_i^{(t)} = (i, t, \epsilon_{im}, e_{ij})$ where $j \in \{1, ..., k\} - \{i\}$ and the signature $\sigma_i(M_i^{(t)})$ is prepared and broadcasted by P_i.
5) P_i decrypts the e_{ij} comes from the other participants, verifies the correctness of the share by the equation explained in Feldmans's Verifiable Secret Sharing Scheme by using $g^{u_{ji}} \stackrel{?}{=} (\epsilon_{j1})^i(\epsilon_{j2})^{i^2}...(\epsilon_{jk})^{i^k}$ *(mod p)*
6) If the messages from other participants are correct, then the above equation holds. Therefore P_i has done the verification and accepted the messages from other participants.
7) P_i updates his own share by $s_i^{(t)} \leftarrow s_i^{(t-1)} + (u_{1i} + u_{2i} + ... + u_{ni})$ *(mod q)* and erases the other variables.

In the above process, if there exists irregularities in the verification part, the dealer must detect the misbehaving participant. Each database server checks the other servers' messages. The participant contacts with the dealer to resolve the inconsistent behaviour when the verification is not done. For this kind of accusation, all honest participants agree on the malicious participant. Then the dealer sends random value to malicious server and wants to encrypt and sign it. If the malicious server (P_d) is not verified by the dealer, then the renewal process updated by the equation:
$s_i^{(t)} \leftarrow s_i^{(t-1)} + \sum_j u_{ji}$ where $j \neq d$ *(mod q)*

D. Retrieving Data

The sensitive data in e-exam, the questions and answers are needed for the exam. The authority should decide which questions are needed for the exam. We separate all questions into two parts: main and reference. The main part of all questions is kept encrypted in the databases. Unlike the main part, reference parts are kept in the main database without any encryption. For preparation of an exam, a committee decides the tags of the questions, i.e., subjects, categories, subcategories and hardness of all questions. With these requirements, the authority selects the questions from reference tables. After that selection, the IDs of questions are determined. For examination, the authority should retrieve the questions with determined IDs. The authority selects the required questions as follows:

1) Let $EXAM = \{ID_1, ID_2, ..., ID_n\}$ be the set of questions to be used in the exam.
2) Let $QUES = \{\mathcal{E}(QM_1), \mathcal{E}(QM_2), ..., \mathcal{E}(QM_n)\}$ be the set of questions to retrieved. Authority assigns each ID in $EXAM$ to $QUES$ respectively.
3) Let $\mathcal{E}(QM_i) = s_r = p^t(r)$ be the secret to retrieved where $r \in \mathcal{B}$ such that $\mathcal{B}$ is the set of servers have incorrect shares.
4) Every database servers $P_i \in \mathcal{D} = \mathcal{A} - \mathcal{B}$ choose k-degree random polynomial δ_i over $\mathbb{Z}_q$ where $\delta_i(r) = 0$ and compute $\delta_{i0} = -\sum_j \delta_{ij} r^j$ *(mod q)*, $j \in \{1, ..., k\}$ where $\mathcal{A}$ is the set of servers.
5) Each P_i broadcasts $\mathcal{E}_j(\delta_i(j))$, $i, j \in \mathcal{D}$.
6) Each P_i's creates new share $s'_r = s_r + \sum_j \delta_j(i)$ and sends to P_r with $\mathcal{E}_r(s'_i)$
7) P_r decrypts the share, and with polynomial interpolation recover s_r
8) The authority uses the key for AES and decrypt $s_r = \mathcal{E}(QM)$ and gets $\overline{\mathcal{D}}(\mathcal{E}(QM)) = QM$ for examination.

IV. Security Analysis

In this section we will give some basic security analysis in order to show that our proposed model is secure. The model based on (t, n)-threshold scheme with verifiable secret sharing scheme assuming that $t - 1 < n/2$.

Theorem 1: If there are at most $t - 1$ malicious database servers, then the proposed protocol reconstruct the secret by honest servers and the system remains secure.

Proof: We assumed that $t - 1 < n/2$ in $(t, n)-$threshold scheme. So that if there are $t - 1$ malicious servers than $t = n/2$ trusted servers. Therefore, the secret is reconstructed by $t = n/2$ trusted server by the definition of secret sharing scheme. ■

Theorem 2: The proposed data storage protocol possesses authenticity of the servers during the renewal of the secret.

Proof: In the renewal of secret phase, each server P_is should sign a message $\mathcal{M}$ with his private key. The signature of the message $\mathcal{M}$ is verified by each server and Feldman's verifiable secret sharing equation holds. This verification provides authenticity of the servers. ■

V. Conclusion and Future Works

The development in technology affects almost every area of our life. Especially, the development in digital technology reveals the confidentiality of sensitive data. Electronic learning is today's one of the most popular subject. Nevertheless, electronic exam is also popular cryptographic subject of this branch. In this paper, we propose a new data storage model for electronic exam. We show that the proposed model is secure under the condition of the periodic renewal of the share with the authenticated servers. The model is based on verifiable secret sharing scheme and long term confidentiality. The secrecy of electronic exam is provided by the secrecy of sensitive data which is questions and answers. With the given model, every question encrypt with symmetric key encryption algorithm and split into pieces with verifiable secret sharing scheme. Every piece of encrypted questions are kept in different databases located in different places. The authority is responsible every part of electronic exam, so that for long term confidentiality, the secret shares should be updated periodically by him. As a future work, the given model should be extended to other applications using great amount of data with many users. Also, quantum resistant schemes can be used for the proposed scheme.

Acknowledgment

The authors would like to thank Ali Doğanaksoy for his valuable contributions and feedback.

References

[1] A. Herzberg, S. Jarecki, H. Krawczyk and M. Yung, *Proactive Secret Sharing Or: How to Cope With Prepetual Leakage*, Lecture Notes in Computer Science, Springer-Verlag, 1995, pp:339-352.
[2] A. Huszti and A. Pethö, *A Secure Electronic Exam System*, Publ. Math. Debrecen, 77/3-4, 2010, pp:299-312.
[3] A. Shamir, *How to Share a Secret*, Communications of ACM, 1979, pp:612-613.
[4] B. Chor, S. Goldwasser, S. Micali and B. Awerbuch, *Verifiable Secret Sharing and Achieving Simultaneous Broadcast*, Proc. of IEEE, 1985, pp:335-344.
[5] J. Braun, J. Buchmann, C. Mullan and A. Wiesmaier, *Long Term Confidentiality: a Survey*, Designs, Codes and Cryptograpyh, Springer, 2012.
[6] J. Castella-Roca, J. Herrera-Joancomarti and A. Dorca-Josa, *A Secure E-Exam Management System*, Proceeding of the First International Conference on Avaliability, Reliability an Security (ARES'06), 2006, pp:864-871.
[7] J. Daemen and V. Rijmen, *The Design of Rijndale: AES - The Advanced Encryption Standard*, Springer Verlag, Berlin, Heidelberg, New York, 2002.
[8] P. Feldman, *A Practical Scheme for Non-Interactive Verifiable Secret Sharing*, Proc. of the 28th IEEE Symposium on the Foundations of Computer Science, 1987, pp:427-437.
[9] R. Ostrovsky and M. Yung, *How To Withstand Mobile Virus Attacks*, Proc. 10th ACM Conf. on Principle of Distributed Systems, 1991
[10] R. Rivest, A. Shamir and L. Adleman, *A Method For Obtaining Digital Signatures and Public-Key Cryptosystems*, Communications of the ACM, 1978, pp:120-126.
[11] *About The TOEFL PBT Test*, ETS, https://www.ets.org/toefl/pbt/about, 2015.
[12] *How ETS Protects The Integrity of The TOEFL Test*, ETS, https://www.ets.org/toefl/institutions/about/security, 2015.

More Efficient Secure Outsourcing Methods for Bilinear Maps

Öznur Arabacı, Mehmet Sabir Kiraz, İsa Sertkaya, and Osmanbey Uzunkol
Mathematical and Computational Sciences
TÜBİTAK BİLGEM, Turkey
{oznur.arabaci, mehmet.kiraz, isa.sertkaya, osmanbey.uzunkol}@tubitak.gov.tr

***Abstract*—Bilinear maps are popular cryptographic primitives which have been commonly used in various modern cryptographic protocols. However, the cost of computation for bilinear maps is expensive because of their realization using variants of Weil and Tate pairings of elliptic curves. Due to increasing availability of cloud computing services, devices with limited computational resources can outsource this heavy computation to more powerful external servers. Currently, the checkability probability of the most efficient outsourcing algorithm is $1/2$ and the overall computation requires 4 point addition in the preimage and 3 multiplications in the image of the bilinear map under the one-malicious version of a two-untrusted-program model. In this paper, we propose two efficient new algorithms which decrease not only the memory requirement but also the overall communication overhead.**

***Index Terms*—Outsourcing computation, Bilinear maps, Secure delegation, Secure Cloud Computing.**

I. Introduction

THE improvements in the cloud computing services result in variety of new security and privacy challenges. Many cryptographic mechanisms involving complex computations such as bilinear maps are proposed to overcome these challenges [1], [2], [3]. Since speeding up the computation of bilinear maps is crucial in real-life applications, many schemes are suggested to reduce the computational cost of pairing computation [4], [5], [2], [6], [7], [8]. Especially, Hess introduced a general framework encompassing different types of pairing functions giving optimum numbers of computation steps [9]. However, these computations are still infeasible or unaffordable for resource constrained devices including mobile phones, tablets, smart or RFID cards.

Since Hohenberger and Lysyanskaya stated the question of how a computationally limited device may outsource its computation to another, potentially malicious, but much more computationally powerful device [10], it has been studied extensively. Outsourcing the complex computations to external powerful devices dates back to Matsumoto, in which the RSA signature generation problem is considered [11]. More specifically, Chevallier-Mames *et al.* proposed a protocol enabling a computationally limited device to outsource the computation of bilinear maps into a more resourceful device [12]. However, this delegation process brought new concerns. Firstly, the external device should learn nothing about the secrets. Also, the computationally-limited device should be able to check whether the external device computed correctly, at least with certain probability. These two concerns can be eliminated by masking the secret values with the cost of some extra computations before sending to the external server, and then removing the masking values together with a way of validating the outsourced computation.

Besides the efficiency constraints, secrecy is the main objective of the security model, since the input and output pair of a client is used for cryptographic purposes. Henceforth, outsource mechanisms surely follow a security model in which the client (the energy limited trusted device that needs to delegate the computation) does not trust the servers (which perform the needed computations). Thus, in the security model, it is assumed that the client is honest but the servers are untrusted. Furthermore, checkability, validation of the computation processes, should be also addressed.

As simulated in Figure 1, outsource computation protocols may utilize one or more servers. Based on the number of servers utilized Tian *et al.* classified them as follows [13]:

- One-Untrusted Program (OUP): One malicious server performs the computation.
- One-Malicious version of a Two-Untrusted Program (OMTUP): Two untrusted servers perform the computation but only one of them may behave maliciously.
- Two-Untrusted Program (TUP): Two untrusted servers perform the computation and both of them may behave maliciously, but they do not maliciously collude.

Following the work of Chen *et al.*, [14], Tian *et al.* proposed two algorithms [13] in the OMTUP setting. First algorithm achieves less computational complexity and the second one improves the checkability at the cost of some additional computations. These outsource protocols are composed of both offline and online computation steps. In the offline phase, the client prepares the necessary values. During the online phase, the client creates masked values based on the precomputed offline values and requests the bilinear map computation.

In this paper, we propose two new algorithms following the steps of Chen *et al.* and Tian *et al.*'s work. We further analyze the protocols under the OMTUP assumption and reduce not only the computational complexity of the offline computations, but also the memory needed to store the values resulting from the offline computations together with the communication overhead. While doing so, we do not increase the computation costs that need to be handled by the client.

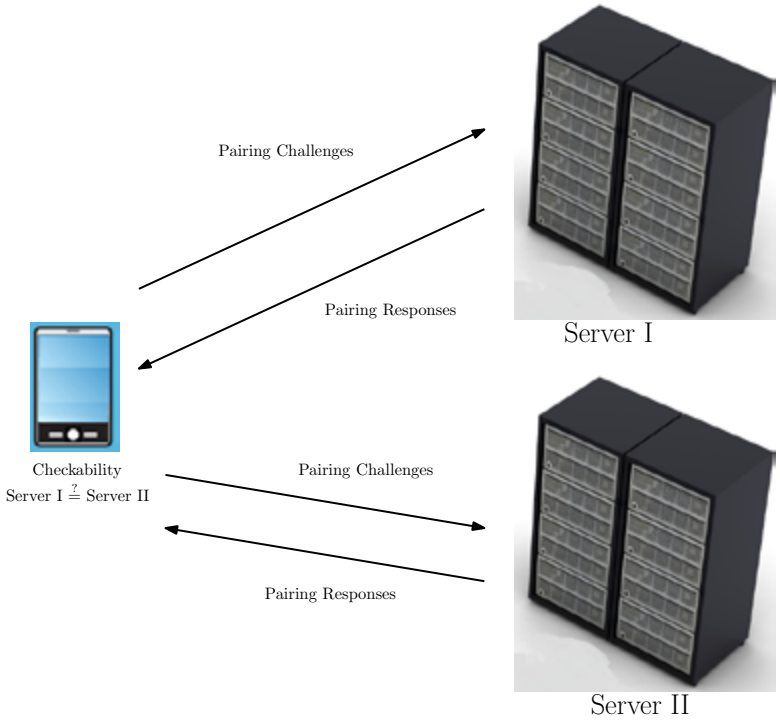


Fig. 1. Outsourcing Bilinear Maps with Two Untrusted Cloud Servers

A. Related work

Weil and Tate pairings are firstly used as cryptanalytic tools for reducing the discrete logarithm problem (DLP) on some elliptic curves to DLP on finite fields [15]. Later, Boneh *et al.* and Joux constructed new cryptographic protocols based on bilinear maps [1], [2], [3]. Reducing the computational cost of bilinear maps are suggested in [4], [5], [2], [6], [7], [8], [16], [9].

First protocol for secure outsourcing of elliptic curve pairings were proposed by Chevallier-Mames *et al.* [12]. The algorithm assumes the OUP setting and it is 1-checkable. However, the algorithm requires expensive computations, namely multiple membership test operation which is equivalent to an exponentiation over the finite field and inversion on the exponents. Later on, under the same OUP assumption, Kang *et al.* [17] and Canard *et al.* [18] improved the computational complexity results. However, the solutions were not feasible since exponentiation, membership test, and inversion were still required. Tsang *et al.* made a taxonomy for pairing based computations and constructed a batch pairing delegation mechanism [19]. Chow *et al.* studied server aided signature verification [20].

Chen *et al.* broke the paradigm by utilizing two servers under the OMTUP assumption and by performing some computations during an idle time of the resource-limited device [14]. As a result, this outsourcing mechanism of bilinear maps was the first one which does not depend on the membership test operations and exponentiations over the finite field. Additionally, this scheme decreased the online computations on the client side. The user had to perform only 5 point additions in $\mathbb{G}_1$ and $\mathbb{G}_2$, and 4 multiplications in $\mathbb{G}_3$, where $e: \mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_3$ is the underlying bilinear map. Later, Tian *et al.* proposed a more efficient algorithm [13] and reduced the online computation phase to 4 point additions in $\mathbb{G}_1$ and $\mathbb{G}_2$, and 3 multiplications in $\mathbb{G}_3$.

B. Our contributions

In this paper, we propose two efficient new algorithms for secure outsourcing of bilinear maps. Compared to the state of the art algorithms (especially Tian *et al.*'s [13]), our algorithms need less offline computations, less memory, and less queries to the servers. In order to manage that, different from the previous studies, we use negation of an input value (it is almost for free since it is located over the elliptic curve), and we also send the same checking computation to both servers. Since it is assumed that the two servers do not collude, we reduce the computation costs without affecting the checkability of the system. The second algorithm may for instance be utilized in signature verification applications, in which we evade from at least one multiplication. For both propositions, we also provide the security model following exactly the lines of the security model of Hohenberger and Lysyanskaya [10]. We conclude the paper by comparing the efficiency of the system with the very recent work of Tian *et al.* [13].

C. Roadmap

In Section II, we give the security definitions for the outsourcing algorithm. Then, we present some background and preliminaries that will be needed throughout the manuscript, and we propose our two main algorithms together with their security analysis in Section III. Next, in Section IV, we analyze complexity of our new algorithms and compare it to the complexity of the best known algorithm [13]. Finally, we conclude the paper in Section V.

II. Security Model

Chen *et al.*. [14] and Tian *et al.* [13] follow the security model proposed by Hohenberger and Lysyanskaya [10]. We remark especially that we also follow exactly their security model [10].

Definition 1: An algorithm is said to obey the outsource input/output specification if it takes five inputs, and produces three outputs. The first three inputs are generated by an honest party, and are classified by how much the adversary $A = (E, U')$ knows about them. The first input is called the honest, secret input, which is unknown to both E and U; the second is called the honest, protected input, which may be known by E, but is protected from U; and the third is called the honest, unprotected input, which may be known by both E and U. In addition, there are two adversarially-chosen inputs generated by the environment E: the adversarial, protected input, which is known to E, but protected from U; and the the adversarial.

Definition 2: Let $\mathsf{Alg}(.,.,.,.,.)$ be an algorithm with outsource-IO. A pair of algorithms (T, U) is said to be an outsource-secure implementation of an algorithm Alg if:
Correctness. T^U is a correct implementation of Alg.
Security. For all probabilistic polynomial-time adversaries $A = (E, U')$, there exist probabilistic expected polynomial-time simulators (S_1, S_2) such that the following pairs of random variables are computationally indistinguishable. Let us say that the honestly-generated inputs are chosen by a process I.

- **Pair One:** $EVIEW^i_{real}$~$EVIEW_{ideal}$ (The external adversary, E, learns nothing.):
  - The view that the adversarial environment E obtains by participating in the following REAL process:

$$\begin{aligned} EVIEW^i_{real} = \{&(istate^i, x^i_{hs}, x^i_{hp}, x^i_{hu}) \leftarrow \\ &I(1^k, istate^{i-1}); \\ &(estate^i, j^i, x^i_{ap}, x^i_{au}, stop^i) \leftarrow \\ &E(1^k, EVIEW^{i-1}_{real}, x^i{}_{,hp}, x^i_{hu}); \\ &(tstate^i, ustate^i, y^i_s, y^i_p, y^i_u) \leftarrow \\ &T^{U'(ustate^{i-1})}(tstate^{i-1}, x^{j^i}_{hs}, x^{j^i}_{hp}, x^{j^i}_{hu}, x^i_{ap}, x^i_{au}) : \\ &(estate^i, y^i_p, y^i_u)\} \end{aligned}$$

    $EVIEW_{real} = EVIEW^i_{real}$ if $stop^i = TRUE$.
    The real process proceeds in rounds. In round i, the honest (secret, protected, and unprotected) inputs $(x^i_{hs}, x^i_{hp}, x^i_{hu})$ are picked using an honest, stateful process I to which the environment does not have access. Then the environment, based on its view from the last round, chooses (0) the value of its $estate_i$ variable as a way of remembering what it did next time it is invoked; (1) which previously generated honest inputs $(x^{j^i}_{hs}, x^{j^i}_{hp}, x^{j^i}_{hu})$ to give to $T^{U'}$ (note that the environment can specify the index j^i of these inputs, but not their values); (2) the adversarial, protected input x^i_{ap}; (3) the adversarial, unprotected input x^i_{au}; (4) the Boolean variable $stop^i$ that determines whether $roundi$ is the last round in this process. Next, the algorithm $T^{U'}$ is run on the inputs $(tstate^{i-1}, x^{j^i}_{hs}, x^{j^i}_{hp}, x^{j^i}_{hu}, x^i_{ap}, x^i_{au})$, where $tstate^{i-1}$ is Ts previously saved state, and produces a new state $tstate^i$ for T, as well as the secret y^i_s, protected y^i_p and unprotected y^i_u outputs. The oracle U' is given its previously saved state, $ustate^{i-1}$, as input, and the current state of U' is saved in the variable $ustate^i$. The view of the real process in $roundi$ consists of $estate^i$, and the values y^i_p and y^i_u. The overall view of the environment in the real process is just its view in the last round (i.e., i for which $stop^i = TRUE$).
  - The IDEAL process:

$$\begin{aligned} EVIEW^i_{ideal} = \{&(istate^i, x^i_{hs}, x^i_{hp}, x^i_{hu}) \leftarrow \\ &I(1^k, istate^{i-1}); \\ &(estate^i, j^i, x^i_{ap}, x^i_{au}, stop^i) \leftarrow \\ &E(1^k, EVIEW^{i-1}_{ideal}, x^i{}_{,hp}, x^i_{hu}); \\ &(astate^i, y^i_s, y^i_p, y^i_u) \leftarrow \\ &\mathsf{Alg}(astate^{i-1}, x^{j^i}_{hs}, x^{j^i}_{hp}, x^{j^i}_{hu}, x^i_{ap}, x^i_{au}); \\ &(sstate^i, ustate^i, Y^i_p, Y^i_u, replace^i) \leftarrow \\ &S_1^{U'(ustate^{i-1})}(sstate^{i-1}, x^{j^i}_{hp}, x^{j^i}_{hu}, x^i_{ap}, x^i_{au}, y^i_p, y^i_u); \\ &(z^i_p, z^i_u) = replace^i(Y^i_p, Y^i_u) + (1 - replace^i)(y^i_p, y^i_u) : \\ &(estate^i, z^i_p, z^i_u)\} \end{aligned}$$

    $EVIEW_{ideal} = EVIEW^i_{ideal}$ if $stop^i = TRUE$.
    The ideal process also proceeds in rounds. In the ideal process, we have a stateful simulator S_1 who, shielded from the secret input x_{hs}, but given the non-secret outputs that Alg produces when run all the inputs for round i, decides to either output the values (y^i_p, y^i_u) generated by Alg, or replace them with some other values (Y^i_p, Y^i_u) (Notationally, this is captured by having the indicator variable $replace^i$ be a bit that determines whether y^i_p will be replaced with Y^i_p.) In doing so, it is allowed to query the oracle U'; moreover, U' saves its state as in the real experiment.
- **Pair Two**: $UVIEW_{real}$~$UVIEW_{ideal}$ (The untrusted software, (U_1, U_2), learns nothing.):
  - The view that the untrusted software U' obtains by participating in the REAL process described in Pair One. $UVIEW_{real} = ustate^i if stop^i = TRUE$.
  - The $IDEAL$ process:

$$\begin{aligned} UVIEW^i_{ideal} = \{&(istate^i, x^i_{hs}, x^i_{hp}, x^i_{hu}) \leftarrow \\ &I(1^k, istate^{i-1}); \\ &(estate^i, j^i, x^i_{ap}, x^i_{au}, stop^i) \leftarrow \\ &E(1^k, estate^{i-1}, x^i{}_{,hp}, x^i_{hu}, y^{i-1}_p, y^{i-1_u}); \\ &(astate^i, y^i_s, y^i_p, y^i_u) \leftarrow \\ &\mathsf{Alg}(astate^{i-1}, x^{j^i}_{hs}, x^{j^i}_{hp}, x^{j^i}_{hu}, x^i_{ap}, x^i_{au}); \\ &(sstate^i, ustate^i) \leftarrow \\ &S_2^{U'(ustate^{i-1})}(sstate^{i-1}, x^{j^i}_{hu}, x^i_{au}) : \\ &(ustate^i)\} \end{aligned}$$

    $UVIEW_{ideal} = UVIEW^i_{ideal}$ if $stop^i = TRUE$.
    In the ideal process, we have a stateful simulator S_2 who, equipped with only the unprotected inputs (x^i_{hu}, x^i_{au}), queries U'. As before, U' may maintain state.

In our security model we assume one-malicious version of a two-untrusted program (OMTUP) model. More concretely, there are two untrusted cloud servers in this model perform-

ing the outsourced computation, where only one of them is assumed to be malicious.

Definition 3: A pair of algorithms (T, U_1, U_2) are an α-efficient implementation of an algorithm Alg if (1) they are an outsource-secure implementation of Alg, and (2) $\forall$ inputs x, the running time of T is $\leq$ an α-multiplicative factor of the running time of $\mathsf{Alg}(x)$.

Definition 4: A pair of algorithms (T, U_1, U_2) are an β-checkable implementation of an algorithm Alg if (1) they are an outsource-secure implementation of Alg, and (2) $\forall$ inputs x, if $U_i', i = 1, 2$ deviates from its advertised functionality during the execution of $T^{(U_1', U_2')}(x), T$ will detect the error with probability $\geq \beta$.

Definition 5: A pair of algorithms (T, U_1, U_2) are an (α, β)-outsource secure implementation of an algorithm Alg if they are both α-efficient and β-checkable.

III. Algorithms for Outsourcing of Bilinear Maps

A. Preliminaries: Bilinear Maps

Let $(\mathbb{G}_1, +)$ and $(\mathbb{G}_2, +)$ be two additive cyclic groups of order q with $\mathbb{G}_1 =< Q >$ and $\mathbb{G}_2 =< P >$, $(\mathbb{G}_3, \cdot)$ be a multiplicative cyclic group of order q, where q is a prime number and $0_{\mathbb{G}_1}, 0_{\mathbb{G}_2}$ and $1_{\mathbb{G}_3}$ are the identity elements of the groups $\mathbb{G}_1$, $\mathbb{G}_2$ and $\mathbb{G}_3$, respectively. Assume that Discrete Logarithm Problem (DLP) is hard in both $\mathbb{G}_1$ and $\mathbb{G}_2$ (i.e., given a random $y \in \mathbb{G}_1$ (or $\in \mathbb{G}_2$), it is computationally infeasible to find an integer $x \in \mathbb{Z}$ such that $y = g^x$). If it is clear from the context we write 0 for the identity elements of $\mathbb{G}_1$, $\mathbb{G}_2$ and 1 for $\mathbb{G}_3$. A *bilinear map* is a map $e: \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_3$ satisfying the following properties [15]:

- **Bilinearity:** For all $P_1, Q_1 \in \mathbb{G}_1, P_1', Q_1' \in \mathbb{G}_2$, e is a group homomorphism in each component, i.e.
  1) $e(P_1 + Q_1, P_1') = e(P_1, P_1') \cdot e(Q_1, P_1')$,
  2) $e(P_1, P_1' + Q_1') = e(P_1, P_1') \cdot e(P_1, Q_1')$.
- **Non-degeneracy:** e is non-degenerate in each component, i.e.,
  1) For all $P \in \mathbb{G}_1$, $P \neq 0$, there is an element $Q \in \mathbb{G}_2$ such that $e(P, Q) \neq 1$,
  2) For all $Q \in \mathbb{G}_2$, $Q \neq 0$, there is an element $P \in \mathbb{G}_1$ such that $e(P, Q) \neq 1$.
- **Computability:** There exists an algorithm which computes the bilinear map e efficiently.

B. Algorithm 1

1) Precomputations: Like all existing outsourcing algorithms, some precomputations are performed to speed up the proposed algorithms following the method of [13]. It includes a static table ST and a dynamic table DT. The values stored in the dynamic table are replaced while they are used, and then the table is reconstructed in an idle time of the device. We next describe the steps of the Rand1 algorithm to generate random group elements.

Rand1

- **Preprocessing Step:** Let P_1 and P_2 be generators of $\mathbb{G}_1$ and $\mathbb{G}_2$, respectively. Generate n random elements $\alpha_1, \cdots, \alpha_n \in \mathbb{Z}/q\mathbb{Z}$. For $j = 1, \cdots, n$ compute $\beta_{j_1} = \alpha_j \cdot P_1$ and $\beta_{j_2} = \alpha_j \cdot P_2$, and store the values of α_j, β_{j_1} and β_{j_2} in ST. Compute $e(P_1, P_2) \in \mathbb{G}_3$ and store it in ST.
- **Generation of Precomputed Values:** A new entry in DT is computed as follows: Generate randomly $S \in \{1, \cdots, n\}$ such that $|S| = k$. For each $j \in S$, select randomly $K_j \in \{1, \cdots, h-1\}$, where $h > 1$ is a small integer. Compute

$$x_1 \equiv \sum_{j \in S} \alpha_j K_j \bmod q.$$

If $x_1 \equiv 0 \bmod q$, start again. Otherwise, compute

$$x_1 \cdot P_1 \equiv \sum_{j \in S} K_j \cdot \beta_{j_1} \bmod q.$$

Following the above procedure, compute similarly the elements $(x_2, x_2 \cdot P_2), (x_3, x_3 \cdot P_1)$, and $(x_4, x_4 \cdot P_2)$. Then compute

1) $2x_1 \cdot P_1$,
2) $-2x_2 \cdot P_2$,
3) $e(P_1, P_2)^{2x_1 x_2}$,
4) $e(P_1, P_2)^{x_3 x_4}$.

The entry

$$(x_1 \cdot P_1, 2x_1 \cdot P_1, x_3 \cdot P_1, x_2 \cdot P_2, -2x_2 \cdot P_2,$$

$$x_4 \cdot P_2, e(P_1, P_2)^{2x_1 x_2}, e(P_1, P_2)^{x_3 x_4})$$

is stored in DT. On each invocation of Rand1, an entry is returned and removed from DT. Further, a new set of values is used as fresh random values.

2) Proposed Algorithm 1: Our algorithm takes $A \in \mathbb{G}_1$, $B \in \mathbb{G}_2$ as inputs and produces $e(A, B)$ as output. In what follows, T denotes a trusted device with limited computation resources, and $U_i(A, B) \to e(A, B), i \in \{1, 2\}$ denotes party U_i taking (A, B) as inputs and returning $e(A, B)$ as output.

- **Initialization:** T calls Rand1 to get random values

$$(x_1 \cdot P_1, 2x_1 \cdot P_1, x_3 \cdot P_1, x_2 \cdot P_2, -2x_2 \cdot P_2,$$

$$x_4 \cdot P_2, \lambda = e(P_1, P_2)^{2x_1 x_2}, e(P_1, P_2)^{x_3 x_4}).$$

- **Computation:** In random orders, T sends the following values to U_1
  1) $U_1(A + 2x_1 \cdot P_1, -B - 2x_2 \cdot P_2) \to \alpha_1$,
  2) $U_1(x_3 \cdot P_1, x_4 \cdot P_2) \to \alpha_1'$.

  Similarly, in random orders, T sends the following values to U_2
  1) $U_2(A + x_1 \cdot P_1, B + x_2 \cdot P_2) \to \alpha_2$,
  2) $U_2(x_3 \cdot P_1, x_4 \cdot P_2) \to \alpha_2'$.
- **Recover:** T checks whether $\alpha_1' \stackrel{?}{=} \alpha_2' \stackrel{?}{=} e(P_1, P_2)^{x_3 x_4}$. If the verifications are successful then it computes

$$\beta = \alpha_1 \alpha_2^2 \lambda$$

and produces β as output. Otherwise, it rejects and gives an "Error".

3) Security Analysis:

Theorem 6: Under the OMTUP assumption, the algorithms $(T,(U_1,U_2))$ of "Algorithm 1" are an outsource-secure implementation of a pairing evaluation, where the input (A,B) may be honest, secret, honest, protected, or adversarial, protected.

Proof: The correctness follows easily by bilinear property of e:

$$\begin{aligned}\beta &= \alpha_1 \cdot \alpha_2^2 \cdot \lambda \\ &= e(A + 2x_1 \cdot P_1, -B - 2x_2 \cdot P_2) \cdot \\ &\quad e(A + x_1 \cdot P_1, B + x_2 \cdot P_2)^2 \cdot e(P_1, P_2)^{2x_1x_2} \\ &= e(A,B)^{-1} \cdot e(P_1,P_2)^{-4x_1x_2} \cdot e(A,B)^2 \cdot \\ &\quad e(P_1,P_2)^{2x_1x_2} \cdot e(P_1,P_2)^{2x_1x_2} \\ &= e(A,B).\end{aligned}$$

We next prove the security of the algorithm. Let (E,U_1',U_2') be a PPT adversary that interacts with a PPT algorithm T in the OMTUP model.

Pair One: $EVIEW^i_{real}$~$EVIEW_{ideal}$ (The external adversary, E, learns nothing.):

For a round i, if the input (A,B) is other than secret, honest, (i.e., honest, protected or adversarial, protected) the simulator S_1 behaves as in the real round. S_1 never requires to access the secret input, since there is none. So suppose that the input (A,B) is honest, secret. In that case, the simulator S_1 behaves as follows: On receiving input in the ith round, S_1 ignores it and instead make random queries to the servers U_1',U_2':

- $U_1'(x_1 \cdot P_1, x_2 \cdot P_2) \rightarrow \alpha_1$,
- $U_1'(x_3 \cdot P_1, x_4 \cdot P_2) \rightarrow \alpha_1'$,
- $U_2'(x_5 \cdot P_1, x_6 \cdot P_2) \rightarrow \alpha_2$,
- $U_2'(x_3 \cdot P_1, x_4 \cdot P_2) \rightarrow \alpha_2'$.

After getting responses from U_1' and U_2', S_1 checks:

- If $\alpha_1' \neq \alpha_2'$ or $e(x_3 \cdot P_1, X_4 \cdot P_2) \neq \alpha_1'$, S_1 produces $Y_p^i =$"Error", $Y_u^i = \emptyset$ and $rep^i = 1$.
- If all responses are correct, S_1 sets $Y_p^i = \emptyset, Y_u^i = \emptyset$, and $rep^i = 0$.
- Otherwise, S_1 selects a random value $s_r \in \mathbb{G}_3$ and sets $Y_p^i = s_r, Y_u^i = \emptyset$ and $rep^i = 1$.

For all cases, S_1 saves the appropriate states.

The input distributions in the real and ideal experiments are computationally indistinguishable for U_1' and U_2'. The inputs to U_1' and U_2' are chosen uniformly at random in the ideal experiment. In a real experiment, each part of all queries that T makes to any one program in the computation step is independently re-randomized, and the re-randomization factors(i.e., outputs of Rand1) are either truly randomly generated by naive table-lookup approach or computationally indistinguishable from random by the assumption of the EBPV generator [21]. Now, there are three possibilities to consider.

- If U_1' and U_2' behave honestly in the round i, S_1 gives the correct output, using Alg, which is the same as the output of T^{U_1,U_2}.
- If one of (U_1',U_2') give an incorrect output in the ith round and it has been detected by both T and S_1 with probability $1/2$, then it will result in an "Error".
- Finally, we consider the case, where one of (U_1',U_2') give an incorrect output in the ith round and it is not caught with probability $1/2$. In the real experiment,the two outputs generated by (U_1',U_2') are multiplied together along with a random value λ.

Thus, a corrupted output looks random to the environment E, in the real experiment, S_1 also simulates with a random value in $\mathbb{G}_3$ as the output. So, $EVIEW^i_{real}$~$EVIEW^i_{ideal}$ even when one of (U_1',U_2') is dishonest. Now, by the hybrid argument, we conclude that $EVIEW_{real}$~$EVIEW_{ideal}$.

Pair Two: $UVIEW_{real}$~$UVIEW_{ideal}$ (The untrusted software, (U_1,U_2), learns nothing.):

Here, regardless of the input type, the simulator S_2 always behaves the same way. Upon receiving an input on round i, S_2 ignores it and instead makes four random queries to U_1' and U_2'. Then S_2 saves its own state and the states of (U_1',U_2'). E can easily distinguish between these real and ideal experiments (output of the ideal experiment is never corrupted), but we want to show that E cannot share this information with (U_1',U_2'). This happens because in the ith round of the real experiment, T always re-randomizes the inputs to (U_1',U_2'), and in the ideal experiment S_2 creates random, independent queries for (U_1',U_2'). So, for each round i, we have $UVIEW^i_{real}$~$UVIEW^i_{ideal}$. Then, by the hybrid argument, we get the desired result $UVIEW_{real}$~$UVIEW_{ideal}$.

Theorem 7: The algorithms $(T,(U_1,U_2))$ of "Algorithm 1" are a $(\mathcal{O}(\frac{1}{\log q}),1/2)$-outsource secure implementation of a pairing evaluation under the OMTUP assumption.

Proof: By the above theorem, U_1 and U_2 cannot distinguish a test query from a real query. Without loss of generality, assume that U_1 is honest while U_2 is dishonest (since we are under the OMTUP assumption). Thus, U_2 fails with a probability $1/2$.

C. Algorithm 2

1) Precomputations: Rand2

- **Preprocessing Step:** Generate n random elements $\alpha_1,\cdots,\alpha_n \in \mathbb{Z}_q$. For $j = 1,\cdots,n$ compute $\beta_{j_1} = \alpha_j \cdot P_1$ and $_{j_2} = \alpha_j \cdot P_2$, and store the values of α_j, β_{j_1} and β_{j_2} in a static table ST. Compute $e(P_1,P_2)$ and store the value in ST.
- **Generation of Precomputed Values:** When a table DT needs a new entry, it is produced as follows. Randomly generate $S \in \{1,\cdots,n\}$ such that $|S| = k$. For each $j \in S$, randomly select $K_j \in \{1,\cdots,h-1\}$, where $h > 1$ is a small integer. Compute

$$x_1 \equiv \sum_{j \in S} \alpha_j K_j \bmod q.$$

If $x_1 \equiv 0 \bmod q$, start again. Otherwise, compute

$$x_1 P_1 \equiv \sum_{j \in S} K_j \cdot \beta_{j_1} \bmod q.$$

Following the above procedure, compute similarly the elements $(x_2, x_2 \cdot P_2), (x_3, x_3 \cdot P_1), (x_4, x_4 \cdot P_1)$, and $(x_5, x_5 \cdot P_2)$. Then compute

	Algoritm [13]	Algorithm 1	Algorithm 2
SM	3	2	1
ME	2	2	1
PA	5(k+h-3)	4(k+h-3)	5(k+h-3)

TABLE I
COMPARISON OF PRECOMPUTATION

	Algoritm [13]	Algorithm 1	Algorithm 2
PA	4	4	4
FM	3	3	2

TABLE II
COMPARISON OF CLIENT'S WORKLOAD

1) $(x_2 - x_1^{-1}x_2x_3) \cdot P_2$,
2) $e(P_1, P_2)^{x_4x_5}$.

The entry

$$(x_1 \cdot P_1, x_3 \cdot P_1, x_4 \cdot P_1, x_2 \cdot P_2, x_5 \cdot P_2,$$

$$(x_2 - x_1^{-1}x_2x_3) \cdot P_2, e(P_1, P_2)^{x_4x_5})$$

is stored in DT. On each invocation of Rand2, an entry is returned and removed from DT. Further, a new set of values is used as fresh random values.

2) Proposed Scheme 2: Algorithm 2 takes $A \in \mathbb{G}_1$, $B \in \mathbb{G}_2$ as inputs and produces $e(A, B)$ as output. In what follows, T denotes a trusted device with limited computation resources, and $U_i(A, B) \rightarrow e(A, B), i \in \{1, 2\}$ denotes party U_i taking (A, B) as inputs and giving $e(A, B)$ as output.

- **Init:** T calls Rand2 to get random values

$$(x_1 \cdot P_1, x_3 \cdot P_1, x_4 \cdot P_1, x_2 \cdot P_2, x_5 \cdot P_2,$$

$$(x_2 - x_1^{-1}x_2x_3) \cdot P_2, e(P_1, P_2)^{x_4x_5}).$$

- **Computation:** In random orders, T sends the following to U_1
  1) $U_1(A + x_1 \cdot P_1, B + x_2 \cdot P_2) \rightarrow \alpha_1$,
  2) $U_1(x_4 \cdot P_1, x_5 \cdot P_2) \rightarrow \alpha_1'$.

  Similarly, in random orders, T sends the following to U_2
  1) $U_2(A + x_3 \cdot P_1, -x_2 \cdot P_2) \rightarrow \alpha_2$,
  2) $U_2(-x_1 \cdot P_1, B + (x_2 - x_1^{-1}x_2x_3) \cdot P_2) \rightarrow \alpha_3$
  3) $U_2(x_4 \cdot P_1, x_5 \cdot P_2) \rightarrow \alpha_2'$.
- **Recover:** T checks whether $\alpha_1' \stackrel{?}{=} \alpha_2' \stackrel{?}{=} e(P_1, P_2)^{x_4x_5}$. If the verifications are successful then it computes

$$\beta = \alpha_1\alpha_2\alpha_3$$

and produces β as output. Otherwise, it rejects and gives an "Error".

3) Security Analysis:

Theorem 8: Under the OMTUP assumption, the algorithms $(T, (U_1, U_2))$ of "Algorithm 1" are an outsource-secure implementation of a pairing evaluation, where the input (A, B) may be honest, secret, honest, protected, or adversarial, protected.

	Algoritm A [13]	Algorithm 1	Algorithm 2
PC	6	4	5

TABLE III
COMPARISON OF SERVER'S WORKLOAD

	Algoritm A [13]	Algorithm 1	Algorithm 2
PC	$\approx$ 0,117 KB	$\approx$ 0,078 KB	0,098 KB

TABLE IV
COMPARISON OF COMMUNICATION OVERHEAD FOR 80-BIT SECURITY

Proof: The correctness is straight forward:

$$\begin{aligned}\beta &= \alpha_1 \cdot \alpha_2 \cdot \alpha_3 \\ &= e(A + x_1 \cdot P_1, B + x_2 \cdot P_2) \cdot e(A + x_3 \cdot P_1, -x_2 \cdot P_2) \cdot \\ &e(-x_1 \cdot P_1, B + (x_2 - x_1^{-1}x_2x_3) \cdot P_2) \\ &= e(A, B)e(P_1, P_2)^{x_1x_2} \cdot e(P_1, P_2)^{-x_2x_3} \cdot \\ &e(P_1, P_2)^{-x_1(x_2 - x_1^{-1}x_2x_3)} \\ &= e(A, B).\end{aligned}$$

The proof of the security part follows analogously to the proof of Theorem 6.

IV. COMPLEXITY ANALYSIS

A. Comparisons

We compare the precomputation algorithms of our proposed schemes with Tian *et al.*'s algorithm [13]. In the following tables, SM denotes scalar multiplication in $\mathbb{G}_1, \mathbb{G}_2$, ME modular exponentiation, PC pairing computation on the server side, FM field multiplication, and PA point addition in $\mathbb{G}_1, \mathbb{G}_2$. Furthermore, k denotes the size of the set S in the algorithms Rand1 and Rand2. Table I compares the precomputation, Table II compares the client's workload, and Table III compares the server's workload. Table IV compares the communication overhead between the client and the servers, and finally Table V gives the memory requirements for ST and DT by means of counting the number of group elements.

V. CONCLUSION

In this paper, we studied outsourcing the computation of bilinear maps and proposed two new efficient algorithms decreasing both the memory requirement and the overall communication overhead. We defined the necessary security model, and proved the correctness and the security of the proposed secure outsourcing algorithms. We further gave the comparisons of our algorithms with a very recent outsourcing mechanism of Tian *et al.* [13] with respect to the offline and online computations, and the memory to be used. In this

	ST	DT
Algorithm 1	3	8
Algorithm 2	3	7

TABLE V
MEMORY REQUIREMENTS FOR Rand ALGORITHMS.

way, we show that our algorithms are more efficient than all previously proposed solutions.

ACKNOWLEDGMENT

Kiraz's and Arabaci's works are supported by a grant from Ministry of Development of Turkey provided to the Cloud Computing and Big Data Research Lab Project. Uzunkol's work is supported by the project (114C027) funded by EU FP7-The Marie Curie Action and TÜBİTAK (2236-CO-FUNDED Brain Circulation Scheme).

REFERENCES

[1] D. Boneh and M. Franklin, "Identity-based encryption from the weil pairing," in *Advances in Cryptology CRYPTO 2001*, ser. Lecture Notes in Computer Science, J. Kilian, Ed. Springer Berlin Heidelberg, 2001, vol. 2139, pp. 213–229. [Online]. Available: http://dx.doi.org/10.1007/3-540-44647-8_13

[2] D. Boneh, B. Lynn, and H. Shacham, "Short signatures from the weil pairing," *Journal of Cryptology*, vol. 17, no. 4, pp. 297–319, 2004. [Online]. Available: http://dx.doi.org/10.1007/s00145-004-0314-9

[3] A. Joux, "A one round protocol for tripartite diffiehellman," *Journal of Cryptology*, vol. 17, no. 4, pp. 263–276, 2004. [Online]. Available: http://dx.doi.org/10.1007/s00145-004-0312-y

[4] P. Barreto, S. Galbraith, C. higeartaigh, and M. Scott, "Efficient pairing computation on supersingular abelian varieties," *Designs, Codes and Cryptography*, vol. 42, no. 3, pp. 239–271, 2007. [Online]. Available: http://dx.doi.org/10.1007/s10623-006-9033-6

[5] J.-L. Beuchat, J. Gonzlez-Daz, S. Mitsunari, E. Okamoto, F. Rodrguez-Henrquez, and T. Teruya, "High-speed software implementation of the optimal ate pairing over barretonaehrig curves," in *Pairing-Based Cryptography - Pairing 2010*, ser. Lecture Notes in Computer Science, M. Joye, A. Miyaji, and A. Otsuka, Eds. Springer Berlin Heidelberg, 2010, vol. 6487, pp. 21–39. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-17455-1_2

[6] F. Hess, N. Smart, and F. Vercauteren, "The eta pairing revisited," *Information Theory, IEEE Transactions on*, vol. 52, no. 10, pp. 4595–4602, Oct 2006.

[7] N. Koblitz and A. Menezes, "Pairing-based cryptography at high security levels," in *Cryptography and Coding*, ser. Lecture Notes in Computer Science, N. Smart, Ed. Springer Berlin Heidelberg, 2005, vol. 3796, pp. 13–36. [Online]. Available: http://dx.doi.org/10.1007/11586821_2

[8] M. Scott, N. Costigan, and W. Abdulwahab, "Implementing cryptographic pairings on smartcards," in *Cryptographic Hardware and Embedded Systems - CHES 2006*, ser. Lecture Notes in Computer Science, L. Goubin and M. Matsui, Eds. Springer Berlin Heidelberg, 2006, vol. 4249, pp. 134–147. [Online]. Available: http://dx.doi.org/10.1007/11894063_11

[9] F. Hess, "Pairing lattices," in *Pairing-Based Cryptography Pairing 2008*, ser. Lecture Notes in Computer Science, S. Galbraith and K. Paterson, Eds. Springer Berlin Heidelberg, 2008, vol. 5209, pp. 18–38. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-85538-5_2

[10] S. Hohenberger and A. Lysyanskaya, "How to securely outsource cryptographic computations," in *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005, Proceedings*, ser. Lecture Notes in Computer Science, vol. 3378. Springer, 2005, pp. 264–282. [Online]. Available: http://www.iacr.org/cryptodb/archive/2005/TCC/3678/3678.pdf

[11] T. Matsumoto, K. Kato, and H. Imai, "Speeding up secret computations with insecure auxiliary devices," in *Advances in Cryptology CRYPTO 88*, ser. Lecture Notes in Computer Science, S. Goldwasser, Ed. Springer New York, 1990, vol. 403, pp. 497–506. [Online]. Available: http://dx.doi.org/10.1007/0-387-34799-2_35

[12] B. Chevallier-Mames, J.-S. Coron, N. McCullagh, D. Naccache, and M. Scott, "Secure delegation of elliptic-curve pairing," Cryptology ePrint Archive, Report 2005/150, 2005.

[13] H. Tian, F. Zhang, and K. Ren, "Secure bilinear pairing outsourcing made more efficient and flexible," in *Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security, ASIA CCS '15, Singapore, April 14-17, 2015*, F. Bao, S. Miller, J. Zhou, and G. Ahn, Eds. ACM, 2015, pp. 417–426. [Online]. Available: http://doi.acm.org/10.1145/2714576.2714615

[14] X. C. 0001, W. Susilo, J. L. 0002, D. S. Wong, J. Ma, S. Tang, and Q. Tang, "Efficient algorithms for secure outsourcing of bilinear pairings." *Theor. Comput. Sci.*, vol. 562, pp. 112–121, 2015. [Online]. Available: http://dblp.uni-trier.de/db/journals/tcs/tcs562.html#ChenSLWMTT15

[15] I. Blake, G. Seroussi, N. Smart, and J. W. S. Cassels, *Advances in Elliptic Curve Cryptography (London Mathematical Society Lecture Note Series)*. New York, NY, USA: Cambridge University Press, 2005.

[16] S. D. Galbraith, K. G. Paterson, and N. P. Smart, "Pairings for cryptographers," *Discrete Applied Mathematics*, vol. 156, no. 16, pp. 3113 – 3121, 2008, applications of Algebra to Cryptography. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166218X08000449

[17] B. G. Kang, M. S. Lee, and J. H. Park, "Efficient delegation of pairing computation," 2005.

[18] S. Canard, J. Devigne, and O. Sanders, "Delegating a pairing can be both secure and efficient," in *Applied Cryptography and Network Security*, ser. Lecture Notes in Computer Science, I. Boureanu, P. Owesarski, and S. Vaudenay, Eds. Springer International Publishing, 2014, vol. 8479, pp. 549–565. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-07536-5_32

[19] P. Tsang, S. Chow, and S. Smith, "Batch pairing delegation," in *Advances in Information and Computer Security*, ser. Lecture Notes in Computer Science, A. Miyaji, H. Kikuchi, and K. Rannenberg, Eds. Springer Berlin Heidelberg, 2007, vol. 4752, pp. 74–90. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-75651-4_6

[20] S. S. Chow, M. H. Au, and W. Susilo, "Server-aided signatures verification secure against collusion attack," *Information Security Technical Report*, vol. 17, no. 3, pp. 46 – 57, 2013, security and Privacy for Digital Ecosystems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1363412712000489

[21] P. Nguyen, I. Shparlinski, and J. Stern, "Distribution of modular sums and the security of the server aided exponentiation," in *Cryptography and Computational Number Theory*, ser. Progress in Computer Science and Applied Logic. Birkhuser Basel, 2001, vol. 20, pp. 331–342.

Efficient Modular Exponentiation Methods for RSA

Hatice Kübra Güner, Murat Cenk and Çağdaş Çalık

***Abstract*—RSA is a commonly used asymmetric key cryptosystem that is used in encrypting and signing messages. The efficiency of the implementation is an important factor in effectively using the system. The RSA algorithm heavily depends on the modular exponentiation operation on large integers. A drawback of this system is that it becomes inefficient so quickly when the parameters are adjusted to increase security. This situation causes the operations to be performed with large numbers. Therefore, implementations require the utilization of faster methods than the traditional ones. One popular modular exponentiation method is the *repeated squaring and multiplication algorithm*. In this study, we examine some of the modular exponentiation algorithms and implement them for comparison with the *repeated squaring and multiplication algorithm*. The results suggest that particular cases of studied methods provide at least 23% improvement over the *repeated squaring and multiplication algorithm*.**

***Index Terms*—RSA, modular exponentiation, m-ary method, efficiency, running time.**

I. Introduction

In RSA cryptosystem, encryption and decryption procedures are based on modular exponentiation operation. According to the RSA algorithm [**?**], the encryption process consists of first representing the plaintext as an integer *M* and then the ciphertext *C* is computed by

$$C = M^e \bmod n.$$

For decryption, the process is the same but with a different exponent, that is, the original message is obtained by

$$M = C^d \bmod n.$$

By looking at these procedures, it is obvious that modular exponentiation operation is at the core of the RSA cryptosystem and efficiency of the system on a particular platform heavily depend on how the modular exponentiation operation is implemented on that platform.

Over the years, the recommended key sizes for RSA have changed because of the increased computation power and the improved methods the cryptanalysts have [**?**]. Today, the minimum accepted RSA modulus length is 2048-bits.

The *repeated squaring and multiplication algorithm* was recommended by Rivest, Shamir and Adleman in [**?**]. Running time of the modular exponentiation increases rapidly when the key size is doubled. In TABLE I, running times of the RSA decryption operation for different modulus sizes are given. The decryption exponent *d* is calculated by choosing the encryption exponent *e* as 65537.

According to the table, if the modulus size is increased from 1024-bits to 2048-bits, running time of decryption increases by a factor of more than 7. If the key modulus size is increased to 3072-bits, the running time increases by a factor of almost 22 times compared to 1024-bits. When the modulus size is increased from 1024-bits to 4096-bits, the running time increases by a factor of 47. One of the way to overcome this fast increase is to find faster modular exponentiation methods . With this motivation, we examine some methods and compute required number of operations for each of them separately. Generally, the existing methods provide efficiency with using memory. So, if there was any memory usage in the method, required memory sizes are also tabulated. These methods are implemented using *MPIR* library with *Microsoft Visual Studio* on Intel Core i7 2.00 GHz., and the execution times were calculated for the *repeated squaring and multiplication algorithm*. In some cases, significant efficiency improvements were observed.

TABLE I
Execution time of RSA decryption operation for various modulus sizes.

RSA modulus size	running time (ms)
1024-bits	4.52
2048-bits	32.55
3072-bits	99.46
4096-bits	214.96

H. K. Güner, M. Cenk and Ç. Çalık are with the Cryptography Program, Institute of Applied Mathematics, METU, Ankara, Turkey. e-mail: kubra.guner@metu.edu.tr, mcenk@metu.edu.tr, ccalik@metu.edu.tr

II. Some Fast Modular Exponentiation Methods

A. The Repeated Squaring and Multiplication Algorithm

The *repeated squaring and multiplication algorithm* was examined by Knuth in [**?**]. This method has an old history [**?**]. There are two variants of the algorithm; the *left-to-right* method and the *right-to-left* method. Here, we only examine the *left-to-right* method in which the procedure starts from the most significant bit of binary expansion of the exponent. Let us represent the binary expansion of private exponent *d* as

$$d = (d_k d_{k-1} \dots d_1 d_0)_2 = \sum_{i=0}^{k} d_i \cdot 2^i \quad \text{,where } d_i \in \{0, 1\}$$

In this algorithm, one squaring is performed in each step, and after that if the corresponding bit is 1, a successive multiplication is also carried out. The pseudo-code is presented in Algorithm 1.

The required number of operations [**?**] is

$$(\lceil log(d) \rceil - 1)S + (H(d) - 1)M$$

where *H(d)* is the Hamming weight of *d*, S represents the squaring operation and M represents the multiplication operation. The cost of exponentiation for different scenarios are given as:

- Worst case: $(\lceil log(d) \rceil - 1)$(S + M)

Algorithm 1 *Left-to-right* method of the *repeated squaring and multiplication algorithm*

Input: $M,\ n,\ (d_k d_{k-1} \dots d_1 d_0)_2$
Output: $C = M^d \bmod n$
1: **if** $d_k = 1$ **then**
2: $C \leftarrow M$
3: **else**
4: $C \leftarrow 1$
5: **end if**
6: **for** i **from** $k-1$ **to** 0 **do**
7: $C \leftarrow C^2 \bmod n$
8: **if** $d_i = 1$ **then**
9: $C \leftarrow C \cdot M \bmod n$
10: **end if**
11: **end for**
12: **return** C

- Average case: $(\lceil log(d) \rceil - 1)$(S + $\frac{1}{2}$M)
- Best case: $(\lceil log(d) \rceil - 1)$S

The number of multiplications and squarings are calculated separately because the squaring operation can be implemented more efficiently than the multiplication operation [**?**].

B. The m-ary Method

The *m-ary* method is explained in [**?**]. Idea of the method is the same with the *left-to-right* method of the *repeated squaring and multiplication algorithm* with addition of a look-up table. A look-up table is prepared prior to the exponentiation to keep powers of *M* such as $M^i \bmod n$ where $i \in \{2, 3, \cdots, m-1\}$. In this method, m-th power of the partial value is computed and if the related digit value of the expansion is different from 0, a subsequent multiplication is also applied in every step. Here, multiplier is obtained from the look-up table by looking at value of the corresponding digit. Base *m* representation of private key *d* is

$$d = (d_l d_{l-1} \dots d_1 d_0)_m = \sum_{i=0}^{l} d_i \cdot m^i,$$

where $d_i \in \{0, 1, \dots, m-1\}$. The pseudo-code is given as Algorithm 2.

The method can be used for any *m* values. But, the more efficient results are obtained when *m* is chosen as power of 2 [**?**]. Here we only focus on the values of *m* in the form $m = 2^r$ and assume that *d* is a k-bit integer. The required number of operations for this method is [**?**]

- For precomputation: S + $(m-3)$M
- For powering: $(\frac{k}{r} - 1) \cdot r$S
- For multiplication: $\frac{m-1}{m} \cdot (\frac{k}{r} - 1)$M

For each RSA modulus size, there is a particular value of *m* for which the method requires the minimum multiplications and squarings [**?**]. In TABLE II, the number of operations are calculated, and optimal *m* values are defined.

By these results, *m*=32 provides the least number of operations for 1024-bits modulus size. For 2048-bits modulus size, 64 is the best option. Taking *m*=128 gives the most efficient results for 3072-bits and 4096-bits modulus sizes.

Algorithm 2 *m-ary* method

Input: $M,\ n,\ (d_l d_{l-1} \dots d_1 d_0)_m$
Output: $M^d \bmod n$
1: $pc[0] \leftarrow M^2 \bmod n$
2: **for** i **from** 1 **to** $m-3$ **do**
3: $pc[i] \leftarrow pc[i-1] \cdot M \bmod n$ { for look-up table}
4: **end for**
5: **if** $d_l = 0$ **then**
6: $C \leftarrow 1$
7: **else**
8: **if** $d_l = 1$ **then**
9: $C \leftarrow M$
10: **else**
11: $C \leftarrow pc[d_l - 2]$
12: **end if**
13: **end if**
14: **for** i **from** $l-1$ **to** 0 **do**
15: $C \leftarrow C^m \bmod n$
16: **if** $d_i = 1$ **then**
17: $C \leftarrow C \cdot M \bmod n$
18: **else**
19: **if** $d_i > 1$ **then**
20: $C \leftarrow C \cdot pc[d_i - 2] \bmod n$
21: **end if**
22: **end if**
23: **end for**
24: **return** C

TABLE II
AVERAGE NUMBER OF OPERATIONS

m	1024-bit	2048-bit	3072-bit	4096-bit
2	1023·S+512·M	2047·S+1024·M	3071·S+1536·M	4095·S+2048·M
4	1023·S+384·M	2047·S+768·M	3071·S+1152·M	4095·S+1536·M
8	1022·S+303·M	2046·S+604·M	3070·S+900·M	4094·S+1199·M
16	1021·S+252·M	2045·S+492·M	3069·S+732·M	4093·S+972·M
32	**1020·S+226·M**	2044·S+425·M	3068·S+588·M	4092·S+822·M
64	1019·S+228·M	**2043·S+396·M**	3067·S+564·M	4091·S+732·M
128	1018·S+269·M	2042·S+414·M	**3066·S+559·M**	**4090·S+705·M**
256	1017·S+380·M	2041·S+507·M	3065·S+635·M	4089·S+762·M

With a suitable choice of *m*, the *m-ary* method provides an improvement over the *repeated squaring and multiplication algorithm*. It should be noted that, additional memory is used to store the look-up table. Required memory sizes are given for different *m* values in TABLE III.

TABLE III
MEMORY USAGE-KB (1 KB = 1024-BYTE)

	4	8	16	32	64	128	256
1024-bit	0.25	0.75	1.75	3.75	7.75	15.75	31.75
2048-bit	0.5	1.5	3.5	7.5	15.5	31.5	63.5
3072-bit	0.75	2.25	5.25	11.25	23.25	47.25	95.25
4096-bit	1	3	7	15	31	63	127

C. The Modified m-ary Method

With the *m-ary* method, required number of operations decreases by considerable amounts for increasing values of

m. On the other hand, number of precomputation operations increases nearly twice when m is doubled. Basically, rapid increases in the required number of operations for large m values are caused by precomputation operations. The *modified m-ary* method decreases these multiplications by a factor of 2 with some modifications in the *m-ary* method. The idea stems from the fact that even powers of a number can be calculated with odd parts of the exponentiation. For instance, we can perform $(M^3)^2 \bmod n$ instead of $M^6 \bmod n$. The algorithm [**?**], [**?**] is given in Algorithm 3.

Algorithm 3 *modified m-ary* method

Input: $M,\ n,\ (d_l d_{l-1} \cdots d_0)_m$
Output: $M^d \bmod n$
1: $pc[0] \leftarrow M^3 \bmod n$
2: **for** i **from** 1 **to** $\frac{m-4}{2}$ **do**
3: $pc[i] \leftarrow pc[i-1] \cdot M^2 \bmod n$ {look-up table multiplications}
4: **end for**
5: **if** $d_l = 0$ **then**
6: $C \leftarrow 1$
7: **else**
8: $t \leftarrow 1$
9: **while** $(d_l \bmod 2) = 0$ **do**
10: $d_l \leftarrow d_l/2$
11: $t \leftarrow 2 \cdot t$
12: **end while**
13: **if** $d_l = 1$ **then**
14: $C \leftarrow M$
15: **else**
16: $C \leftarrow pc[\frac{d_l-3}{2}]$
17: **end if**
18: $C \leftarrow C^t \bmod n$
19: **end if**
20: **for** i **from** $l-1$ **to** 0 **do**
21: **if** $d_i = 0$ **then**
22: $C \leftarrow C^m \bmod n$
23: **else**
24: $t \leftarrow 1$
25: **while** $(d_i \bmod 2) = 0$ **do**
26: $d_i \leftarrow d_i/2$
27: $t \leftarrow 2 \cdot t$
28: **end while**
29: $C \leftarrow C^{\frac{m}{t}} \bmod n$
30: **if** $d_i = 1$ **then**
31: $C \leftarrow C \cdot M \bmod n$
32: **else**
33: $C \leftarrow C \cdot pc[\frac{d_i-3}{2}] \bmod n$
34: **end if**
35: $C \leftarrow C^t \bmod n$
36: **end if**
37: **end for**
38: **return** C

Average number of required operations is:

- For precomputation: S + $\frac{m-2}{2}$·M
- For powering: $(\frac{k}{r} - 1) \cdot r$·S
- For multiplication: $\frac{m-1}{m} \cdot (\frac{k}{r} - 1)$·M
- For division by 2: $\frac{k}{m \cdot r} \cdot (m - r - 1)$(Division by 2)

For the *modified m-ary method*, the required memory sizes are in TABLE IV . As we see from the table, the required memory sizes decrease to exactly half. Therefore, running time of preparation of the look-up table decreases to half.

TABLE IV
MEMORY USAGE-KB (1 KB=1024-BYTE)

	4	8	16	32	64	128	256
1024-bit	0.125	0.375	0.875	1.875	3.875	7.875	15.875
2048-bit	0.25	0.75	1.75	3.75	7.75	15.75	31.75
3072-bit	0.375	1.125	2.625	5.625	11.625	23.625	47.625
4096-bit	0.5	1.5	3.5	7.5	15.5	31.5	63.5

D. Reducing Precomputation Multiplications

Especially with small exponents, all base m numerals are not expected to be seen in the base m expansion of the exponent. In these situations, calculating all look-up table values becomes unnecessary. With the *reducing precomputation multiplications*, we ignore these unnecessary precomputations and reduce the number of required multiplications.

The *reducing precomputation multiplications* cannot be applied for large exponents since all base m numerals are expected to be seen in the expansion. Therefore, this method cannot be used in decryption effectively. But, it allows considerable improvements in encryption when the public key exponent e = 65537 [**?**] is used.

III. IMPLEMENTATION RESULTS OF STUDIED METHODS

In this section, implementation results are presented. These results are obtained using the *MPIR* library with *Microsoft Visual Studio* on Intel Core i7 2.00 GHz. Running times of these methods are compared with the *repeated squaring and multiplication algorithm* and the most efficient cases are noted. Encryption and decryption steps were discussed separately.

Aim of this study is to find more efficient methods than the *repeated squaring and multiplication algorithm* for modular exponentiation. Thus, the RSA cryptosystem parameters were chosen as in [**?**]. In the following tables, there are modulus size, m value, running time, comparison with m=2 and saving columns. With saving column, we intend to express obtained improvement over percent. If the obtained result is worse than m=2 case, then we express them with minus sign.

A. Encryption

The public key e was taken as 65537 in all experiments.

The *m-ary* method's running time results for different plaintext sizes are tabulated in TABLE V. According to the information in the table, the *m-ary* method is not suitable for encryption with e = 65537. Running time increases rapidly when m grows. The reason behind this increase is due to the computation of a larger look-up table. But, only one multiplication is needed or there is no need for any multiplication during the procedure. This means that all of the look-up table multiplications are never used. For example, if we take m

TABLE V
RUNNING TIME RESULTS FOR THE M-ARY METHOD ON ENCRYPTION

M	*m*	running time (ms)	comparison with *m*=2	saving (%)
1024-bit	2	0.22	1	0
	4	0.26	1.179	-17.9
	8	0.34	1.513	-51.3
	16	0.52	2.344	-134.4
	32	0.86	3.856	-285.6
	64	1.53	6.827	-582.7
	128	3.57	15.978	-1497.8
	256	9.49	42.447	-4144.7
2048-bit	2	0.98	1	0
	4	1.09	1.11	-11
	8	1.28	1.298	-29.8
	16	1.82	1.846	-84.6
	32	2.74	2.781	-178.1
	64	4.73	4.809	-380.9
	128	10	10.166	-916.6
	256	23.53	23.913	-2291.3
3072-bit	2	2.08	1	0
	4	2.34	1.123	-12.3
	8	2.73	1.309	-30.9
	16	3.88	1.858	-85.8
	32	5.85	2.803	-180.3
	64	9.98	4.785	-378.5
	128	20.19	9.679	-867.9
	256	44.85	21.498	-2049.8
4096-bit	2	3.51	1	0
	4	3.98	1.134	-13.4
	8	4.52	1.286	-28.6
	16	6.28	1.787	-78.7
	32	9.34	2.656	-165.6
	64	15.9	4.508	-350.8
	128	31.6	8.998	-799.8
	256	68.6	19.519	-1851.9

TABLE VI
RUNNING TIME RESULTS FOR THE REDUCING PRECOMPUTATION MULTIPLICATION METHOD ON ENCRYPTION

M	*m*	running time (ms)	comparison with *m*=2	saving (%)
1024-bit	2	0.05	1	0
	4	0.05	0.981	1.9
	8	0.06	1.086	-8.6
	16	0.05	0.979	2.1
	32	0.06	1.078	-7.8
	64	0.06	1.086	-8.6
	128	0.06	1.079	-7.9
	256	0.05	0.98	0.2
2048-bit	2	0.18	1	0
	4	0.18	0.976	2.4
	8	0.18	0.997	0.3
	16	0.17	0.965	0.5
	32	0.18	0.991	0.9
	64	0.18	0.993	0.7
	128	0.18	0.983	1.7
	256	0.17	0.959	4.1
3072-bit	2	0.37	1	0
	4	3.66	0.988	1.2
	8	3.69	0.995	0.5
	16	3.64	0.982	1.8
	32	3.63	0.980	2
	64	3.67	0.989	1.1
	128	3.68	0.997	0.3
	256	3.62	0.976	2.4
4096-bit	2	5.78	1	0
	4	5.76	0.997	0.3
	8	5.82	1.007	-0.7
	16	5.77	0.998	0.2
	32	5.83	1.008	-0.8
	64	5.81	1.005	-0.5
	128	5.81	1.005	-0.5
	256	5.72	0.995	0.5

= 256, there is no need for a look-up table. However, 254 precomputation multiplications are computed because of the *m-ary* method execution. To overcome this, we should consider the *reducing precomputation multiplications*. In TABLE VI, running time results of the method for different plaintext sizes are shown. By the table, small savings are obtained with the use of small amounts of memory. Obtained ratios are very close to 1. This means that the method does not provide a significant efficiency over the *repeated squaring and multiplication algorithm*. But, the *reducing precomputation multiplications* offers an alternative way for encryption.

B. Decryption

Decryption is the most time consuming part of the RSA algorithm because of the size of *d*. Hence, making improvements on this part has great importance. In TABLE VII running times of the *m-ary* method, comparision of the results with *m*=2 case and also amount of savings are listed.

By the table, for 1024-bits modulus size, choosing *m*=32 gives 28% improvement by using 3.75 KB memory. If we choose 2048-bits modulus size, *m*=64 provides 23% improvement with a cost of 15.5 KB memory. For 3072-bits modulus size, the best option is *m*=64 with 23% improvement and the required memory size is 23.25 KB. If the modulus size is chosen to be 4096-bits, *m*=128 provides 23% acceleration by using 64 KB of memory.

With the *m-ary* method, efficient cases are obtained for each modules size. But, we know that when the *modified m-ary* method is used, efficiency can be increased by using less memory. To observe running time relations between the *modified m-ary* and the *m-ary* method, these methods were compared according to the running times for different modulus sizes. Obtained running times and ratios are tabulated in the TABLE VIII. By the table, the *modified m-ary* method becomes more efficient than the *m-ary* method after *m*=16. If we look at the recommended cases for the *m-ary* method, these *m* values are larger than 16. Therefore, we should consider the *modified m-ary method* to get more efficient results.

In TABLE IX, running times of the *modified m-ary* method, comparisons with *m*=2 case and obtained savings are shown. In accordance with the information in the table, for 1024-bits modulus size choosing *m*=32 provides 28% acceleration with using 1.875 KB memory. For 2048-bits modulus size, *m*=64 provides 26% efficiency with using 7.75 KB memory. When the key size is 3072-bits, we should choose *m*=128 to get the best result. In this case, performance improvement is nearly 28% and required memory size is 23.625 KB. For 4096-bit key size, *m*=128 provides 25% acceleration with using 31.5 KB memory. The improvement does not seem to be significant compared to the *m-ary* method but the required memory size decreases to exactly the half in for each case.

IV. CONCLUSION

In this paper, we present some fast modular exponentiation methods. The aim is to find particular parameters for this oper-

TABLE VII
RUNNING TIME RESULTS FOR THE M-ARY METHOD WITH POWER OF 2

modulus size	*m*	running time (ms)	comparison with *m*=2	saving (%)
1024-bit	2	23.6	1	0
	4	20.24	0.857	14.3
	8	18.69	0.791	20.9
	16	17.6	0.745	25.5
	32	**17.09**	**0.723**	**27.7**
	64	17.39	0.736	26.4
	128	19.44	0.823	17.7
	256	25.49	1.078	-7.8
	512	51.31	2.171	-117.1
	1024	148.04	6.264	-526.4
2048-bit	2	173.77	1	0
	4	156.14	0.899	10.1
	8	145.95	0.84	16
	16	139.44	0.802	19.8
	32	**135.34**	**0.779**	**22.1**
	64	**134.5**	**0.774**	**22.6**
	128	138.91	0.799	22.1
	256	156.01	0.898	10.2
	512	177.77	1.023	-2.3
	1024	301.24	1.734	-73.4
3072-bit	2	98.01	1	0
	4	88.84	0.907	9.3
	8	82.59	0.843	15.7
	16	78.92	0.805	19.5
	32	76.64	0.782	21.8
	64	**75.6**	**0.771**	**22.9**
	128	77.51	0.791	20.9
	256	86.9	0.887	11.3
	512	123.39	1.259	-25.9
	1024	263.14	2.685	-168.5
4096-bit	2	1219.61	1	0
	4	1108.288	0.909	9.1
	8	1037.402	0.85	15
	16	990.694	0.812	18.8
	32	962.834	0.789	21.1
	64	**947.358**	**0.777**	**22.3**
	128	**944.52**	**0.774**	**22.6**
	256	977.716	0.802	19.8
	512	1043.298	0.855	14.5
	1024	1256.364	1.03	-3

TABLE VIII
RUNNING TIME RESULTS FOR COMPARISON OVER MODIFIED M-ARY METHOD & M-ARY METHOD

modulus size	*m*	*modified m-ary* (ms)	*m-ary* (ms)	ratio
1024-bit	4	4.14	3.99	1.038
	8	3.72	3.65	1.02
	16	3.44	3.43	1.005
	32	3.35	3.39	0.988
	64	3.32	3.54	0.938
	128	3.56	4.52	0.787
	256	4.48	8	0.56
2048-bit	4	29.94	29.49	1.015
	8	27.43	26.44	1.037
	16	25.28	25.13	1.005
	32	24.51	24.589	0.997
	64	23.92	24.44	0.979
	128	24.14	26.07	0.926
	256	25.9	32.53	0.796
3072-bit	4	88.27	87.9	1.004
	8	82.84	82.84	1
	16	79.24	79.24	1
	32	76.72	76.85	0.998
	64	74.96	76.09	0.985
	128	74.71	78.01	0.957
	256	77.01	87.39	0.881
4096-bit	4	193.04	192.08	1.005
	8	179.93	178.66	1.007
	16	171.32	170.72	1.004
	32	165.16	165.16	1
	64	161.17	162.4	0.992
	128	159.77	164.13	0.973
	256	162.23	176.66	0.918

ation which have better performance compared to the *repeated squaring and multiplication algorithm*. For encryption, we implemented the *m-ary* method and the *reducing precomputation multiplications*. According to the results, the *m-ary* method is not suitable for encryption. When the *reducing precomputation multiplications* is implemented, some alternatives are observed which use small memory sizes. However, we do not recommend an option that is better than the *repeated squaring and multiplication algorithm* for encryption.

For decryption, the *m-ary* and the *modified m-ary* methods were implemented. According to the results, the *m-ary* method provides a considerable improvement in running time. This improvement is greater than 23% for each modulus size. For the *modified m-ary* method, more efficient results are obtained compared to the *m-ary* method. With this method at least 25% improvement is achieved by using less memory than the *m-ary* method.

Consequently, acceleration of the RSA algorithm especially for decryption is very important to get efficient implementations. With the analyzed methods in this paper, we observed considerable improvements in the running time for some of the methods compared to the *repeated squaring and multiplication algorithm*.

V. ACKNOWLEDGEMENTS

The second author is partially supported by TÜBİTAK under Grant No. BIDEB114C052.

REFERENCES

[1] E. Akyldz, Ç. Çalık, M. Özarar, Z. Tok, O. Yayla, RSA Kriptosistemi Parametreleri için Güvenlik Testi Yazlm, ISCTURKEY 2013, Proceedings of 6th International Security & Cryptology Conference, pp.124-127
[2] D. Boneh, Twenty Years of Attacks on the RSA Cryptosystem, Notices of the AMS, pp. 203-213, February 1999
[3] J. Chung, M. A. Hasan, Asymmetric Squaring Formulae, Computer Arithmetic.ARITH'07. 18th IEEE Symposium on, pp.113-122, 2007
[4] D. M. Gordon, A survey of Fast Exponentiation Methods, Journal of Algorithms, Vol.27, Issue 1, pp.129-146, April 1998
[5] B. Kaliski, The Mathematics of the RSA Public-Key Cryptosystem, RSA Laboratories, 2006
[6] D. E. Knuth, The Art of Computer Programming, Vol.2/Seminumerical Algorithms, Second Edition, Addision-Wesley Publishing Company, 1981
[7] Ç. K. Koç, High-Speed RSA Implementation, RSA Laboratories, 1994
[8] A. Kumar, S. Jakhar, S. Makkar, Comparative Analysis Between DES and RSA Algorithm's, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 7, 2012
[9] A. J. Menezes, P. C. van Oorschot, S. A. Vanstone, Handbook of Applied Cryptography, CRC Press, 1996
[10] R. L. Rivest, A. Shamir, L. Adleman, A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, Communications of the ACM, 1978
[11] M. Welschenbach, Cryptography in C and C++, Second Edition, Apress, 2005

TABLE IX
RUNNING TIME RESULTS FOR THE MODIFIED M-ARY METHOD WITH POWER OF 2

modulus size	m	running time (ms)	comparison with m=2	saving (%)
1024-bit	2	4.7	1	0
	4	4.07	0.867	13.3
	8	3.67	0.781	21.9
	16	3.5	0.745	25.5
	32	**3.38**	**0.719**	**28.1**
	64	3.39	0.721	27.9
	128	3.59	0.766	23.4
	256	4.54	0.967	3.3
2048-bit	2	32.32	1	0
	4	29.03	0.898	10.1
	8	26.20	0.811	18.9
	16	24.94	0.772	22.8
	32	24.14	0.747	25.3
	64	**23.74**	**0.735**	**26.5**
	128	23.9	0.74	26
	256	25.58	0.791	20.9
3072-bit	2	103.96	1	0
	4	93.48	0.899	10.1
	8	87.08	0.838	16.2
	16	81.79	0.787	21.3
	32	77.33	0.744	25.6
	64	76.02	0.731	26.9
	128	**75.32**	**0.724**	**27.6**
	256	77.63	0.747	25.3
4096-bit	2	213.79	1	0
	4	194.59	0.91	9
	8	182.15	0.852	14.8
	16	173.11	0.81	19
	32	168.24	0.787	21.3
	64	163.15	0.763	23.7
	128	**161.25**	**0.754**	**24.6**
	256	164.07	0.767	23.3

A Survey of Zero Correlation Linear Cryptanalysis

Ziya AKCENGİZ, *Middle East Technical University,* Muhiddin UĞUZ, *Middle East Technical University,*
Fatih SULAK, *Atılım University,* Hacı Ali ŞAHİN, *Middle East Technical University,*

***Abstract*—There are many symmetric-key algorithms composed of permutation, substitution, xor and summation operations. Due to the increases of block cipher encryption algorithms, the different kind of cryptanalysis methods developed and improved. The differential cryptanalysis and linear cryptanalysis are two of the most popular cryptanalysis methods which are widely used to reveal the weaknesses of substitution-permutation network ciphers. Whereas the differential cryptanalysis needs high or zero probability distinguishers, the linear cryptanalysis needs high, zero or low correlation distinguishers. In this paper, we show the contribution of zero correlation key recovery attacks into different type of block ciphers and analyse these results in terms of time and memory.**

***Index Terms*—Cryptanalysis, Zero Correlation Linear Attack, Key Recovery**

I. Introduction

Linear cryptanalysis [1] and differential cryptanalysis [2] are two of the most effective methods improved against block ciphers in early 1990s. Later, many block cipher algorithms were designed to be resistant against linear cryptanalysis and differential cryptanalysis. Due to efficent development of security of block ciphers, new cryptanalysis methods were improved such as impossible differential cryptanalysis [3], truncated differential cryptanalysis [4], improbable differential cryptanalysis [5]. Zero correlation linear cryptanalysis [6], relatively new approach, is one of these cryptanalysis methods which we analyse in the article. In the literature, there are many cryptanalysis algorithms and cryptanalysis attacks. Therefore, these attack algorithms are not only used to overcome the security of system but also to measure the security level of algorithms.

Nowadays, the relationship between linear and differential cryptanalysis attracts attention of cryptanalysts. Therefore, new theorems about relation between zero correlation and impossible differential cryptanalysis started to be appeared [7]. Since zero correlation cryptanalysis is a new approach and has an important role in the relation between linear and differential attacks, it is very helpful to have a survey about zero correlation cryptanalysis. Therefore, we give a survey explaining all steps of zero correlation cryptanalysis and effects of this attack to some block ciphers such as CLEFIA [8], CAMELLIA [9] and HIGHT [10].

II. Zero-Correlation Linear Cryptanalysis

Linear cryptanalysis is based on the correlation value of some certain input bits and output bits. Constructions of distinguisher and key recovery are two steps of linear cryptanalysis.

In the construction of distinguishers, we need to determine input and output masks of each round according to the correlation values. In the zero correlation linear attacks, input and output masks which give zero correlation are used. Since the correlation values of distinguisher should be evaluated exactly, we need to have at least half of plaintext and ciphertext values for this part which is the one of the main differences of zero correlation attacks from impossible differential attacks [6].

In the key recovery part, more rounds than covered by the zero correlation distinguisher are used and key values in extra rounds are guessed. Since we know that the correlation value is zero key-independently in the linear hull, we can check correlation values for input and output values of linear hull and decide whether it is the correct guess or not.

In this section, we firstly give the definitions of scalar product, linear hull and trail, correlation and probability values of each round and block cipher. Then, we explain construction of zero correlation distinguishers. After that, we explain how to apply key recovery attack to cipher by using the distinguishers. Finally, results of zero correlation attacks are demonstrated and efficiency of the attack algorithm is discussed. In the conclusion part, we give a brief overview of the article and discuss our future works about zero correlation attack.

A. Definitions

We need to have some definitions, scalar product, probability and correlation in order to define correlations and relations for each round. If we want to mask n-bit x value with n-bit a value, we will use the following scalar product formula

$$a \diamond x = \bigoplus_{\forall 0 \le i < n} a_i \cdot x_i$$

Also, the correlation is defined as

$$C = 2 \cdot p - 1$$

where the p is the probability of function $f(x)$ with patterns α and β such that

$$p = Pr_x\{\alpha \diamond x = \beta \diamond f(x)\}$$

When α and β are zero, $\forall x$ value, $\alpha \diamond x$ and $\beta \diamond f(x)$ are equal to 0. Therefore, for all x values, $\alpha \diamond x = \beta \diamond f(x)$ and $p = 1$ and the correlation C is

$$2 \cdot p - 1 = 1$$

If only one of α and β is zero, the equality $\alpha \diamond x = \beta \diamond f(x)$ is true with probability $\frac{1}{2}$. Therefore, $\alpha \to 0$ and $0 \to \beta$ have correlation C

$$2 \cdot \frac{1}{2} - 1 = 0$$

where $\alpha, \beta \neq 0$. Then, we can say that

$$C_{0\to 0} = 1, C_{\alpha\to 0} = C_{0\to\beta} = 0$$

Since block ciphers consists of rounds and we need to have correlation value for each round, i^{th} round of block cipher is shown as f_i and correlation value is shown as $C^{f_i}_{\alpha,\beta}$.

Assume that the block design is composed of r rounds and u_i is the input mask of i^{th} round.

$$C^{f}_{\alpha,\beta} = \prod_{1\leq i\leq r} C^{f_i}_{u_{i-1},u_i}$$

where $f = f_r \circ f_r - 1 \circ ...f_1$ and $\alpha = u_1$ and $\beta = u_r$ are input and output masks, respectively.

Finally, $(u_1, u_2, ...u_r)$ is called as linear trail and $\alpha \to \beta$ is called as linear hull.

B. Construction of Zero Correlation Cryptanalysis Distinguishers

Similar to differential cryptanalysis, we need to construct a distinguisher for substitution-permutation network designs. However, the most important point in this attack is that the distinguisher $\alpha \to \beta$ has a zero correlation for all linear trails $(u_1, u_2, ...u_r)$ where $u_1 = \alpha$ and $u_r = \beta$.

Since block cipher consists of permutation, XOR operation and branching points, we will use following lemmas in order to find a zero correlation hulls.

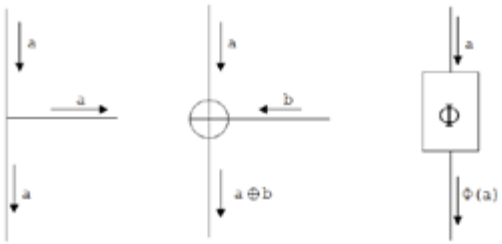


Fig. 1. Branching point, XOR Operation and Permutation

Lemma 1: Either three linear selections patterns at a XOR operation are equal or the correlation is exactly zero.

Lemma 2: Either the summation of three linear selection pattern at a branching point is zero or the correlation is zero.

Lemma 3: If both of input mask α and output mask β of permutation f are neither both zero nor nonzero, then the correlation $C^{f}_{\alpha,\beta}$ is zero.

Proof: Assume that z_1 and z_2 are input and output mask of permutation and n is bit length of input. The correlation C is equal to $2 \cdot p - 1$ where

$$p = \frac{|\{x|x \diamond z_1 = \phi(x) \diamond z_2\}|}{2^n} = \frac{|\{x|x \diamond z_1 \oplus \phi(x) \diamond z_2 = 0\}|}{2^n}$$

If z_1 is zero and z_2 is not zero, then $z_1 \diamond x$ is always zero. Therefore, the probability $p = Pr\{z_2 \diamond \phi(x) = 0\} = 1/2$. Similarly, If the z_2 is zero and z_1 is not zero, then $p = Pr\{x \diamond z_1 = 0\} = 1/2$. Therefore, if only one of input and output masks is zero, then probability p is 1/2 and correlation C is $2 \cdot p - 1 = 0$. ■

Other lemmas can be similarly proven by the definition of correlation defined on the block design components. By using three lemmas above, we can list sufficient conditions for nonzero correlations.

If we have linear trails U which justify following conditions, then the correlation C_U is always nonzero

- All masks at XOR are equal to each other,
- A summation of all masks at a branching point is zero,
- Both of input and output masks of permutation are either zero or nonzero.

C. Construction Method

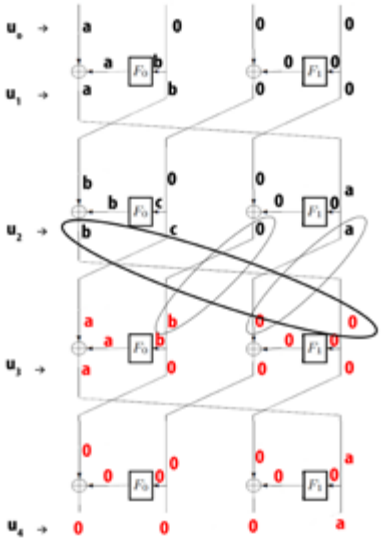
Assume that $\alpha \to \beta$ is a linear hull of block cipher f. We would like to find an α and β such that for all linear trails U, there exist the zero correlation $C^{f_i}_{u_i,u_{i+1}}$ for some $1 \leq i \leq r$.

Since there are $(2^b)^{r-1}$ possible linear trails where b is the block size and r is a number of rounds, it is not sufficient to check whether the correlation values of all linear trails are zero or not. Therefore, we start from input mask by justifying three conditions above. Also, we start from output mask similarly. Since the correlation of linear trail U is the multiplication of correlations of all rounds, it is enough that at least one round have zero correlation for each linear trail. Therefore, in the middle of design, there should be a case which does not justify three condition lemma. Since there is always inconsistency in the middle of distinguisher, the correlation is always zero for all linear trails.

We construct following example design by simplifying HIGHT algorithm[10] in order to illustrate the construction of distinguisher by using conditions above.

Example: In this example, F_0 and F_1 are permutations from

Fig. 2. A simple example of zero correlation linear cryptanalysis



$\mathbb{F}_2^8$ to $\mathbb{F}_2^8$. Also, Let

$$b = 32, r = 4,$$

$$F_0(b) = a, F_0(c) = b,$$

$$u_0 = \alpha = (a, 0, 0, 0), u_4 = \beta = (0, 0, 0, b)$$

$$u_1 = (a, b, 0, 0), u_2 = (b, c, 0, a)$$

$$u_3 = (a, 0, 0, 0)$$

where a, b, c are 8-bit non-zero mask values and $\alpha \rightarrow \beta$, $(u_0, u_1, u_2, u_3, u_4)$ are linear hull and linear trail.

We begin with input and output masks and continue until the middle of design by simply performing three conditions above. It is beneficial to justify the XOR condition firstly because we can set input-output of permutation and branching points with respect to results of XOR operations. However, input-output masks must be equal to each other by Lemma 1, we could not set these masks in terms of other masks.

In the Figure 2, we see that output of second round and the input of third round are not equal to each other. Therefore, $\alpha \rightarrow \beta$ are linear hull which give a zero correlation distinguisher. Since only one non-equivalence is enough to have zero correlation distinguisher, the number of rounds in this example can be increased and more efficient distinguishers can be created.

D. Key Recovery Part of Zero Correlation

Key recovery is the second part of zero correlation attack. After we generate a key independent distinguisher, we guess some sub-keys in block ciphers.

We perform key recovery into more rounds than the number of rounds used by linear hulls. However, all key values in extra rounds should be guessed to perform key recovery. Then, we partially encrypt plaintexts or decrypt ciphertexts in order to reduce the number of rounds into the number of rounds used in the linear hull. Later, we use input and output selection patterns determined by the distinguisher and evaluate the correlation value. Since we know that the distinguisher has zero correlation value key-independently, we check whether the correlation value is still zero or not. If it is zero, then our guess in extra key is correct.

Assume that we have r rounds in linear hulls and e extra rounds used in key recovery and each round has m-bit key. We can list the key recovery part,

- Guess $e \cdot m$ bit key values and if extra rounds are first rounds, then partially encrypt plaintexts and generate inputs for distinguishers. If extra rounds are last rounds, then partially decrypt ciphertexts and generate outputs for distinguishers.
- Evaluate the correlation value of linear hull for input and output values and check whether the correlation value is zero or not. If it is zero, then $e \cdot m$ key bits are correctly guessed. Otherwise, we repeat first step.
- Guess another $e \cdot m$ bit key values so that we can perform linear hulls in the r rounds.

III. Analysis of Zero Correlation Linear Attack to Algorithms

There are many ciphers in which zero correlation attack can be applied successfully such as ARIA, Camellia and Clefia. In this part, we give brief explanations about algorithms and analyse results of zero correlation attack on these algorithms.

In order to illustrate the method of distinguisher construction, we explain how to find distinguishers for some algorithms and give tables which explain the attack efficiencies in terms of time and memory complexity, number of rounds and amount of data. Time complexity shows how long the attack run until recovering key. Memory complexity shows how much memory space we use through the attack.

In the tables, $ZC.FFT$ is the method which reduces time complexity of zero correlation attack by using Fast Fourier Transform in the matrix vector multiplications. $Mutlidim.ZC$ is another extension of zero correlation attack which reduces the memory complexities and keeps time and data complexities unchanged.In the key recovery part of multidimensional zero correlation attack, we use non-zero distinguisher combinations and analyse candidate key distributions by computing the statistic T in $section$ 2.3 in [11]. Finally, $ZC.partial{-}Sum$ is the last method shown in tables. In the key recovery attack we need to do partial encryption or decryption by guessing some part of keys. Partial-sum method provides more efficient way to encrypt or decrypt data than direct encryption or decryption [13].

A. Camellia

Camellia is the block cipher which supports 128-bit block size and 128, 192 and 256-bit key size. The cipher mainly consists of functions namely F, P, FL and S-boxes. It is claimed that Camellia has stong security compared to MARS [22], RC6[24], Rijndael[23], Serpent[25] and Twofish[26]. In the Table I summary of cryptanalysis is given.

TABLE I
SUMMARY OF ZERO CORR. ATTACKS ON CAMELLIA [11]

Key Size	round	Attack Type	Data	Time	Memory
128	11	ZC. FFT	$2^{122.5}$	$2^{123.5}$	2^{101}
192	12	ZC. FFT	$2^{125.7}$	$2^{188.8}$	2^{112}

Since we need to know properties of all components of the algorithm, properties of FL functions should be determined in order to find suitable distinguisher. Therefore, properties given in [11] are used together with lemmas explained. One of properties is the following.

Property 1:

If the input mask of FL is $(0, 0, 0, 0, 0, 0, 0, i)$, then the output mask of FL is $(?, 0, 0, ?, ?, 0, 0, ?)$ where $?$ is an unknown value.

There are similar input and output masks for FL^{-1}.Since we know how mask values change when the FL is applied, we can find suitable input and output mask values whose correlation value is zero.In [11], 7 round distinguisher is constructed by applying same methods illustrated in $section$ 2. The distinguisher starts with an input mask $(b, 0, 0, b, 0, b, b, b)$

- $(0,0,0,0,0,0,0,0)$ and an output mask $(0,0,0,0,0,0,0,0)$
- $(h,0,0,h,0,h,h,h)$ where 0, h and b are 8-bit zero and non-zero values, respectively.

B. ARIA

Aria supports 128-bit block size and 128,192 and 256 key size [12]. It consists of basic operations such as XOR operations and S-boxes of Rijndael[23]. There is a 16×16 involutory binary matrix with branch number 8 in order to avoid differential and linear attacks of Rijndael. In the Table II summary of cryptanalysis is given.

TABLE II
SUMMARY OF ZERO CORR. ATTACKS ON ARIA [13]

Key Size	round	Attack Type	Data	Time	Memory
128	6	ZC. Partial-Sum	$2^{123.6}$	2^{121}	$2^{90.3}$
128	6	ZC. FFT	$2^{124.1}$	$2^{121.5}$	$2^{90.3}$
256	7	ZC. Partial-Sum	$2^{124.6}$	$2^{203.5}$	2^{152}
256	7	ZC. FFT	$2^{124.7}$	$2^{209.5}$	2^{152}

C. Clefia

Clefia is the block cipher with 128-block size and 128, 192 and 256-key size. Because of new evaluation techniques, it is composed of cost-efficient components such as Fiestel structure with 4 data lines, 32-bit F functions and diffusion matrices. One of the most important points in the algorithm is that diffusion matrix and two different S-box used by Clefia is constructed specifically in order to increase algebraic immunity of the cipher[8]. In the Table III summary of cryptanalysis is given.

TABLE III
SUMMARY OF ZERO CORR. ATTACKS ON CLEFIA [11]

Key Size	round	Attack Type	Data	Time	Memory
192	14	Multidim. ZC	$2^{127.5}$	$2^{180.2}$	2^{115}
256	15	Multidim. ZC	$2^{127.5}$	$2^{127.5}$	2^{115}

In order to attack 14-round Clefia, 9-round distinguisher is constructed. Since Clefia is composed of diffusion matrix, S-boxes, XOR and substitution, it is enough to apply lemmas 1,2 and 3. Connections of some successive bytes cause diffusion in the algorithm. Therefore, distinguisher with limited number of rounds can be constructed.In figure 3, construction of 9-round are displayed and there is an inconsistency at function F_0 of round 5. In the Clefia, input and output masks are set as $(a,0,0,0)$ and $(0,0,0,a)$, respectively where a is non-zero value.

D. TEA and XTEA

TEA is the short program and consists of large number of iterations. Also, the set-up time of the algorithm is shorter than other algorithms. Subtraction and summation operations are used for encryption and decryption instead of XOR operation in this algorithm. In the algorithm, diffusion is provided at the sixth iteration. Therefore, TEA and XTEA should consist of at least 6 rounds [14].

XTEA is constructed as an extension algorithm of TEA in order to eliminate some weakness of TEA[15]. In the TableIV summary of cryptanalysis is given.

Fig. 3. Zero Correlation Distinguisher for Clefia

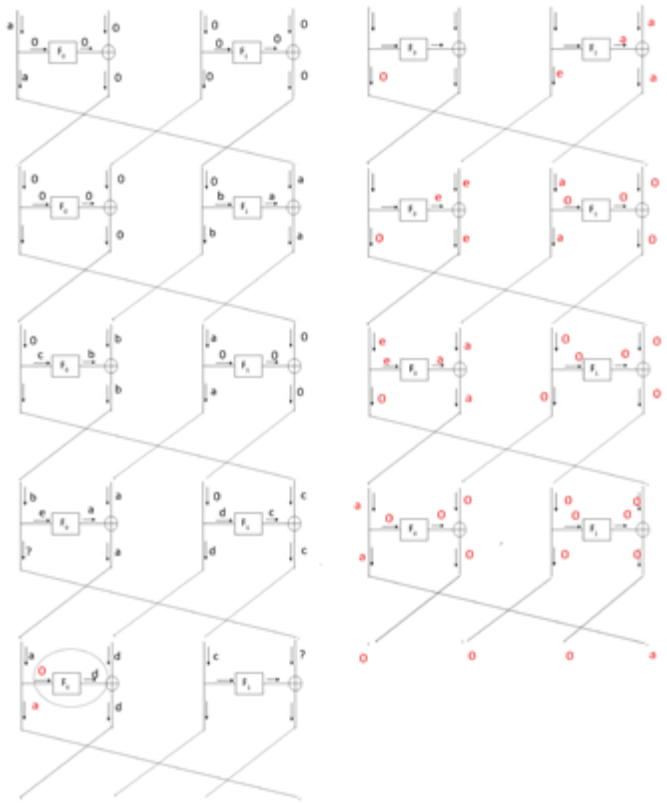


TABLE IV
SUMMARY OF ZERO CORR. ATTACKS ON TEA AND XTEA [16]

Key Size	round	Attack Type	Data	Time	Memory
126	21	ZC	$2^{62.62}$	$2^{121.52}$	negligible
126	23	ZC	2^{64}	$2^{119.54}$	negligible
126	25	ZC	$2^{62.62}$	$2^{124.53}$	2^{30}
126	27	ZC	2^{64}	$2^{120.71}$	negligible

E. HIGHT

HIGHT is a 32-round algorithm with 64-bit block size and 128-bit key size. There are two transformation functions before and after all rounds. In each round, there are two different Feistel functions, XOR and summation operations [10]. In the Table V summary of cryptanalysis is given.

TABLE V
SUMMARY OF ZERO CORR. ATTACKS ON HIGHT [17]

Key Size	round	Attack Type	Data	Time	Memory
128	26	ZC	$2^{62.79}$	$2^{119.1}$	2^{43}
128	27	ZC	$2^{62.79}$	$2^{120.78}$	2^{43}

Since block size of Hight algorithm is 64-bit, there are 8 bytes which can be masked in Hight. In the table 3 [17], there is a distinguisher applyed into 26 rounds. Only fourth byte is nonzero in the input mask.Since there is a connection between some of successive bytes, masked value of fourth byte influences the fifth byte. Also, unconnected bytes influence each other due to rotations applied at the end of each rounds.

Similar to input mask, all output mask values are zero except for third byte. Because of the connection between third and second bytes, mask value in second byte of 16^{th} round becomes non-zero. After continuing from input mask and output mask,there is an inconsistency in the third byte of ninth round. Therefore, the correlation value of the distinguisher is never non-zero.

F. CAST256

It is the algorithm constructed by framework used in DES algorithm. Some components of CAST are designed in terms of strict avalanche criterion. Since components in each round are constructed more securely than DES. CAST256 with fewer rounds than DES algorithm has same security level as DES algorithm. Each round mainly consists of Feistel functions and XOR operations[18].

Cast256 is the 256 block size version of CAST algorithm. In the Table VI, summary of cryptanalysis is given.

TABLE VI
SUMMARY OF ZERO CORR. ATTACKS ON CAST256[19]

Key Size	round	Attack Type	Data	Time	Memory
256	28	Multidim. ZC	$2^{98.8}$	$2^{246.9}$	2^{68}

In the CAST256, 24-round distinguishers can be constructed in order to apply 28-round Key recovery zero correlation attack. Construction of the distinguisher is so similar to illustration given in the $section\ 2$, $part\ C$ because CAST256 is composed of only XOR, Fiestel functions and branching poinrs. The distinguisher has input mask $(0, 0, 0, a)$ and output mask $(0, 0, 0, b)$ where a and b are 8-bit non-zero mask value in the figure 5 [19].

G. LBLOCK

LBLOCK is the lightweight block cipher with 64-bit block size and 80-bit key size. Each round consists of Fiestel structure which is composition of diffusion and confusion functions. In the table 5 [20], efficiency of hardware implementation is compared to other algorithms such as XTEA, HIGHT and DES. In the Table VII, summary of cryptanalysis is given.

TABLE VII
SUMMARY OF ZERO CORR. ATTACKS ON LBLOCK[21]

Key Size	round	Attack Type	Data	Time	Memory
80	22	ZC	2^{64}	$2^{70.54}$	2^{64}
80	22	ZC	2^{62}	$2^{71.27}$	2^{64}
80	22	ZC	2^{60}	2^{79}	2^{64}

IV. CONCLUSION

In this paper, we firstly explain the zero correlation linear distinguishers. Since it is necessary to have a distinguisher with zero correlation, relations between input and output mask values and correlations are stated and partially proved in lemmas. Also, the method of distinguisher construction is explained and illustrated. Then, usage of distingiusher and key recovery parts are itemized and the explanation of the attack is completed. After that, seven different block cipher designs are selected and brief informations about ciphers are given. In each cipher, we construct a table which demonstrate Key size,round and data-time complexity. In conclusion, we stated steps of zero correlation attack, give illustration of construction of the distinguisher and displayed tables for all selected block ciphers. We aim to increase the number of block ciphers effected by zero correlation attack and study on the relation of zero correlation attack with other effective attacks.

REFERENCES

[1] M. Matsui, Linear Cryptanalysis Method for DES Cipher, Advances in Cryptology-EUROCRYPT '93 (Lecture Notes in Computer Science no. 765), Springer-Verlag, pp. 386-397, 1994.

[2] E. Biham and A. Shamir, Differential Cryptanalysis of DES-like Cryptosystems, Journal of Cryptology, vol. 4, no. 1, pp. 3-72, 1991.

[3] Eli Biham, Alex Biryukov, Adi Shamir, Miss in the Middle Attacks on IDEA and Khufu, proceedings of Fast Software Encryption 6, Lecture Notes in Computer Science 1636, pp. 124138, Springer-Verlag, 1999.

[4] L. R. Knudsen. Truncated and higher order differential. In B. Preneel, editor, Fast Software Encryption-Second International Workshop, volume 1008 of Lecture Notes in Computer Science, pages 196211. Springer-Verlag, 1995.

[5] Tezcan, C.: The improbable differential attack: cryptanalysis of reduced round CLEFIA. In: Gong, G., Gupta, C.K. (eds.) INDOCRYPT 2010. LNCS, vol. 6498, pp. 197209. Springer, Heidelberg (2010)

[6] Andrey Bogdanov, Vincent Rijmen. Zero-Correlation Linear Cryptanalysis of Block Ciphers. In *FSE'2012*, volume ??, pages 29-48. Springer, 2012

[7] Blondeau, C., Nyberg, K.: New Links between Differential and Linear Cryptanalysis. In: Johansson, T., Nguyen, P.Q. (eds.) EUROCRYPT 2013. LNCS, vol. 7881, pp. 388404. Springer, Heidelberg (2013)

[8] Shirai, Taizo, et al. "The 128-bit blockcipher CLEFIA." Fast software encryption. Springer Berlin Heidelberg, 2007.

[9] Aoki, K., Ichikawa, T., Kanda, M., Matsui, M., Moriai, S., Nakajima, J., Tokita, T.: Camellia: A 128-bit block cipher suitable for multiple platforms. In: Stinson, D.R., Tavares, S. (eds.) SAC 2000. LNCS, vol. 2012, pp. 4154. Springer, Heidelberg (2001)

[10] Deukjo Hong, Jaechul Sung, Seokhie Hong, Jongin Lim, Sangjin Lee, Bonseok Koo,Changhoon Lee, Donghoon Chang, Jaesang Lee, Kitae Jeong, Hyun Kim,Jongsung Kim, and Seongtaek Chee. HIGHT: A New Block Cipher Suitable for Low-Resource Device. In Louis Goubin and Mitsuru Matsui (editors), CHES 2006, volume 4249 of Lecture Notes in Computer Science, pages 46-59. Springer, 2006.

[11] Bogdanov, Andrey, Geng Huizheng, Wang Meiqin, Wen Long, Collard Baudoin. "Zero-correlation linear cryptanalysis with FFT and improved attacks on ISO standards Camellia and CLEFIA." Selected Areas in Cryptography–SAC 2013. Springer Berlin Heidelberg, 2014. 306-323.

[12] Daesung Kwon, Jaesung Kim, Sangwoo Park, Soo Hak Sung, Yaekwon Sohn, Jung Hwan Song, Yongjin Yeom, E-Joong Yoon, Sangjin Lee, Jaewon Lee, Seongtaek Chee, Daewan Han, Jin HongShow less, "New Block Cipher: ARIA," ICISC, LNCS, vol. 2971, Springer, 2004, pp.432-445

[13] Yi, Wen-Tan, Shao-Zhen Chen, and Kuan-Yang Wei. "Zero-Correlation Linear Cryptanalysis of Reduced Round ARIA with Partial-sum and FFT." KSII Transactions on Internet and Information Systems (TIIS) 9.1 (2015): 280-295.

[14] Wheeler, D.J., Needham, R.M.: TEA, a Tiny Encryption Algorithm. Technical report, Computer Laboratory, University of Cambridge (1995)

[15] Needham, R. M., Wheeler, D. J. (1997). Tea extensions. Report, Cambridge University, Cambridge, UK (October 1997).

[16] Bogdanov, Andrey, and Meiqin Wang. "Zero correlation linear cryptanalysis with reduced data complexity." Fast Software Encryption. Springer Berlin Heidelberg, 2012.

[17] Long Wen, Meiqin Wang, Andrey Bogdanov, Huaifeng Chen. "Multidimensional zero-correlation attacks on lightweight block cipher HIGHT: Improved cryptanalysis of an ISO standard." Information Processing Letters 114.6 (2014): 322-330.

[18] Adams, Carlisle M. "Constructing symmetric ciphers using the CAST design procedure." Selected Areas in Cryptography. Springer US, 1997.

[19] Andrey Bogdanov, Gregor Leander, Kaisa Nyberg, Meiqin Wang. ”Integral and multidimensional linear distinguishers with correlation zero.” Advances in CryptologyASIACRYPT 2012. Springer Berlin Heidelberg, 2012. 244-261.
[20] Wu, Wenling, and Lei Zhang. ”LBlock: a lightweight block cipher.” Applied Cryptography and Network Security. Springer Berlin Heidelberg, 2011.
[21] Soleimany, Hadi, and Kaisa Nyberg. ”Zero-correlation linear cryptanalysis of reduced-round LBlock.” Designs, Codes and Cryptography 73.2 (2014): 683-698.
[22] Carolynn Burwick and Don Coppersmith and Edward D’Avignon and Rosario Gennaro and Shai Halevi and Charanjit Jutla and Stephen M. Matyas and Luke O’Connor and Mohammad Peyravian and David Safford and Nevenko Zunic, The MARS Encryption Algorithm, 1999
[23] Joan Daemen and Vincent Rijmen, AES Proposal: Rijndael, 1998
[24] Scott Contini and Ronald L. Rivest and M. J. B. Robshaw and Yiqun Lisa Yin, The Security of the RC6 TM Block Cipher, 1998
[25] Ross Anderson and Eli Biham and Lars Knudsen, Serpent: A Proposal for the Advanced Encryption Standard
[26] Bruce Schneier and John Kelsey and Doug Whiting and David Wagner and Chris Hall and Niels Ferguson, Twofish: A 128-Bit Block Cipher, In First Advanced Encryption Standard (AES) Conference, 1998

ISBN: 978-605-86904-3-1

BİLGİ GÜVENLİĞİ DERNEĞİ

Maltepe Mahallesi Tuncer Sokak No: 2/8
06570 Çankaya-ANKARA
0 (312) 231 18 10
bilgi@bilgiguvenligi.org.tr