



SPCSJ

**SCIENTIFIC AND PRACTICAL
CYBER SECURITY JOURNAL**

**VOL3 No2
June 2019**

ISSN 2587-4667

Securing User's Attributes on Transit to the Cloud using AES-128 bits Cryptography and DCTM3 Steganography Techniques

Maria M. Abur, Sahalu B. Junaidu, Saleh E. Abdullahi and Afolayan A. Obinyi
Department of Computer Science, Ahmadu Bello University, Zaria.

ABSTRACT. Cloud adoption is increasing day by day as such, more and more trades and enterprises are moving their vital IT structure and data to the cloud. This move is driven by the remarkable potential of cloud platforms that promise exceptional functioning, efficacy, productivity, agility, elasticity and cost-effectiveness. Although every technology has its strengths and weaknesses, the nature of the cloud makes it vulnerable to the following issues: Performance, Security and Cloud Interoperability with the main problem being security and to be even more specific are the privacy concern which cloud users really fear. The lack of privacy is the inability to protect user's attributes (or Personal Identifiable Information (PII)) as a result of data leakage, breaches and loss of data. This had made users' sceptical about sending their sensitive data to the cloud. Although there are other solutions to protect user's data during transit such as securing user's attribute with the Rivest–Shamir–Adleman (RSA) cryptography. However, RSA have been practically broken and user's sensitive information compromised. Also data leakages still hamper the security of user's data during transmission on the network to the Identity provider (IdP) on the Cloud. This paper presents an Enhanced PII Privacy Protection solution using Advanced Encryption Standard AES-128 and Discrete Cosine Transform Modulus Three (DCT-M3) Steganography techniques in order to protect user's attributes from being leaked when it is being transmitted and stored on the IdP in the cloud. The supremacy of the proposed model over the existing model was also measured based on the encryption techniques used, undetectability and robustness of the Stego image.

KEYWORDS: Cloud, Personal Identifiable Information, Security, Identity provider, Network, Steganography, Cryptography techniques and Transit

I INTRODUCTION

Identity Management refers to set of principles, policies, procedures, and technologies that offer automated identifications to persons and preserve confidential facts about the owners of those identifications and help in finding out and allowing access to resources. Identity Management System defines a system that contains information and group of technologies that can be used for innovativeness or inter-network identity management. Examples of Identity Management Systems (IMS) are Shibboleth, OpenID, OpenAm, CardSpace, Liberty Alliance and OAuth (Chadwick (n.d.); Suriadi *et al.*, (2007) and Teena *et al.*, (2017)). The features of IMS includes: Undetectability, Unlinkability and Confidentiality. These features are related with each other due to the fact that they deal with the descriptive actions involving parties in access to a range of private and sensitive data Teena *et al.*, (2017). The undetectability feature shields communications done by the user and prevents the exposure of the user actions in a given system. While the unlinkability feature hides the communication between user identities and history of transactions (e.g., subjects, messages, events, actions) and then the confidentiality feature enables users exercise control over the dissemination of their attributes.

Shibboleth as an Identity Management System (SIMS) is liable for forming the identity of a user, i.e. creating, maintaining and managing identity information for the user, and also managing access to services by the user. It involves the: Identity Provider and Service Provider, (Hogan (n. d.); Chadwick (n. d.); Suriadi *et al.*, (2007); Weingartner & Westptall (2014) and Abur *et al.*, (2018)). Important concepts on SIMS based on this research are: User's attributes, Service Providers (SPs) and the Identity Provider (IdP). The user's attributes which refers to any information usually used to uniquely identify a person whether alone or by combination with other public data that could be connected to a particular person. It is also called Personal Identifiable Information (PII). In this paper, PII is interchangeably used with the user's attributes. Examples of user's attributes are: first name, dates of birth, addresses, student number and the likes. The service providers denotes organizations that provide services or resources desired by a user, by requesting for the submission of valid credentials such as attributes or pseudonyms from the user's IdP. Finally the Identity Provider is "the entity that creates, maintains and manages identity information for the users and provides users' authentication to other service providers within IMS.

Cloud adoption is increasing day by day as such, more and more trades and enterprises are moving their vital IT structure and data to the cloud. This move is driven by the remarkable potential of cloud platforms that promise exceptional functioning, efficacy, productivity, agility, elasticity and cost-effectiveness. Although every technology has its advantages and weaknesses, the nature of the cloud makes it vulnerable to the following issues: Performance, Security and Cloud Interoperability with the main problem being security and to be even more specific are the privacy concern (i.e. data leakage, breaches and loss of data) has made Cloud users' sceptical about sending their sensitive data to the Cloud. Although there are other solutions to protect user's data during transit to IdP such as securing user's attribute with the Rivest–Shamir–Adleman (RSA) cryptography. However, RSA have been practically broken and user's sensitive information compromised. Also data leakages still hamper the security of user's data during transmission on the network to the IdP in the cloud.

Securing of user's attributes in the cloud environment must comprise of definite features that consider the intricacy of the environment. If the security of user's attributes is not guaranteed in the cloud, users will remain sceptical about transferring data to the cloud. This paper presents an Enhanced PII Privacy Protection solution using Advanced Encryption Standard AES-128 and Discrete Cosine Transform Modulus Three (DCT-M3) Steganography techniques in order to protect user's attributes from being leaked when it is being transmitted and stored on the IdP in cloud. The supremacy of the proposed system over the existing system was also measured based on the encryption techniques used, undetectability and robustness of the Stego image. This paper proposes to ensure that user's attributes is secured during transit and when stored in the IdP on Cloud.

The subsequent sections are organized as follows: Section II describes related work; Section III illustrates the PII privacy protection model; Section IV discusses models for formalizing attributes protection on IdP; Section V presents a comparative analysis of the prototype model versus existing model; Finally, section VI concludes the paper.

II RELATED WORK

Shibboleth is a joint project of Internet2 and IBM. It is open-source based system that supports inter-institutional resource sharing with access. Shibboleth enables the secure exchange of interoperability of services. It employs the idea of federated identity and Single Sign-On (SSO) authentication where there is interaction between the IdP, SP and User. However, there are still limited features and functionality that threatens user's privacy and identity if they store and process personal information with inadequate protective measures Aldeen *et al.*, (2010). Usually user's attributes are entered as plaintext when they sign up into the system and when these attributes passes through the network they become exposed to data leakage and network issues such as sniffing, spoofing, eavesdropping and malicious insider attack, Asha *et al.*, (2016) & CSA, (2016). This is not healthy for the privacy of the user when their data are transmitted to the cloud.

Switch *et al.*, (2010) added uApprove plugin – a user consent module for shibboleth identity providers to address problem of the existing system. The uApprove plugin displays to the user, the Personal Identifiable Information (PII) that the IdP shall release to the requesting SP on behalf of the user. It also offers awareness of data release when accessing some services. However, users cannot make a choice of data that should be divulged to the Service Provider. Similarly, the client has no option assenting/dissenting with His/her PII disclosure. During the dissemination of user's attributes to SPs on the network, data leakage is envisaged as user's attributes are usually exposed to security threat such as spoofing, sniffing, eavesdropping, malicious attack, Asha *et al.*, (2016) & CSA, (2016) and the SPs having received these attributes use them maliciously either directly or indirectly against the users without their consent and then leading to collusion. This way, the users' privacy is being threatened. Also during signup by users into their IdP, through the network, user's attributes are transmitted as plaintext which are prone to data leakage and network issues.

Orawinwatakul *et al.*, (2010) added “uApprove.jp, a user consent acquisition system (UCAS) with an attribute-filter mechanism for a Shibboleth based SSO system”. uApprove.jp request the “user's consent and enable the user control the release of his/her original attributes” values or PII values whether mandatory/optional from the IdP to the SP and then allows the user to determine which of his/her original attributes values are meant to be sent to the SP in order to access services provided by Service Providers. The uApprove.jp is an extension of uApprove Switch *et al.*, (2010). However, the user's control of his/her PII is ineffective making them vulnerable as there is still data leakage. Also original users' attributes values are released to the SP by the IdP are sometimes maliciously shared among other SPs or even used without the user's consent to either harm the user directly or indirectly hence leading to violation of the user's privacy and causing collusion problem. Similarly, during signup by users into their IdP, through the network, user's attributes are transmitted as plaintext which are prone to data leakage and network issues, Asha *et al.*, (2016) & CSA, (2016).

Weingartner & Westptall (2014) improved on the research of Orawiwattanakul *et al.*, (2010) by adding two objects namely: Template Data Dissemination and Cryptography Encryption Key Technique on Shibboleth/uApprove.jp framework. The former object helps users during the course of dissemination of their attributes from the IdP to the

SPs; while the later enable users enter their attributes as plaintext and then get them encrypted with Key I before sending them to IdPs. During a transaction when some PII data is needed by SPs, users would be entreated to open that encrypted data with key II in order to disseminate them to the requesting SP for the release of resources back to user. The Cryptography Encryption Key Technique used by Weingartner & Westptall (2014) is the Rivest–Shamir–Adleman (RSA). However, three basic weaknesses were identified on this system. The first is centred on the encryption techniques used i.e. the RSA which have been be practically broken and user’s sensitive information compromised Kumar *et al.*, (2011); Tripathi & Agrawal, (2014) and Hackers News, (2017). The second weakness identified is the exposure of the transmitted attributes on the communication medium, which attracts potential hackers due to data leakage, Asha *et al.*, (2016) & CSA, (2016) when transmitting to the IdP. Thirdly, SPs having received these attributes may keep them and use them without the users consent or even maliciously use them against users in the future. This leads to collusion.

Leandro *et al.*, (2014) demonstrated the flow of operations of the Shibboleth architecture and discussed its main components in details. It was observed that the flow of information from user to SP, then to IdP and then back to the user is lengthy and thereby exposing users to security threat such as data leakage on the network. Secondly, there is delay in WAYF populating the IdP for User’s selection. Thirdly, during dissemination of user’s attributes from the IdP to the SP, the original user’s attributes value are sent to the SP. This attributes may be kept or further reused again by the SPs without the consent of the user for malicious purposes and thus, violating the user’s privacy. Furthermore, during signup by users into their IdP, through the network, user’s attributes are transmitted as plaintext which are prone to data leakage and network issues.

In comparison with existing system of Weingartner & Westptall (2014), based on the features in the proposed system it is expected that this will outperform the existing system in the area of securing users attributes on the IdP. This is as a result of the fact that the AES-128 is more secured than RSA encryption techniques and further hiding of the encrypted attributes in the Stego image provides additional security. Hence, joining both of them will provide a stronger security solution. This paper aims at providing a better solution of securing users attributes on transit to the IdP on cloud and ensuring that the privacy of the user is preserved. Figure 1(a) depicts the existing PII privacy model used for securing user’s attributes and Figure 1(b) illustrates insecure communication between user & IdP and then IdP & SP on the Cloud in the existing system. This paper is focused on the insecure communication between user & IdP.

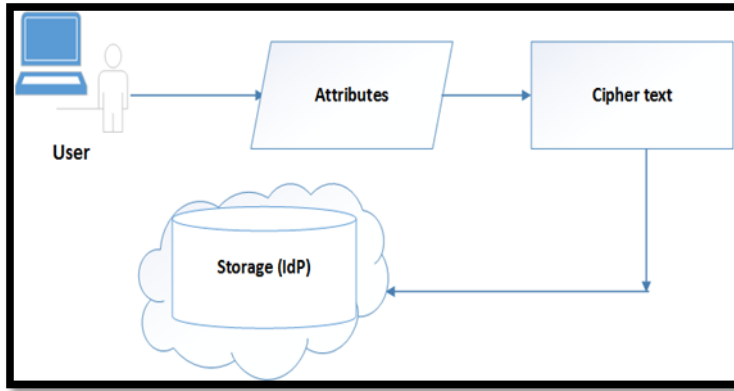


Figure 1(a): Existing PII privacy model for securing user's attributes. Weingartner & Westptall (2014)

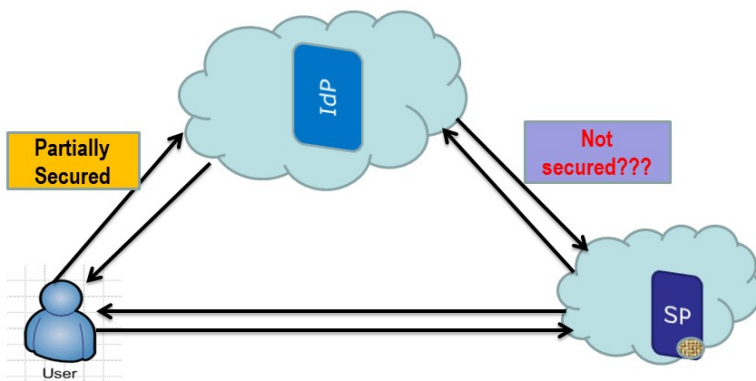


Figure 1(b): Insecure communication between user, IdP and SP on the Cloud on the existing system.

III PROPOSED PII PRIVACY PROTECTION MODEL

This is an improvement on the existing PII privacy protection model of Weingartner & Westptall (2014). Users would enter their attributes as plaintext into identity provider (IdP) in the same way like on the existing model. The proposed system builds a more secured system which will make up for the weaknesses of the existing model discussed in section III. Two basic weaknesses were identified in the existing system. The first is centred on the encryption techniques used i.e. the Rivest–Shamir–Adleman (RSA) which have been practically broken and user's sensitive information compromised. The second weakness identified is exposure of the transmitted attributes on the communication medium, which attracts potential hackers. The proposed model will however address these two basic weaknesses identified with the existing system.

Firstly to take care of the encryption techniques the proposed model uses the AES-128 encryption technique. The use of this encryption technique is hinged on the fact that attributes encrypted with this technique have not been practically broken, though theoretically broken (Kaminsky *et al.*, (2010); Song *et al.*, (2014); Sachdev *et al.*, (2013); Lokhande *et al.*, 2014 and Aleisa (2015)). Although Literature had revealed that AES-128 is yet to be 100% completed and that at year 2020 AES-128 will be completely broken, (Kaminsky *et al.*, (2010); Song *et al.*, (2014)). An attempt to wait for the practical breaking of the AES-128 before enhancing the security of user's attributes will create a vacuum in the security of user's attributes on the cloud. It is a wise thing to do by augmenting the weaknesses

of AES-128 on the theoretical aspect and of course to forestall the effect of the practical breaking of AES-128 in the nearest future.

Secondly, to circumvent the exposure to potential hackers on the existing model and to forestall the effect of the practical breaking of AES-128 in the nearest future, the proposed model introduces DCT-M3 steganography technique (Subhedar & Mankar, 2014 and Attaby *et al.*, 2017). This is to ensure that the already encrypted attributes are hidden in a cover image given rise to a Stego image before onward transmission to the IdP. This process has potentials to make the encrypted attributes undetectable while being transmitted to the IdP.

It is expected that this will outperform the existing model in the area of securing users attributes on the IdP. This is as a result of the fact that the AES-128 is more secured than RSA encryption techniques and further hiding of the encrypted attributes in the Stego image provides additional security. Hence, joining both of them will provide a stronger security solution. The supremacy of the proposed model over the existing model shall be measured based on the encryption techniques, undetectability and robustness of the Stego image. MatLab shall be used to test the undetectability and robustness of the proposed model. This shall be presented in section IV B. Figure 2 shows the proposed PII privacy protection model. The procedure for hiding the User’s attribute using AES-128 + DCT-M3 techniques and the respective extraction stage as illustrated on Figure 2 are described in a) and b) respectively.

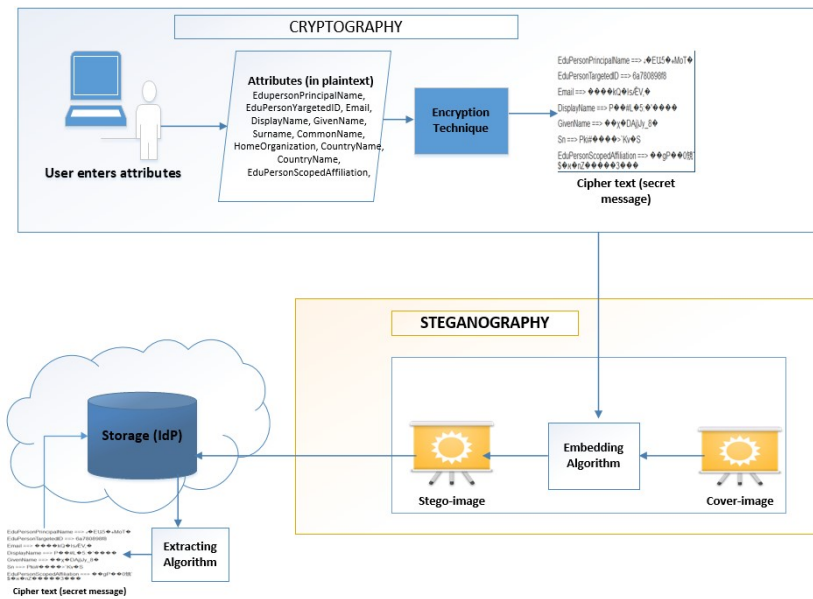


Figure 2: Proposed PII privacy protection model

a) The user’s attribute hiding stage can be summarized as:

- i. The user’s attributes (i.e. secret message) are first received as plaintext and converted to ciphertext by encrypting with the AES-128 encryption algorithm.
- ii. Obtain the binary representation of the ciphertext.

- iii. Cover image is selected and then “switch the RGB color layers of the cover image into three different components (Y, Cb and Cr)”.
- iv. After that, translate “the image into transform domain by transforming the pixel data into 8 * 8 block DCT coefficients” using the equation:
- v.
$$F(u, v) = \frac{1}{4}C(u)C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) * \cos\left[\frac{(2x+1)u\pi}{16}\right] \times \cos\left[\frac{(2y+1)v\pi}{16}\right]$$
- vi. Produce a “randomized” order with secret key, “K using pseudo random method”.
- vii. Select a static “place of two DCT coefficients which” shall be changed to implant cipher (thereby avoiding “the DC component of each DCT coefficients block”).
- viii. Inside every block of 64 coefficients implant “only two bits (Pair) as follows:
 - i.) Compute the difference between non-overlapping pair of AC coefficients which are selected.
 - ii.) Change DCT coefficients values based on the original values and the message bits accordingly”.
- ix. Conclude the implanting till the message bit stream is ended.
- x. Reinstate original order “of the DCT blocks using the key, K”.
- xi. “Quantize the image using a quantization table.
- xii. Re-order the values using Zig-Zag ordering”.
- xiii. “Use Huffman lossless compression coding to compress the image.”

b) The decoding phase can be condensed as:

- i. Alter the stego-image into transform domain by changing the pixel data into 8 * 8 block DCT coefficients.
- ii. Produce randomized order “with key, K using pseudo random method”.
- iii. Inside every block of 64 coefficients decode two bits (Pair).
- iv. Join the decoded sub message pairs to get a stream of bits.
- v. “Uncompressing the stream of bits to get the original message (ciphertext).”

IV MODELS FOR FORMALIZING ATTRIBUTES PROTECTION ON IDP

A. Formalising Shibboleth User’s attributes Protection:

In the rest of the paper, the symbol “ \rightarrow ” shall be used to denote mathematical function, the symbol “ \Rightarrow ” shall be used to denote transmission between two entities, the symbol “ \Leftrightarrow ” will be used to denote two-way transmission between entities and the symbol “ \downarrow ” represents constraint on the transmission between two points.

A. Formalising the Existing PII privacy model:

Let α_f be a function that is used for encrypting user’s attributes with public key and passphrase of the existing model. Let set $A = \{a_1, a_2, a_3 \dots a_n\}$ represent user’s attributes needed to be secured; where n is a positive integer ≥ 1 . Let $B = \{b_1, b_2, \dots, b_n\}$, be a set of encrypted values where $b_i = \alpha_f(a_i)$, $a_i \in A$. Now the encryption can be defined in (1) as:

$$\alpha_f: A \rightarrow B \tag{1}$$

Let F be a function representing the communication channel that transmits B to the storage C on the IdP in the Cloud. Let c_i be the elements of B that are stored in C , as shown on (2):

$$F: B \Rightarrow C \tag{2}$$

Since B must pass through F in order to get to C , data leakage through the communication medium, F has been identified by Asha *et al.*, (2016) and CSA, (2016) as a challenge that needs to be addressed when B travels to C . Let W_q be the data leakage problem that may occur between B and C , which can be modeled as shown in (3):

$$F: B \xRightarrow{\downarrow W_q} C \tag{3}$$

The (1) to (3) which were conceptualized by this research have been able to reveal each of the critical stages required for transmission of attributes (A) to the storage C on the IdP in the cloud. The challenge of data leakage and several network attacks are things that should not be ignored. Hence, there is the need to increase the security of the system in order to be able to minimize the influence of network attacks on the transmitted data. To increase the security “of the existing system” in order to prevent W_q , (4) is then introduced.

B. Mathematical model for the proposed PII privacy protection

Let set A represents user’s attributes to be secured, the proposed model employs encrypting function which is based on the symmetric encryption technique. Let α_z be a function denoting the AES-128 cryptography technique of the proposed system that is applied to the elements of A to produce elements of Y , where Y is the set of ciphertext in the proposed solution. Literature have shown that the AES-128 techniques has been theoretically broken, and has not been practically broken (Kaminsky *et al.*, (2010), Song *et al.*, (2014), Sachdev *et al.*, (2013); Lokhande *et al.*, (2014) and Aleisa, 2015) hence, the preference for this encryption technique in the proposed system. This is illustrated in (4):

$$\alpha_z: A \rightarrow Y \tag{4}$$

However, it has been mention in Kaminsky *et al.*, (2010) that at year 2020, AES-128 will be completely broken. The introduction of Steganography is to shield the encrypted information and increase the rigor of breaking the AES-128 cryptography technique. The next step is to introduce a function β representing the DCT-M3 Steganography Technique which will embed the ciphertext in Y into an image. Let v be the image that has the capacity to accommodate all the encrypted attributes which is to be stored in K , where K is a set of the Stego-images. This step is as described by equation (5):

$$\beta: Y \rightarrow K \quad (5)$$

In summary, Let U be the composition of α_z and β which will take the user attributes from set A to the set of stego images K . This is shown in (6):

$$U: A \rightarrow K \quad (6)$$

where $U = \beta \circ \alpha_z$

The next step is to transmit the stego image from K to through F to the storage C on the IdP in the cloud. Let F be a function representing the communication channel that transmits K to C . This is demonstrated on (7) as:

$$F: K \Rightarrow C \quad (7)$$

So that for the stego image k_i in the set K there is an element c_i in C such that $c_i = F(k_i)$ to be stored in C on the IdP.

V COMPARATIVE ANALYSIS OF THE PROPOSED MODEL VERSUS THE EXISTING MODEL

The comparison analysis of the proposed model over the existing model shall be measured based on the encryption techniques used and then the undetectability and robustness of the Stego image of the proposed model.

A. *Comparative analysis based on the complexity of the existing and proposed model with respect to the Wq attack on the Cloud C.*

Comparison of the proposed model with the existing model is based on the security of the transmitted data through F (communication channel which is the network) to the storage C in the cloud. This comparison is based on the encryption technique used i.e. the RSA versus AES-128 which were employed in the existing technique and the proposed technique respectively. The RSA is an asymmetric technique which is a two key method which uses one of the keys for encryption and the other for decryption. This method is susceptible and has been known to be broken (Kumar *et. al.*, 2011; Tripathi & Agrawal, 2014 and Hackers News, 2017). The method is slow in computation and inefficient in terms of implementation (Kumar *et. al.*, 2011; Tripathi & Agrawal, 2014 and Hackers News, 2017). Also, longer key generation time and high computational overload characterize the method. On the other hand, the AES-128 technique uses a single key for encryption as well as decryption. It is characterized by the following advantages: has not “been practically broken in reality”; very easy to design; stronger and faster in nature; efficient in Speed and code compactness on a wide range of platforms, has low computational overload, effective generation time and has short key size (Smith,2003; Sachdev *et al.*, 2013; Lokhande *et al.*, 2014 and Aleisa, 2015).

The introduction of DCTM3 Steganography technique is an additional security measure taken to further enhance the security of the encrypted attributes. This hides the ciphertexts in Stego-image so that in case of data leakage and network attacks (sniffing, spoofing and malicious attack); the encrypted attributes are hidden from the prying eyes of the attacker.

Let $W_q(b)$ represent W_q attack as shown on (3). Since $b = \alpha_f(a)$ then the attack launched on b shall be $W_q(\alpha_f(a))$

On the other hand, from (7) representing the proposed system, the attack is carried out on the Stego-image k .

To successfully attack the transmitted data, the Stego-image has to be broken first so that $W_q(k) = W_q(\beta(y))$, which needs to be broken further to access the ciphertext becomes $W_q(\beta(\alpha_z(a)))$.

In summary, the decrypted attributes in the existing model is given as:

$$a = \alpha_f^{-1}b \tag{8}$$

This shows that W_q needs only to employ α_f^{-1} to successfully attack the transmitted attributes. In the proposed model however, the W_q attack is carried out on the Stego-image k using the key β^{-1} to set the ciphertext y then use the key α_z^{-1} on y to successfully decrypt the data. Hence W_q needs to combine α_z^{-1} and β^{-1} keys to successfully attack the data. So that,

$$a = \alpha_z^{-1}(\beta^{-1}(k)) \tag{9}$$

Comparing the W_q attack as presented in (8) for the existing model and (9) for the proposed model, one could see that recovering the original attributes is easily achievable in the existing model since only one key α_f^{-1} is required to successfully attack and obtain the original attributes.

However, the task of W_q attack on the proposed model requires a composition of two keys; $(\alpha_z^{-1} \circ \beta^{-1})$ which is more difficult than the existing model. This is so due to the fact that α_z^{-1} has not been practically obtained. Hence, the security of data transmitted over the proposed model is more guaranteed in comparison to that of data transmitted over the existing system.

B.

Performance evaluation of the security of the proposed model

This section presents further evaluation of the security of the prototype model. The objective of this session is to analyse the cover image and stego-image and then measure the security of the proposed model based on the following metric: undetectability and robustness of the Stego image.

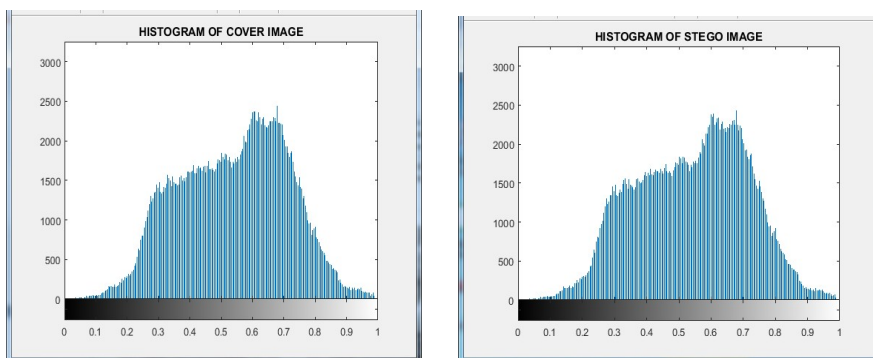
1.) *Analysis of the Cover Image and Stego Image*

The image baboon.jpg with size of 512*512 was used in the research as the cover image to hide the ciphertext containing user's attribute given rise to the resultant stego- image. This was implemented in Java programming language. The cover and stego image were taken and then processed with Matrix Laboratory (MATLAB) version 2017a on Windows 10 with 8GB RAM size, and the processor of Intel(R) Core(TM), i3-5005U CPU @ 2.00GHz. Furthermore, the computer is a 64-bit operating system and x64-based processor. The cover image and stego image are represented on figure 3 (a) and (b) respectively. From the experimental result, it is difficult to differentiate between the Stego-image from the cover image, since the similarity between them is really high and the closer the stego image is to the cover image implies a greater security.



(a) (b)
Figure 3: (a) Cover Image and (b) Stego Image

Comparing the RGB histogram of both cover image and Stego image shows no significant difference graphically between them as depicted on figures 4 (a) and (b) respectively. Similarly, the closer the stego image is to the cover image implies a greater security.



(a) (b)

Figure 4: (a) Histogram of Cover Image and (b) Histogram of Stego Image

2.) *Mean Square Error (MSE) of the Stego image*

MSE describes a minimal non- perceptual error metric that is acquired from the cover image and Stego image where lesser values for MSE demonstrate negligible detectability on the image processed. In this research the MSE was used to measure the undetectability of the stego image. The smaller the value of MSE the lower the error rate and thus guaranteeing security of the stego image, (*Kamdar et al.*, (2013); *Hemalatha et al.*, (2013) and *Attay et al.*, (2017)). MSE is calculated with the given formula on (10).

$$MSE = \frac{1}{(M \times N)} \sum_{x=0}^M \sum_{y=0}^N (C_{ij} - S_{ij})^2 \quad (10)$$

“Where C denote the cover image, S is the Stego image, M and N are the width and height (i.e. M * N) of the cover image C and stego image S”. When MSE is calculated for baboon.jpg with 36.5 kb of data embedded in the image, the MSE is 0.076248. The MSE measures the undetectability of the proposed model with value falling within the standard range for measuring MSE. The proposed model guarantees security of Cloud user’s data.

3.) Peak Signal-to-Noise Ratio (PSNR) of the Stego image

PSNR refers to the ratio between a signal’s maximum power and the power of the signal’s noise, *Seyyed et al.*, (2014). Signals can have a broad dynamic series, so PSNR is basically measured in decibel (db), which is a logarithmic scale. In this research the MSE was used to measure the robustness of the stego image. A bigger PSNR value specifies a better feature of the Steganography algorithm used. The Human Visual System (HVS) would not be able to discern the images with PSNR greater than 36 db, *Seyyed et al.*, (2014). The PSNR of an image can be calculated using the following formula on (11).




$$PSNR = 10 \log_{10} \frac{Max^2}{MSE} db \quad (11)$$

Where Max^2 is the maximum pixel intensity that exists in the cover image and the PSNR of the same image, baboon.jpg is 59.308508db. Thus the PSNR of the image for the proposed model falls within the approved value range and proves that the Image quality is conserved at commendable level. Similarly, a high PSNR value is necessary in order to prevent the Stego image to be recognized by intruders that crawl in across the network.

4.) Results of the prototype system with other Images

The proposed model was used on other images such as lena.jpg and pepper.jpg apart from baboon.jpg. The results are shown in table 1 and 2 respectively. Table 1 shows the undetectability of the proposed model with value falling within the standard range for measuring it. Also, Table 2 demonstrates robustness of the proposed system with values falling within the specified standard range. This conserves the quality of the proposed system. Similarly, Figure 5 and 6 shows the respective graphs for the MSE and PSNR for the proposed model.

Table1: Comparison of Mean Square Error (undetectability) of the proposed model with the existing model based on different images

| S/No | Name of Image | Image | Image Size (bits) | Mean Square Error | |
|------|---------------|---|-------------------|-------------------|----------------|
| | | | | Proposed model | Existing model |
| 1. | Baboon.jpg |  | 512*512 | 0.0762 | N/A |
| 2. | Lena.jpg |  | 512*512 | 0.0794 | N/A |
| 3. | Peppers.jpg |  | 512*512 | 0.0790 | N/A |

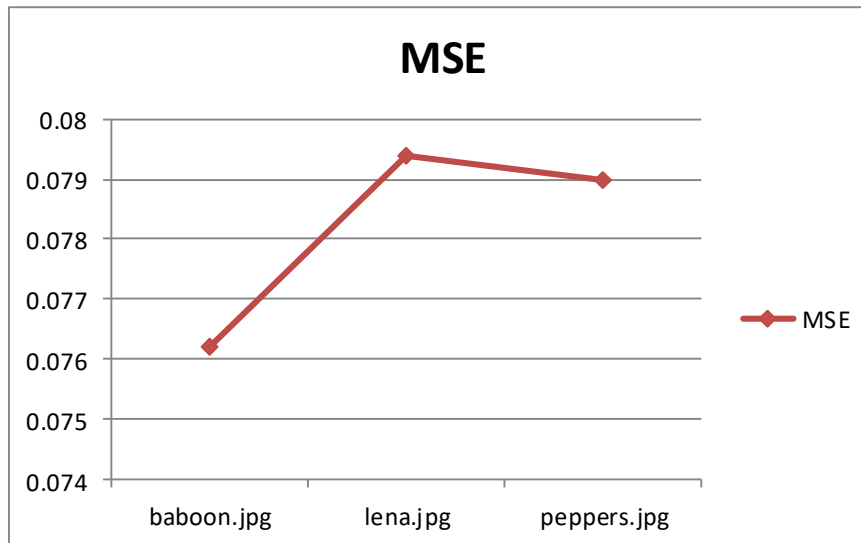





Figure 5: Mean Square Error of the proposed model

Table2: Comparison of Peak Signal-to-Noise Ratio (Robust) of the proposed model with the existing model

| S/No | Name of Image | Image | Image Size (bits) | PSNR (db) | |
|------|---------------|---|-------------------|----------------|----------------|
| | | | | Proposed model | Existing model |
| 1. | Baboon.jpg |  | 512*512 | 59.309 | N/A |
| 2. | Lena.jpg |  | 512*512 | 59.134 | N/A |
| 3. | Peppers.jpg |  | 512*512 | 59.157 | N/A |

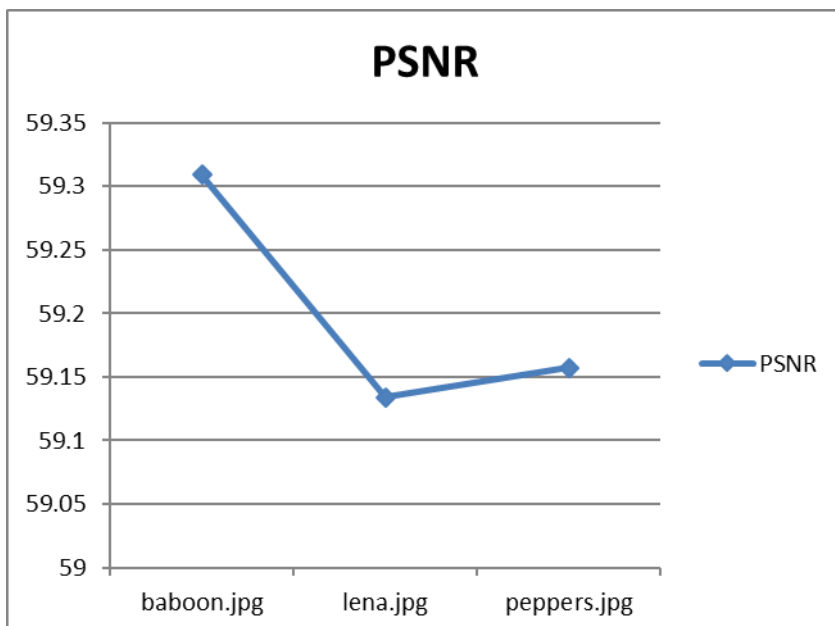


Figure 6: Peak Signal Noise Ratio of the proposed system

5. Discussion

Since the network is usually porous to attacks such as data leakage (sniffing, spoofing and malicious attack) user attributes are hereby endangered. The proposed model has been designed to ensure the security of user's attributes during transmission over the Network. The results of the proposed system have shown that the proposed model have performed more effectively and stronger. The proposed system performance has been evaluated with the performance parameters: Mean Squared Error (MSE) with value of 0.076248 for undetectability and Peak signal to noise ratio (PSNR) with value of 59.308508db for robustness. The performance has shown that the proposed system is undetectable, robust and effective for the security of the user's attributes. The proposed system has been tested on other images and the results have proved the proposed system strong and effective.

VI CONCLUSION

The paper presented an Enhanced PII Privacy Protection solution (using Advanced Encryption Standard AES-128 and Discrete Cosine Transform Modulus Three (DCT-M3) Steganography techniques) for protecting user's attributes from being leaked when transmitted and stored on the cloud. The supremacy of the proposed system over the existing system was also measured based on the encryption techniques used, undetectability and robustness of the Stego image. The result showed that the proposed system is undetectable, robust and effective for the security of the user's attributes on the cloud. Also the proposed system has been tested on other images and the results have proved the proposed system strong and effective for securing user's attributes.

REFERENCES

- M. M. Abur, S. B. Junaidu, S. Danjuma, S. Arlis, R. Ritonga, T. Herawan (2018): Towards a Privacy Mechanism for Preventing Malicious Collusion of Multiple Service Providers (SPs) on the Cloud. In: V. Bhateja, B. Nguyen, N. Nguyen, S. Satapathy, Le DN. (eds) Information Systems Design and Intelligent Applications. Advances in Intelligent Systems and Computing, Singapore: Springer, vol 672.
- M. M. Abur, O. S. Adewale & S. B. Junaidu, (2015): Cloud Computing Challenges: A review on Security and Privacy issues. Proceedings of the ACM International Conference on Computer Science Research and Innovations (CoSRI), Ibadan pp. 89-92.
- M. M. Abur, S. B. Junaidu, A. A. Obiniyi and S. E. Abdullahi (2018) "Privacy Protection and Collusion Avoidance Solution for Cloud Computing Users", 1st International Conference on Education and Development (ITED 2018), Base University, Abuja
- Y. A. Aldeen, M. Salleh & M. Abdur Razzaque, (2015): "A Survey Paper on Privacy Issue in Cloud Computing". *Research Journal of Applied Sciences, Engineering and Technology*, 10 (3): 328-337.

N. Aleisa (2015): A comparison of the 3DES and AES encryption standards. *International Journal of Security and its Applications* 9(7):241-246 <http://dx.doi.org/10.14257/ijisia.2015.9.7.21>.

P. N. Asha, T. Mahalakshmi, S. Archana and S. C. Lingareddy, (2016): Wireless Sensor Networks: A Survey on Security Threats Issues and Challenges. *International Journal of Computer Science and Mobile Computing*, 5(5), 249-267

Attaby A. A., Mursi Ahmed F. and Alsammak A. K., (2017) Data hiding inside JPEG images with high resistance to steganalysis using a novel technique: DCT-M3. *Ain Shams Engineering Journal* <http://dx.doi.org/10.1016/j.asej.2017.02.003>

Chadwick D. W. (n. d.). Federated Identity Management: Computer Laboratory, University of Kent, Canterbury, CT2, &NF, UK.

Chen D. & Zhao H. (2012): Data Security and Privacy Protection Issues in Cloud Computing. *Proc. of the 1st International conference on Computer Science and Electronics Engineering*, Hangzhou China. Doi: 10.1036/0071393722.

Cloud Security Alliance (CSA). 2013: The Nine Notorious Threats. Top threats working group.

Cloud Security Alliance (CSA). 2016: The Treacherous 12 - Cloud Computing Top Threats

Hacker News (2017): Researchers Crack 1024-bit RSA Encryption in GnuPG Crypto Library. Retrieved July 3, 2017 from wiki: <https://thehackernews.com/2017/07/gnupg-libgcrypt-rsa-encryption.html>

S., Hemalatha, A. U. Dinesh, A. Renuka, & P. R. Kamath (2013). A Secure and High Capacity Image Steganography Technique. *Signal & Image Processing: An International Journal (SIPIJ)* 4(1), 83-89.

N. P. Kamdar, D. G. Kamdar, D. N. khandhar, (2013). Performance Evaluation of LSB based Steganography for optimization of PSNR and MSE *Journal of Information, Knowledge and Research in Electronics and Communication Engineering* 2(2), 505-509.

Kaminsky A., Kurdziel M. & Radziszowski S. (2010). An overview of cryptanalysis research for the advanced encryption standard (AES). Military Communications Conference (MILCOM), San Jose, USA. Pp1-8.

Y. Kumar, R. Munjal and H. Sharma, (2011) Comparison of Symmetric and Asymmetric Cryptography with Existing Vulnerabilities and Countermeasures. *International Journal of Computer Science and Management Studies (IJCSMS)* 11(3), 60-63.

M. A. P. Leandro, T. J. Nascimento, D. Santos, C. M. Westphall & C. B. Westphall (2014). Multi-Tenancy Authorization System with Federated Identity for Cloud-Based Environments Using Shibboleth. *In proceeding of the Eleventh International Conference on Networks (NetWare2014)*, Lisbon, Portugal. pp. 42-67.

- U. Lokhande and A. K. Gulve (2014): Steganography using Cryptography and Pseudo random numbers. *International Journal of Computer Applications*, 96 (19), 41-45.
- T. Orawiwattanakul, K. Yamaji, M. Nakamura, T. Kataoka & N. Sonehara (2010): "User-controlled privacy protection with attribute-filter mechanism for a Federated SSO environment using Shibboleth," in P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), International Conference on IEEE, pp.243-249.
- A. S. Seyyed & N. Ivanov (2014). Statistical Image Classification for Image Steganographic Techniques I.J. Image, Graphics and Signal Processing, 8, 19-24 DOI: 10.5815/ijigsp.2014.08.03
- J. Song, K. Lee, and H. Lee, (2014). Biclique Cryptanalysis on the Full Crypton-256 and mCrypton-128. *Journal of Applied Mathematics*. 2014, 1-10, <http://dx.doi.org/10.1155/2014/529736>.
- S. Suriadi, E. Foo and A. Josang (2007). A User-Centric Federated Single-On System. IFIP International Conference on Network and Parallel Computing Workshops.
- R. Smith: Understanding encryption and cryptography basics (2003) Retrieved August 15, 2018 from wiki <https://searchsecurity.techtarget.com/Understanding-encryption-and-cryptography-basics>
- SWITCH, (2010). "uapprove - user consent module for shibboleth identity providers," retrieved: [Online]. Retrieved from: <https://www.switch.ch/aai/support/tools/uApprove.html/03/03/2016>
- A. M. Teena and M. Aaramuthan (2017): Federated Cloud Identity Management: A Study on Privacy Tactics, Tools and Technologies. *Journal of Computer Engineering (IOSR-JCE)* e-ISSN: 2278-0661, p-ISSN: 2278-8727, 19(6), 34-40.
- R. Tripathi & S. Agrawal (2014). Comparative Study of Symmetric and Asymmetric Cryptography Techniques. *International Journal of Advance Foundation and Research in Computer (IJAFRC)* 1(6), 68-76.
- R. Weingartner, C. M. Westphall, (2014) "Enhancing privacy on identity providers", Emerging Security Information Systems and Technologies (SECURWARE). The *Eighth International Conference* Lisbon, Portugal pp. 1-7.
- M. Zhou, R. Zhang, W. Xie, W. Quian & A. Zhou, (2010) Security and Privacy in cloud: Survey. In Proc. Of the 6Th *International Conference on Semantics, Knowledge and Grids, IEEE*. Pages 105-112.

Using Data Mining Classifiers to Predict Academic Performance of High School Students

Yecheng Yao, The University of Chicago (Chicago, USA)
Zebang Chen, The University of California, San Diego (San Diego, USA)
Sumin Byun, Hankuk Academy of Foreign Studies (Yongin, Korea)
Yizhu Liu, Pius XI Catholic High School (Milwaukee, USA)

ABSTRACT: The use of data mining techniques for educational datasets is being referred to as educational data mining. This study uses popular classifier algorithms in data mining with secondary school student data to estimate their success rate. Student success depends on various factors related to the student's personal, family and surrounding environment, among others. This study's dataset has attributes related to parental education, job information, student travel time, study time, financial status, extracurricular activities, access to the Internet, family relationship, alcoholic consumption, student health condition, regular school attendance. This study analyzes the correlations between these attributes and identifies the attributes that contribute to students' test achievement for better prediction and management of student performance. The study also compares the performance of top classification algorithms in data mining and concludes J48 classifier and oneR to outperform the other classifiers.

KEYWORDS: Data Mining, Educational Data Mining, Classifier, High School Data Mining

I. Introduction

Data mining entails extracting knowledge from large volumes of data. This work attempts to uncover insightful patterns and relationships useful for the decision-making process. Data from various sources are processed using various methods and algorithms to uncover useful patterns and insights. Data mining and knowledge discovery have gained increasing attention for their usefulness for decision making and now have become an essential part of any organization, including educational institutions. Schools and teachers need to know their students' academic achievement in terms of what factors influence that achievement. To better estimate of student performance, data mining algorithms and statistical methods have been used for analysis purposes.

The application of data mining to educational datasets is referred to as educational data mining (EDM), which uses data mining techniques, machine learning methods, and statistical analysis methods. EDM is expected to innovate novel methods to mine educational data for useful insights on student learning and achievement and the factors that influence them.

Educational data mining focuses on analysing students' academic data sets, classifying students using decision trees, generating some association rules for better decisions for enhancing

academic performance. The main goal is to predict students' future learning and academic performance, and this goal depends on identifying all attributes related to learning characteristics and behaviors. Another goal is to identify students' interests so that the curriculum could be planned to accommodate their educational standards and learning styles.

The rest of this paper is organized as follows: Section II provides a literature review. Section III discusses the artificial neural network and the multi-layer perceptron, and Section IV provides the materials and methods. Section V discusses the results, and Section VI concludes.

II. Literature Review

Saibaba et al. [1] analyzed student performance using various clustering techniques with the WEKA tool. The proposed system considered student data with 10th percentage, intermediate percentage, and B.Tech I Year, II Year and III Year marks using decision trees. They analyzed these decision trees and forecasted the likelihood of students getting jobs after graduation. They also proposed to consider attributes such as social networking interests, parental economic status, and parental educational achievement.

Ahmed et al. [2] compared four data mining techniques including J48 decision trees, MLP, NB, and SMO and found that the attribute evaluation method to be helpful in predicting instructor performance. The results for the attributed evaluation method show that SMO outperformed other algorithms with 85.5% accuracy. J48 DT outperformed other algorithms with 84.8% accuracy.

Kalpana and Venkatalakshmi [3] presented a case study on educational data mining and discussed how data mining can be used in higher education for graduate student performance. They used engineering students' data for 5 years and applied data mining techniques, using various clustering methods along with distance- and density-based approaches to improve prediction accuracy for graduate student performance.

Brijesh Kumar and Saurabh [4] used various data mining techniques to achieve high quality in the higher education system by extracting insights into students' exam performance. They used a dataset from 50 students from VBS Purvanchal University (Jaunpur of MCA course) from 2007 to 2010.

Kalpesh et al. [5] proposed an architecture to more accurately predict student performance. They considered 17 student attributes using K-means clustering and Naïve-Bayes algorithms and concluded that Naïve Bayes outperformed with 96% accuracy.

Taier et al. [6] discussed the effectiveness of data mining in improving the performance of graduate students by using data mining techniques such as association rules, classification, clustering, and outlier detection.

Cortez et al. [7] proposed a method for predicting student grades by analyzing past grades and found that school, family, and social environments to be key factors. They employed binary and five-level classifications and regression using decision tree, random forest, neural network, and support vector machine.

Jovanovic et al. [8] used classification techniques to analyze and predict student academic performance and applied clustering to student groups based on e-learning, concluding that their model can help identify students' academic strength.

Pandey and Pal [9] examined student performance using 600 students from Dr. R.M.L. Awadh University, Faizabad, India, and used Bayes classification in category, language and background qualifications to predict new students' academic performance.

Hijazi and Naqvi [10] analyzed student performance using 300 students (225 males and 75 females) from Punjab University of Pakistan and considered student attributes including student attitudes towards class attendance, student family income, maternal age, maternal education level, and hours spent studying. They used simple linear regression and concluded maternal education level and student family income to be highly correlated with academic performance.

Khan [11] conducted a performance analysis of 400 students composed of 200 males and 200 females from Senior Secondary School of Aligarh Muslim University, Aligarh. These students were selected using cluster sampling where the whole population was divided into groups or clusters. They concluded that females with high socioeconomic status showed higher science achievement and males of low socioeconomic status, higher academic achievement in general.

Pandey and Pal [12] conducted a performance analysis of 60 students from Dr. R. M. L. Awadh University, Faiza bad, India, using association rule mining and found students' interest in class teaching language.

Bray [13] examined private tutoring and observed the percentage of students receiving private tutoring in India to be higher than that in Malaysia, Singapore, Japan, China, and Sri Lanka. They also observed enhanced academic performance through private tutoring.

Ayesh et al. [14] used the K-means clustering algorithm to predict students' learning activities and concluded the model to be useful in predicting academic performance.

Al-Radaideh et al. [15] used the decision tree model to predict final grades of students taking a C++ course at Yarmouk University, Jordan. They compared three classification techniques including ID3, C4.5 and Naïve Bayes and concluded the decision tree model to outperform other models.

III. Materials and Methods

Dataset:

This study considers data on student achievement in secondary education of two Portuguese schools. Data attributes include student grades, demographic, social, and school-related factors, and the data were collected using school reports and questionnaires. Two datasets focused on two distinct subjects: Mathematics (mat) and Portuguese language (por). These two datasets were combined. Here the target attribute G3 has a strong correlation with attributes G2 and G1, which is due to G3 being the final year grade (issued in the 3rd period), while G1 and G2 correspond to 1st and 2nd period grades. Predicting G3 without G2 and G1 is more difficult, but this kind of prediction is considered to be much more useful.

Table 1 shows the attributes.

Table 1. Various attributes

| Attribute | Description (Domain) |
|------------------|--|
| Sex | student's sex (binary: female or male) |
| Age | student's age (numeric: from 15 to 22) |
| school | student's school (binary: Gabriel Pereira or Mousinho da Silveira) |
| address | student's home address type (binary: urban or rural) |
| Pstatus | parent's cohabitation status (binary: living together or apart) |
| Medu | mother's education (numeric: from 0 to 4a) |

| | |
|------------|--|
| Mjob | mother's job (nominalb) |
| Fedu | father's education (numeric: from 0 to 4a) |
| Fjob | father's job (nominalb) |
| guardian | student's guardian (nominal: mother, father or other) |
| famsize | family size (binary: ≤ 3 or > 3) |
| famrel | quality of family relationships (numeric: from 1 – very bad to 5 – excellent) |
| reason | reason to choose this school (nominal: close to home, school reputation, course) |
| traveltime | home to school travel time (numeric: 1 – < 15 min., 2 – 15 to 30 min., 3 – 30 min. to 1) |
| studytime | weekly study time (numeric: 1 – < 2 hours, 2 – 2 to 5 hours, 3 – 5 to 10 hours or 4 – > 10) |
| failures | number of past class failures (numeric: n if $1 \leq n < 3$, else 4) |
| schoolsup | extra educational school support (binary: yes or no) |
| famsup | family educational support (binary: yes or no) |
| activities | extra-curricular activities (binary: yes or no) |
| paidclass | extra paid classes (binary: yes or no) |
| internet | Internet access at home (binary: yes or no) |
| nursery | attended nursery school (binary: yes or no) |
| higher | wants to take higher education (binary: yes or no) |
| romantic | with a romantic relationship (binary: yes or no) |
| freetime | free time after school (numeric: from 1 – very low to 5 – very high) |
| Gout | going out with friends (numeric: from 1 – very low to 5 – very high) |
| Walc | weekend alcohol consumption (numeric: from 1 – very low to 5 – very high) |
| Dalc | workday alcohol consumption (numeric: from 1 – very low to 5 – very high) |
| Health | current health status (numeric: from 1 – very bad to 5 – very good) |
| absences | number of school absences (numeric: from 0 to 93) |
| G1 | first-period grade (numeric: from 0 to 20) |
| G2 | second-period grade (numeric: from 0 to 20) |
| G3 | final grade (numeric: from 0 to 20) |

Table 1 provides the following insights:

- i. There are 1044 instances and 33 attributes for both Mathematics (mat) and Portuguese language (por) datasets.
- ii. G3 is the output label (all 32 attributes other than G3 are independent variables for the dependent variable G3).
- iii. G3 has a range of [0, 20], and the classification model must predict 1 class out of the possible class labels.
- iv. There is a mix of numerical and nominal attributes. There are no missing values for any given attributes.

IV. Results & Discussion

The downloaded dataset was in .csv format, and this was converted into .arff to accommodate the WEKA environment. The converted .arff file was given as input for the WEKA explorer, as shown as follows:

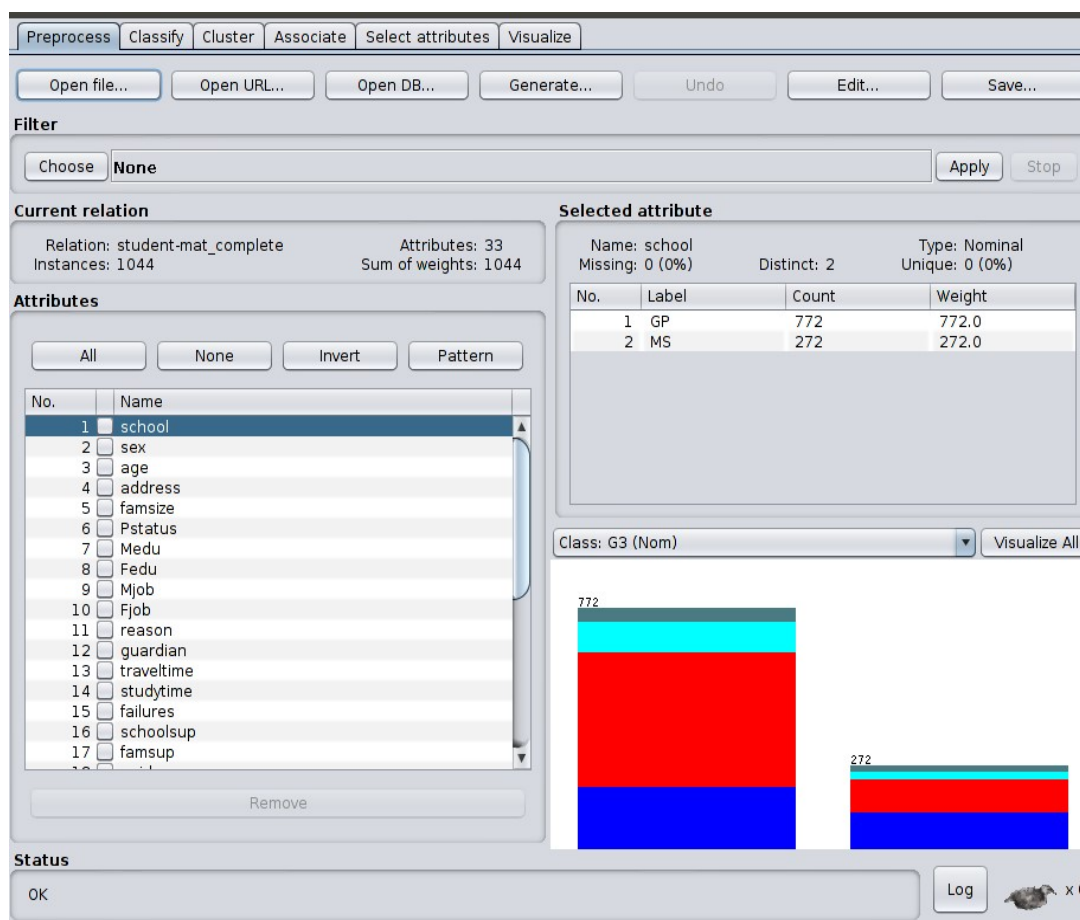


Figure 1. Student performance dataset on WEKA

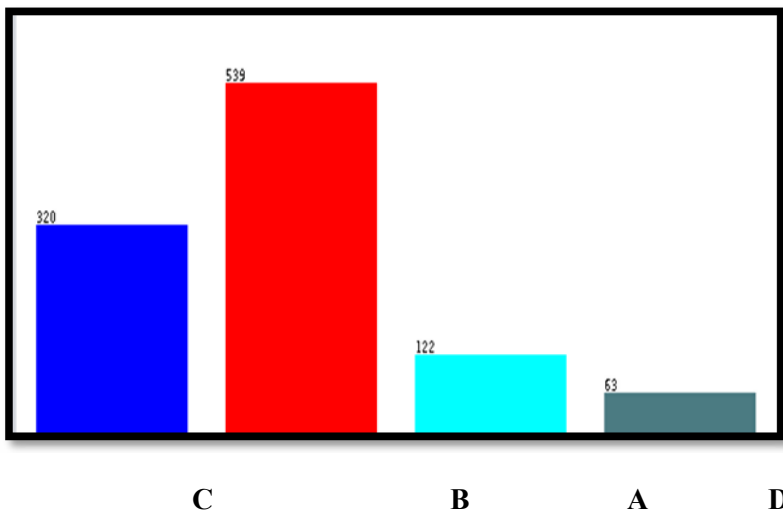
The experiment was conducted in three stages: data preprocessing, data classification without feature selection, and data classification with feature selection.

1. Data Preprocessing:

In the data preprocessing phase, target attributes are identified, and the dataset is loaded into WEKA for cleaning and preprocessing. The target attribute is categorized into 4 classes A, B, C, D, as shown in Table 2.

Table 2. Mapping of clusters based upon initial values

| The range of initial class | New Class Label |
|----------------------------|-----------------|
| 0 ~ 5 | D |
| 6 ~ 10 | C |
| 11 ~ 15 | B |
| 16 ~ 20 | A |



2. Implement of classifiers on full dataset

| Classifiers | Correctly Classified Instances (%) | In-correctly Classified Instances (%) | Kappa statistic | Mean absolute error | Root mean square error |
|--------------------------|---|--|------------------------|----------------------------|-------------------------------|
| ZeroR (baseline) | 51.63 | 48.37 | 0 | 0.3114 | 0.3944 |
| OneR | 84.19 | 15.81 | 0.7421 | 0.079 | 0.2811 |
| J48 | 81.32 | 18.68 | 0.6971 | 0.1171 | 0.2839 |
| Naive bayes | 76.53 | 23.47 | 0.6332 | 0.1303 | 0.2964 |
| IBk (k-nearest neighbor) | 45.98 | 54.02 | 0.1283 | 0.2706 | 0.5186 |
| Decision Tree | 83.81 | 16.19 | 0.7351 | 0.1388 | 0.2477 |
| Logistic Regression | 81.61 | 18.39 | 0.7023 | 0.1125 | 0.2663 |
| PART | 78.83 | 21.17 | 0.6587 | 0.1135 | 0.3133 |
| RandomForest | 82.38 | 17.62 | 0.7099 | 0.1618 | 0.2615 |
| JRip | 81.90 | 18.10 | 0.7113 | 0.1273 | 0.2715 |

Table 3. Comparison of results for different classifiers

Table 3 shows the results for different classifiers for correctly classified instances, incorrectly classified instances, Kappa statistic, mean absolute error, and root mean square error. Using the above method, OneR shows a higher accuracy value of 84.19, while ZeroR (baseline) shows the lowest value of 51.63. Figures 3 and 4 provide the graphical representation of results for different classifiers.

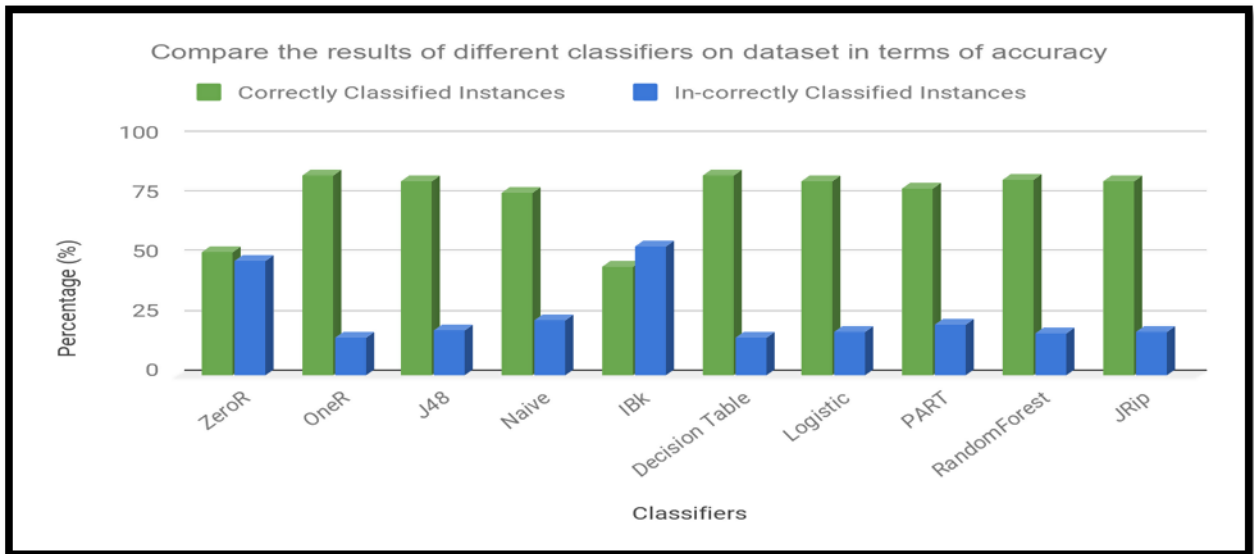


Figure 3. Graphical representation of different classifiers

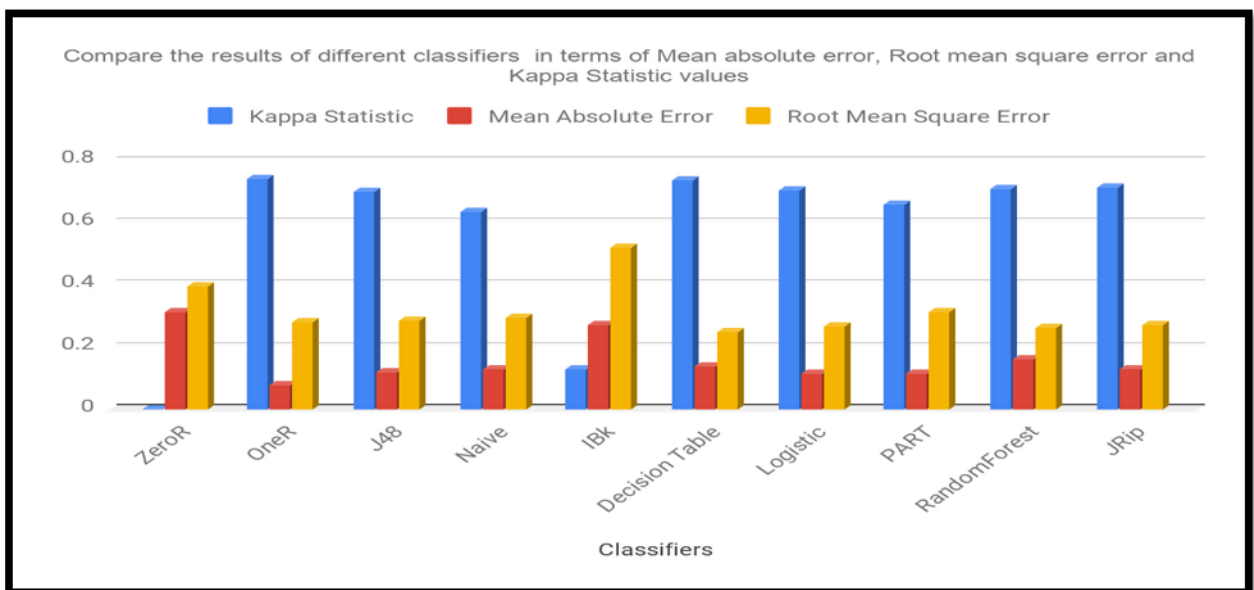


Figure 4. A comparison of accuracy for MASE, RMSEs, Kappa statistic

Table 4. A comparison of different classifiers for execution time

| Classifiers | Time taken to build a model (in a sec) |
|-------------|--|
| ZeroR | 0.001 |
| OneR | 0.01 |

| | |
|----------------|-------|
| J48 | 0.14 |
| Naive Bayes | 0.03 |
| IBk | 0.001 |
| Decision Table | 0.38 |
| Logistic | 0.54 |
| PART | 0.19 |
| RandomForest | 0.47 |
| JRip | 0.31 |

Table 4 compares different classifiers in terms of execution time (seconds) without using feature selection for the given dataset. ZeroR and IBk (k-nearest neighbor) take the shortest time, and logistic classifiers take the longest execution time. Figure 5 graphically shows the execution time for different classifiers.

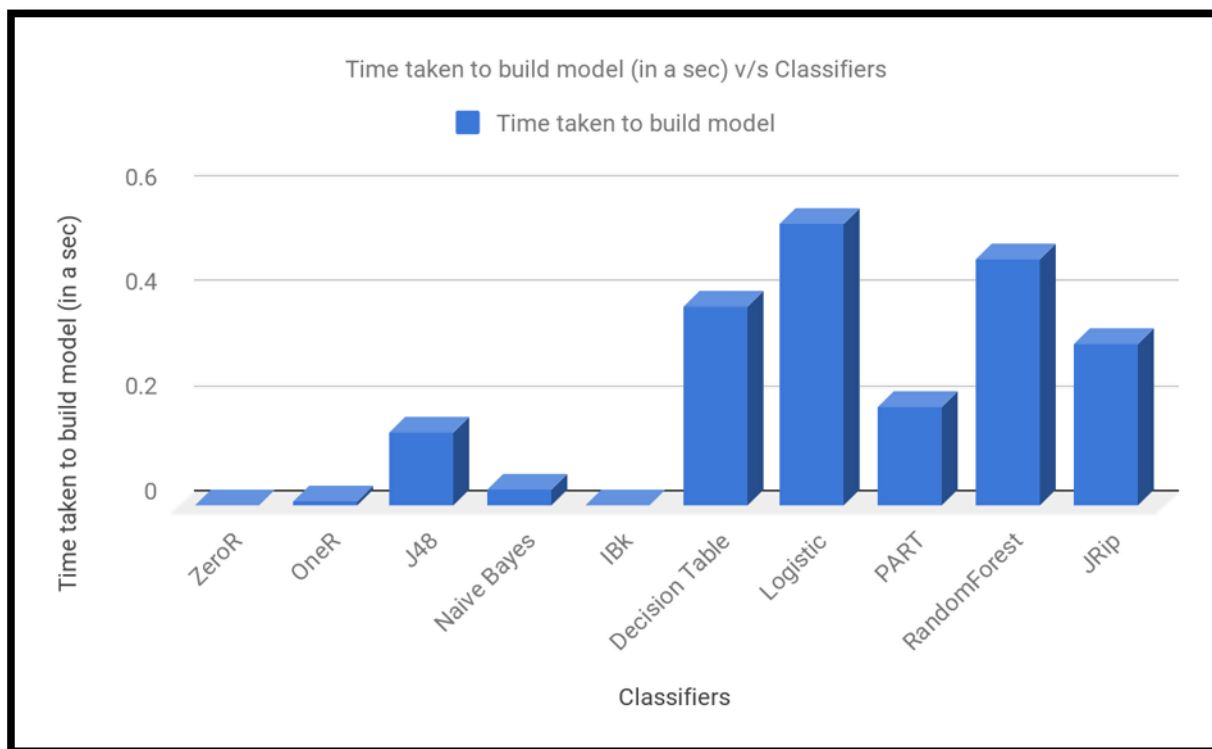


Figure 5. Time taken to build the classifier

3. Feature Selection

Feature selection is used to select attributes from the dataset. The feature evaluator and search method was employed to perform feature selection, and based on selected attributes techniques, the following 6 features/attributes were selected:

- (i) famsize
- (ii) Fedu
- (iii) failures
- (iv) absences
- (v) G1 and
- (vi) G2 for better results from the dataset

After selecting these attributes, the dataset was prepared to implement different classifiers to compare with and without feature selection method.

Table 5. A comparison of results for different classifiers for feature selection

| Classifiers | Correctly | In-correctly | Kappa | Mean | Root mean |
|-------------|-----------|--------------|-------|------|-----------|
| | | | | | |

| | Classified Instances (%) | Classified Instances (%) | statistic | absolute error | square error |
|--------------------------|---------------------------------|---------------------------------|------------------|-----------------------|---------------------|
| ZeroR (baseline) | 51.63 | 48.37 | 0 | 0.3114 | 0.3944 |
| OneR | 84.19 | 15.81 | 0.7421 | 0.079 | 0.2811 |
| J48 | 84.39 | 15.61 | 0.7465 | 0.1142 | 0.2515 |
| Naive bayes | 81.23 | 18.77 | 0.7026 | 0.1197 | 0.2672 |
| IBk (k-nearest neighbor) | 76.53 | 23.47 | 0.6185 | 0.1193 | 0.3316 |
| Decision Tree | 84.20 | 15.80 | 0.7434 | 0.1353 | 0.2429 |
| Logistic Regression | 84.10 | 15.90 | 0.7403 | 0.1147 | 0.2413 |
| PART | 81.03 | 18.97 | 0.6948 | 0.1153 | 0.2728 |
| RandomForest | 80.84 | 19.16 | 0.6912 | 0.1153 | 0.263 |
| JRip | 83.24 | 16.76 | 0.7305 | 0.1271 | 0.263 |

Table 5 shows the results for different classifiers with the feature selection method for correctly classified instances, incorrectly classified instances, Kappa statistic, mean absolute error, and root mean square error. With this method, J48 shows a higher accuracy value of 84.39, and ZeroR (baseline), the lowest value, 51.63. Figures 6 and 7 graphically represent different classifiers results.

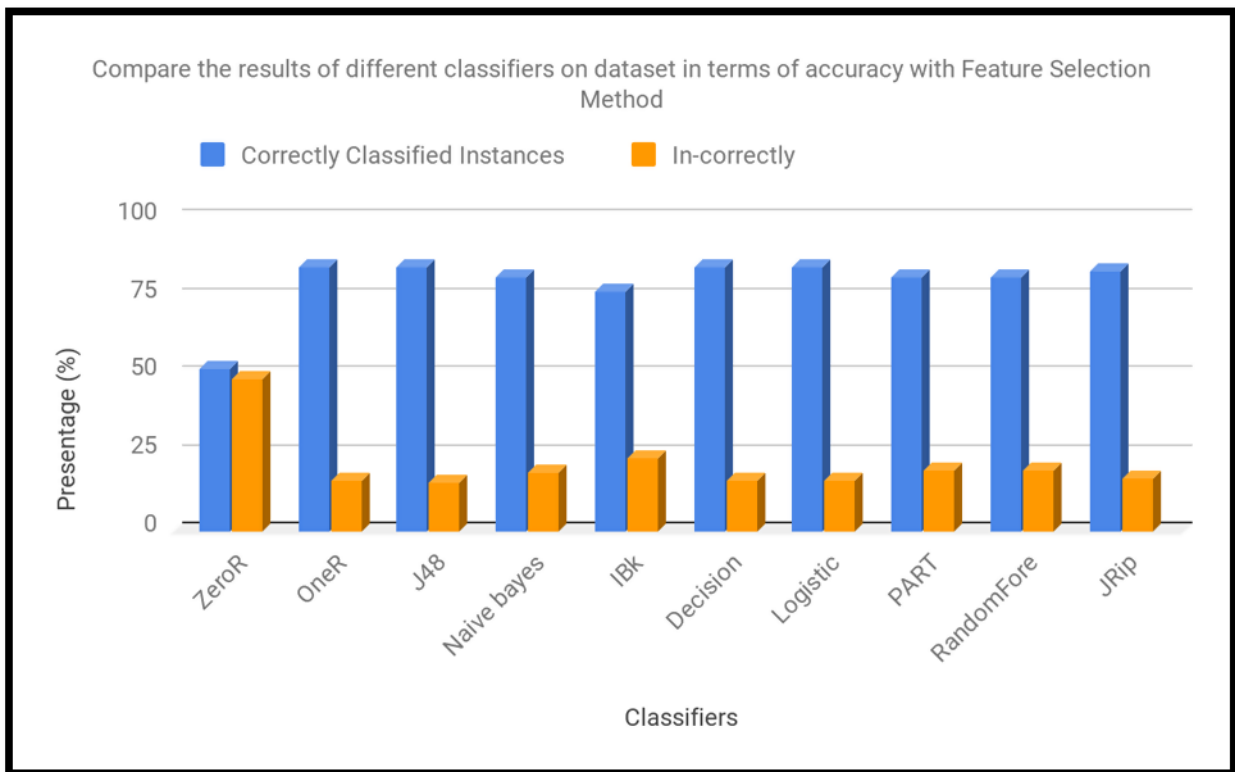


Figure 6. Number of correctly and incorrectly classified instances

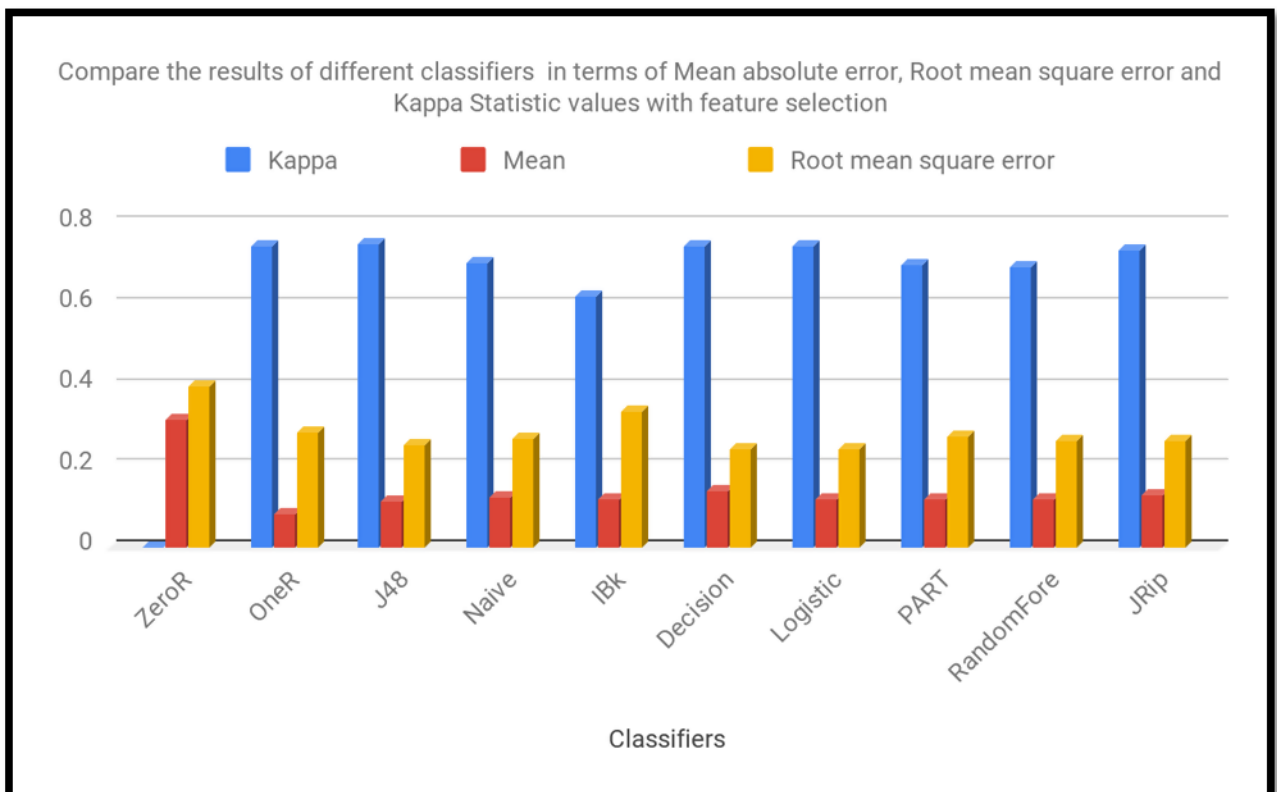


Figure 7. Accuracy for different classifiers

Table 6 compares different classifiers for execution time (seconds) with and without feature selection. Table 6 shows that ZeroR and IBk (k-nearest neighbor) take the shortest time with and without feature selection. The longest time is taken by Logistic classifiers without feature selection, and RandomForest takes the longest time with feature selection.

Table 6. A comparison of different classifiers for execution time

| Classifiers | Time taken to build model (in a sec) | Time taken to build model with feature Selection (in a sec) |
|---------------------|---|--|
| ZeroR | 0.001 | 0.001 |
| OneR | 0.01 | 0.01 |
| J48 | 0.14 | 0.11 |
| Naive Bayes | 0.03 | 0.02 |
| IBk | 0.001 | 0.001 |
| Decision Tree | 0.38 | 0.10 |
| Logistic Regression | 0.54 | 0.11 |
| PART | 0.19 | 0.06 |
| RandomForest | 0.47 | 0.29 |
| JRip | 0.31 | 0.18 |

Figure 8 graphically shows different classifiers for execution time using both methods.

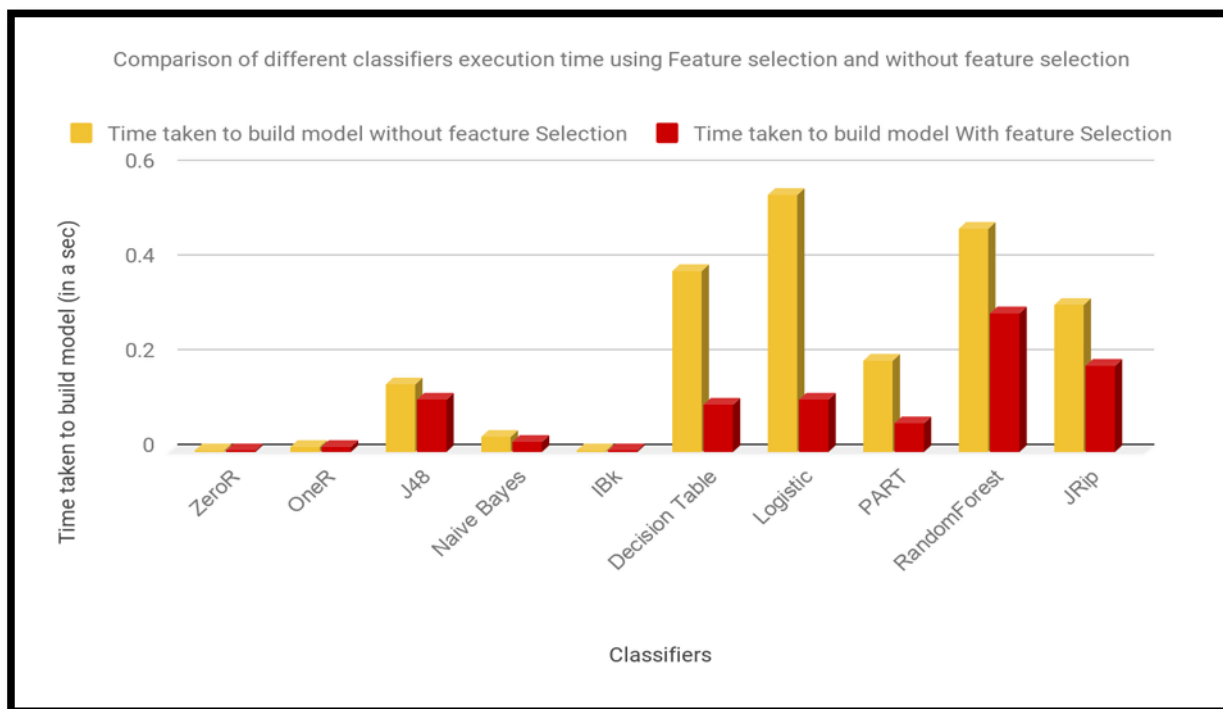


Figure 8. Time taken to build a model with and without feature Selection

Table 7 shows the results for different classifiers for accuracy with and without feature selection. The highest accuracy is shown for J48 with feature selection.

Table 7. A comparison of different classifiers for accuracy

| Classifiers | Correctly Classified Instances with feature selection (%) | Correctly Classified Instances without Feature Selection (%) |
|---------------------|---|--|
| ZeroR | 51.63 | 51.63 |
| OneR | 84.19 | 84.19 |
| J48 | 84.39 | 81.32 |
| Naive bayes | 81.23 | 76.53 |
| IBk | 76.53 | 45.98 |
| Decision Tree | 84.2 | 83.81 |
| Logistic Regression | 84.1 | 81.61 |
| PART | 81.03 | 78.83 |

| | | |
|--------------|-------|-------|
| RandomForest | 80.84 | 82.38 |
| JRip | 83.24 | 81.9 |

Figure 9 graphically shows the results for different classifiers for accuracy (correctly classified instances).

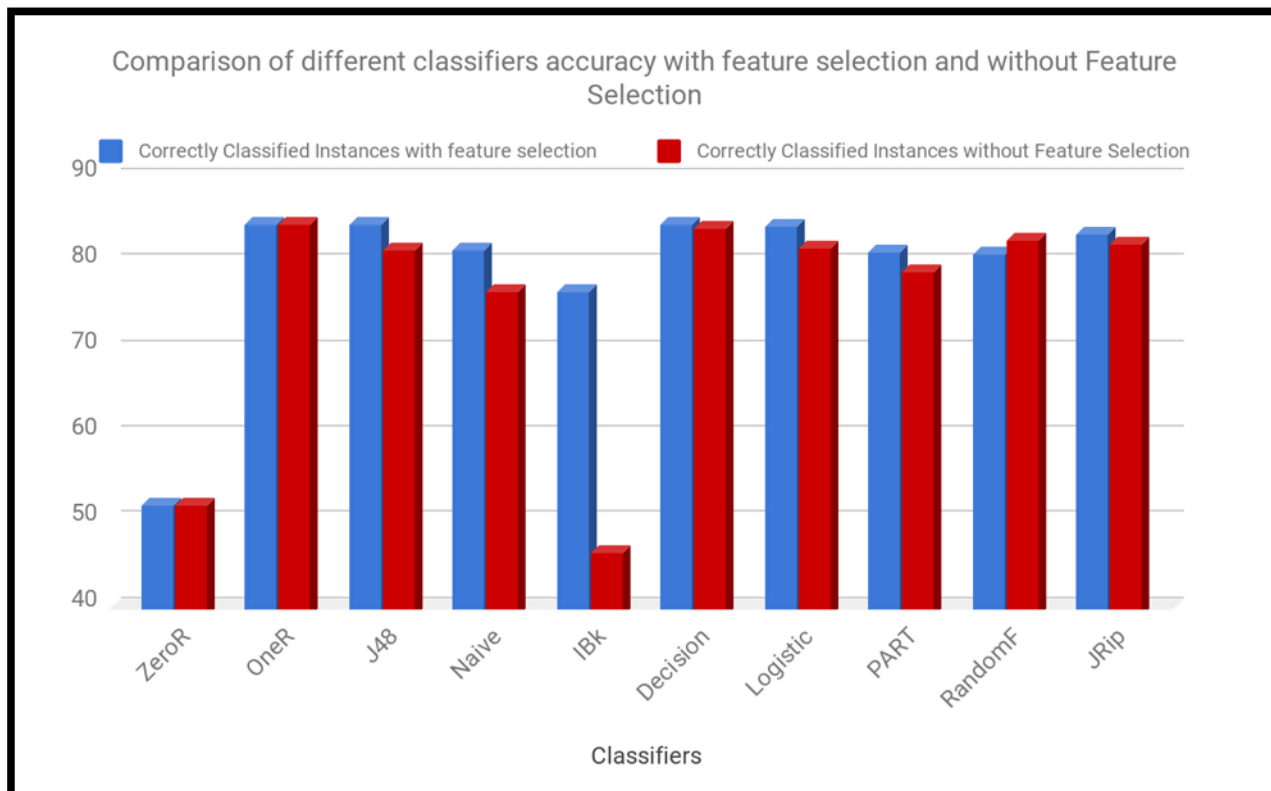


Figure 9. A comparison of different classifiers

V. Conclusions

This paper mines data from high school students to predict their academic performance. The dataset included relevant educational and personal attributes. Top 10 classifiers were identified based on previous research. The results show that without feature selection, oneR provided the best performance, followed by the decision tree, random forest and J48. The results were tabulated, and corresponding plots were obtained. Then the feature selection method was applied to identify famsize, fedu, failures, absence, G1 and G2 as the required attributes. According to the results, J48 showed the best performance with fewer attributes, followed by oneR, decision tree and logistic regression. The tabulated and plotted results suggest that J48 and oneR provided good results for classifying the dataset and predicting student performance.

References

- [1] Sai Baba et al, "Student Performance Analysis Using Classification Techniques", International Journal of Pure and Applied Mathematics, Vol.115, No.5 2017, pp.1-7
- [2] Mohamed Ahmed et al., "Using Data Mining to Predict Instructor Performance", 12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2016, pp. 137-142
- [3] Kalpana and Venkatalakshmi, "Intellectual Performance Analysis of Students by Using Data Mining Techniques", International Journal of Innovative Research in Science, Engineering and Technology, Vol.3, 2014.
- [4] Brijesh Kumar Baradwaj and Saurabh Pal, "Mining Educational Data to Analyze Students Performance", International Journal of Advanced Computer Science and Applications", Vol. 2, No. 6, 2011.
- [5] Kalpesh et al., "Student Performance Prediction System using Data Mining Approach", International Journal of Advanced Research in Computer and Communication Engineering, Vol.6, Issue 3, 2017.
- [6] Tair and El-Halees 2012. Mining Educational Data to Improve Student's performance: A Case Study. International Journal of Information and Communication Technology Research.
- [7] P. Cortex and A. Silva. Using data mining to predict secondary school student performance.
- [8] M. Jovanovic, M. Vukicevic, M. Milovanovic and M. Minovic (2012). Using data mining on student behaviour and cognitive style data for improving e-learning systems: a case study. International Journal of Computational Intelligence Systems.
- [9] U. K. Pandey and S. Pal, Data Mining: A Prediction of performer or under performer using classification , (IJCSIT), International Journal of Computer Science and Information Technology, Vol 2(2), pp. 686-690.
- [10] S. T. Hijazi and R.S.M.M. Naqvi, "Factors affecting students performance: A case of Private Colleges". Bangladesh e-Journal of Sociology, Vol. 3 No.1, 2006.
- [11] Z. N. Khan , "Scholastic achievement of higher secondary students in science stream", Journal of Social Sciences, Vol. 1. No. 2 , pp. 84-87, 2005.
- [12] U. K. Pandey and S. Pal , " A Data mining view on class room teaching language", (IJCSI), International Journal of Computer Science Issue, Vol. 8, Issue 2, pp. 277-282, 2011.
- [13] M. Bray, The Shadow education system : Private tutoring and its implications for planners, (2nd edition), UNESCO, PARIS, France, 2007.

[14] Sheela Ayesha et. al , “Data mining model for higher education system”, European Journal of Scientific Research”, Vol. 43, No.1. pp. 24-29, 2010.

[15] Q. A. Al-Radaideh, et. al., “ Mining Student data using decision trees”, International Arab Conference on Information Technology (ACIT2006), Yarmouk University, Jordan, 2006.

Cyber security European standards in business

Maksim Iavich¹, Sergiy Gnatyuk², Giorgi Iashvili¹ Andriy Fesenko³

1. Caucasus University 2. National Aviation University 3. Taras Shevchenko Kyiv National University

ABSTRACT. In the paper we consider the attacks on large and small businesses. We analyze the European standards and legislation. In the paper we also describe the European trainings materials. Using these materials is made the experiment where we have collected the group of 10 people and assessed them using cyber security exercises created using the attacks on Georgia and Ukraine. Afterwards we trained these students for 20 hours using ENISA materials and made them to answer the similar questions once more. We have got rather good results. Based on our research we offer corresponding recommendations for the representatives of small and big businesses.

KEYWORDS: cyber security, European standards, security in business

Nowadays, entrepreneurs clearly know that a cyber attack carries not only potential financial and information losses, but also damages the company's reputation, which threatens to drain customers and reduce investment attractiveness.

The Internet is becoming more and more dangerous every year and continues to be commercialized. This contributes to the fact that the motives of bad hackers, or the so-called "black hats" are becoming more greedy, so businesses should prepare for such possible hacker attacks in a timely manner. Organizations must improve IT security system so that sensitive data does not leak. The must care about security of the company in advance, involving employees and educating them.

Here are three main aspects of information security:

Confidentiality – people out of the organization and employers that are not intended to see sensitive data must not have access to it;

Integrity – existing data in the system must not be changed;

Availability - employees and authorized clients can access the necessary information at any time;

Small businesses are also victims of hacker attacks. If earlier cybercrime was chosen mainly by large companies, now no small firms or even individual entrepreneurs are insured against its "networks". A loud confirmation of this was the breaking of the OneLogin cloud authorization service, which occurred in June 2017. As a result, attackers gained access to authentication data of more than 2000 companies from almost fifty countries of the world.

According to the opinion cyber security experts, small business is very interesting for modern hackers. Such enterprises have large stocks of finance and other resources than ordinary users. At the same time, the level of their protection is noticeably lower in comparison with large firms and corporations. This combination makes small businesses vulnerable and, at the same time, attractive to cybercriminals of all levels. Hackers are aimed at small companies and individual entrepreneurs due to the fact that the latter do not pay enough

attention to their cyber security. They are confident that the attackers will not waste time on hacking their Internet resources, and therefore underestimate the potential level of risk.

European companies have a very big practice in cyber security. In 2015, almost every 5th European company faced the risk of data loss or theft due to cyber attacks.

The European Network and Information Security Agency (ENISA, European Union Agency for Network and Information Security) has signed a memorandum with the largest companies in the semiconductor industry on the joint development of cybersecurity on the European continent.

The memorandum “General positions on cyber security issues” was signed by Infineon, NXP, STMicroelectronics and ENISA.

Actually, the development of recommendations for priority actions to strengthen cybersecurity, a kind of "road map", is the subject of the memorandum. There is a huge need for a clear identification of the main threats, standardization, the introduction of basic levels of cybersecurity, the development of procedures and measures for the development and implementation of cybersecurity technologies at the software, network and hardware level. The document is rather interesting both for specialists and for ordinary citizens, the safety and well-being of which increasingly depends on the strength of the barrier against invisible threats on the Web.

From the memorandum we can see, that it is very important in cyber security to use the proven solutions and the generally accepted level of security and confidentiality of data attached to the Network and “ smart ” devices - both of these solutions are necessary, and we recommend their implementation, so that Europe takes full advantage of the Internet of Things. As such, standardization and certification are identified as priority areas to accelerate the implementation of a cybersecurity level system that allows citizens, organizations, and institutions to gain trust in the network environment.

As part of the EU Cybersecurity strategy the European Commission proposed the EU Network and Information Security directive. The NIS Directive is the first piece of EU-wide cybersecurity legislation.

NIS Directive has three basic parts:

- National capabilities: EU Member States must have certain national cybersecurity capabilities of the individual EU countries, e.g. they must have a national CSIRT, perform cyber exercises, etc.
- Cross-border collaboration: Cross-border collaboration between EU countries, e.g. the operational EU CSIRT network, the strategic NIS cooperation group, etc.
- National supervision of critical sectors: EU Member states have to supervise the cybersecurity of critical market operators in their country: Ex-ante supervision in critical sectors (energy, transport, water, health, and finance sector), ex-post supervision for critical digital service providers (internet exchange points, domain name systems, etc).

The most weak layer in cyber security is the person. According our research the training of employees greatly decrease the number of successful cyber-attacks on the organization[2].

“ENISA” - The European Union Agency for Cybersecurity is working on making Europe cyber secure since 2004. It cooperates with members states and the private sector to improve the capabilities of defenses against cyber attacks. It also supports the development of a cooperative response to large-scale cross-border cybersecurity incidents or crises. Since 2019 the agency works on creating cybersecurity certification schemes[1].

ENISA CSIRT training material was introduced in 2008. In 2012, 2013 and 2014 it was complemented with new exercise scenarios containing essential material for success in the CSIRT community and in the field of information security. In these pages you will find the ENISA CSIRT training material, containing Handbooks for teachers, Toolsets for students and Virtual Images to support hands on training sessions. In order to deliver trainings more efficiently with better and longer lasting results, the following resources can be used.

ENISA introduced cyber security training materials in 2008, and has grown continuously ever since. The ENISA training material are focused on: technical, operational, setting up a CSIRT and legal and cooperation.

We have collected the group of 10 people and assessed them using cyber security exercises created using the attacks on Georgia and Ukraine. Each students had to answer 10 questions. We got the following results:

| Student ID | Score |
|------------|-------|
| 1 | 4/10 |
| 2 | 2/10 |
| 3 | 4/10 |
| 4 | 5/10 |
| 5 | 6/10 |
| 6 | 2/10 |
| 7 | 3/10 |
| 8 | 4/10 |
| 9 | 5/10 |
| 10 | 2/10 |

As we can see the average number of correct questions is 3.7.

We trained these students for 20 hours using ENISA materials and made them to answer the similar questions once more. We've got the following results:

| Student ID | Score |
|------------|-------|
| 1 | 8/10 |
| 2 | 6/10 |
| 3 | 7/10 |
| 4 | 7/10 |
| 5 | 6/10 |
| 6 | 8/10 |
| 7 | 5/10 |
| 8 | 7/10 |
| 9 | 6/10 |
| 10 | 6/10 |

As we can see the average number of correct questions is 6.6.

This experiment shows us the European training program give rather good results.

Based on our research we offer global and small business owners to use only proven solutions, to work according EU-wide cybersecurity legislation and to spend resources on the awareness raising in cyber security fields using European training materials.

REFERENCES:

1. <https://www.enisa.europa.eu/publications/corporate-documents/enisa-programming-document-2019-2021>
2. Kim, BH., Kim, KC., Hong, SE. et al. Multimed Tools Appl (2017) 76: 6051.
<https://doi.org/10.1007/s11042-016-3495-y>

MATHEMATICAL MODEL OF COUNTER-TERRORIST ACTIVITY

Volodymyr Khoroshko, National Aviation University, Doctor of Engineering Science, Full Professor, Kiev, Ukraine
Mykola Brailovskyi, Taras Shevchenko National University of Kyiv, PhD in Engineering Science, Associate Professor
Kiev, Ukraine

Valeri Kozura, National Aviation University, PhD in Engineering Science, Associate Professor Kiev, Ukraine

ABSTRACT. In this paper, we construct a model of an arbitrary terrorist group with a strictly justified hierarchy using a weighted undirected graph. The proposed mathematical model makes it possible to use new methods for processing graphs that are not used earlier for solving the problem of destroying the modeled grouping, as well as a numerical estimate of the damage inflicted on the enemy through counterterrorist actions.

Аннотация В данной работе осуществляется построение модели произвольной террористической группы со строго обоснованной иерархией при помощи взвешенного неориентированного графа. Предлагаемая математическая модель дает возможность для использования новых, не применяемых раньше, методов обработки графов для решения задачи о разрушении моделируемой группировки, а также численной оценки ущерба, наносимого противнику посредством контртеррористических действий.

KEYWORDS: terrorist group, graph mathematical models, weights, matrix theory, counterterrorism activity.

Введение. Геополитическая ситуация и события, которые происходят в мире, особенно после трагических событий 2008 года в Грузии и в Украине начиная с 2014, чрезвычайно остро поднимают вопрос создания адекватных математических моделей различных террористических организаций с целью привлечения строгого математического аппарата и вычислительной техники для автоматизации разработки возможных путей борьбы с ними, а также для формализации анализа результатов контртеррористических действий.

Традиционным путем для представления группы людей с указанием взаимных отношений между ними является использование теории графов [1,2]. Это обусловлено рядом факторов, среди которых:

- наглядность получаемой модели,
- возможность адекватного отражения при помощи стандартных операций на графах реальных действий над группами и событиями, а также событий в группах,
- существованием разработанного математического аппарата для работы с графами, включая большое количество хорошо зарекомендовавших себя на практике эвристических методов обработки.

В настоящий момент в научном мире чрезвычайно активизировались работы по математическому моделированию террористических организаций [2-5]. Однако существующие модели, информация о которых доступна из определенных источников, далеки от совершенства. Так графовые представления террористической организации, представленные в [3,4], носят ограниченный и недостаточно информативный характер, поскольку не учитывают иерархию организации. Попытка такого учета была предпринята в [2] за счет введения в рассмотрение упорядоченного множества вершин графа, хотя автор не определяет строго на рассматриваемом множестве необходимое

бинарное отношение, обладающее свойствами рефлексивности, транзитивности и антисимметричности, без введения которого рассмотрение упорядоченности множества невозможно. Аналогичный результат очевидно был бы легко получен автором [2] при помощи перехода от неориентированного графа к ориентированному. Такой переход для повышения информативности графа за счет учета иерархии предлагается в [6]. Однако ориентация ребра между двумя вершинами-индивидуумами, основанная на учете лишь количества связей каждого из них (учет степеней соответствующих вершин графа) ставит под сомнение адекватность получаемой модели, так как возможна такая организация террористической группы, когда лидеры будут иметь минимальное количество непосредственных связей с подчиненными, оказываясь таким образом совсем не на верхних ступенях иерархической лестницы.

Таким образом, до настоящего момента не существует адекватной математической модели террористической группировки, полностью отражающей ее реальную иерархию и взаимосвязь между членами модели, позволяющей удовлетворительно формализовать решение традиционных в этой предметной области задач (о разрушении террористической организации, ограничение ее деятельности и т.д.).

Цель настоящей работы – построение модели произвольной террористической группы со строго обоснованной иерархией при помощи взвешенного неориентированного графа, что не делалось ранее. Введение значений веса для вершин и ребер происходящей при максимальном использовании априорной информации о моделируемой террористической группе (противника).

Предлагаемая математическая модель дает возможность для использования новых, не применяемых раньше, методов обработки графов для решения задачи о разрушении моделируемой группировки, а также численной оценки ущерба, наносимого противнику посредством контртеррористических действий.

Основная часть. Рассмотрим задачи, связанные с организацией контртеррористических действий, решение которых осуществляется с использованием графовых математических моделей противника. Отдельные индивидуумы представляются в такой модели в виде узлов (вершин), пары которых соединяются ребром (вершины при этом называются смежными) при существовании определенной взаимосвязи между соответствующими членами рассматриваемой группы.

Пусть террористическая организация в своей иерархии имеет три основных уровня: лидера (руководителя) или нескольких лидеров; представителей связующего звена (руководство на местах) и непосредственных исполнителей. При построении самой простой графовой модели (неориентированный невзвешенный граф) каждому члену организации противника соответствует вершина, ребра графа соединяемой вершины в том случае, если между соответствующими им членами существует непосредственная связь. Пример такой модели представлена на рис.1(а), где узлы, соответствующие лидерам организации, среднему звену и исполнителям, для наглядности имеют соответственно разные цвета (красный, синий и черный цвет). Граф очевидно является связным. Как правило, непосредственной связи между лидерами и исполнителями не существует, хотя такая возможность и не исключается. Традиционно графовые модели противника служат для решения следующих задач: определение членов террористической группы, (противника), блокирования (удаления) которых реально возможно осуществить при этом блокирование приведет к распаду организации противника на несколько несвязанных между собой частей. Результатом такого распада может оказаться как полное уничтожение группы, так и снижение ее боеспособности, эффективности деятельности.

На языке графов данная задача будет формироваться следующим образом: необходимо определить множество узлов (множество, содержащее минимальное количество узлов), удаление которых приведет к распаду связей графа на несколько компонентов. Если такое множество содержит один узел, то он называется точкой сочленения. В примере, приведенном на рис.1(а), точкой сочленения является S1. Блокирование этого единственного члена организации противника приводит

- 1) Его осведомленности об объекте, на который направлено внимание;
- 2) Материальных и временных возможностей для осуществления отведенной данному члену группировки роли в террористической операции;
- 3) Значимость рассматриваемого члена в организации.

Учет всех перечисленных выше составленных частей весовых коэффициентов автоматически выделит лидеров (вершины с наибольшими значениями весов) и остальных менее значимых членов группы.

Все ребра определяются в зависимости от:

а) реальной ценности информации, передаваемой при помощи данной линии связи, (например, информация, передаваемая от руководителей группы подчиненным, является более значимой, чем информация, циркулирующая между непосредственными исполнителями);

б) надежности рассматриваемой линии связи (например, связь при непосредственном контакте является более надежной, чем при использовании телефонной линии).

Пример взвешенного графа – модели приведен на рис.2 (порядок нумерации соответствует иерархии членов организации, в середине узла – его номер, рядом с узлом – его вес, рядом с ребром в скобках – вес ребра).

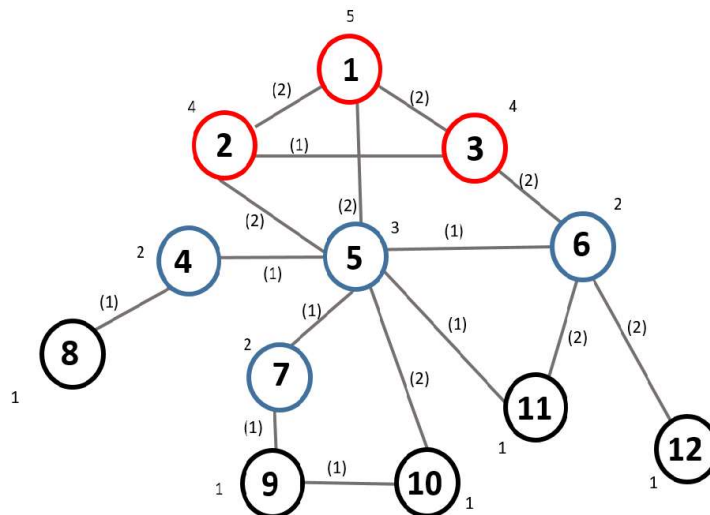


Рис.2 Модель террористической группы в виде взвешенного графа

Замечание 1. Использование взвешенного графа при моделировании организации противника дает возможность учесть ее иерархию, не переходя к ориентированному графу. Такой переход, как правило, осложняет процесс обработки графа. Кроме того, матрица смежности неориентированного графа обладает симметричностью, что дает возможность в некоторых графовых алгоритмах значительно сократить количество необходимых арифметических операций [7].

Задачи, связанные с организацией контртеррористических действий, были сформулированы в общем виде. Результат удаления некоторых членов террористической группировки или блокирования каких-то связей, приводящих с точки зрения снижения дееспособности противника. Например, если удалить связь между членами S4 и S5 (мост в графовой модели противника (рис. 1(a))), это вряд ли нанесет ощутимый удар по всей группировке, т.к. оставшись без D5 часть группировки сохранит как абсолютное большинство своих членов, так и наличие всех иерархических звеньев.

Одним из основных вопросов при моделировании террористических сетей и активных действий над ними является вопрос о том, когда рассматриваемую структуру можно считать разрушенной, или уничтоженной. Вариант уничтожения всех членов группы рассматриваться не будет, т.к. несмотря на то, что такой вариант часто является приемлемым и даже желаемым, он с

большой долей вероятности может оказаться принципиально невыполнимым (либо невыполнимым за определенный ограниченный промежуток времени при наличии определенного ограниченного материального ресурса).

Рассмотрим возможное решение для первой задачи, приводящее к уничтожению террористической группировки в соответствии с [2]. В [2] для решения этой задачи по графовой модели группы определяется множество всех простых, или «командных», цепей согласно [1], начало и конец которых отвечает лидеру и непосредственному исполнителю соответственно. По полученному множеству определяется совокупность узлов графа, каждый из которых присутствует хотя бы в одной цепи, причем каждая цепь вносит в эту совокупность единственный узел.

Удаление из графа такой совокупности, разрушает все существующие «командные» цепи. В этом случае в [2] делается вывод об уничтожении террористической группировки. Однако непосредственного алгоритма предполагаемой «разрушительной» операции не приводится. Более того, при учете иерархии моделируемой террористической организации, проводимом в [2], само выделение «командной» цепи становится проблематичным.

Рассмотрим возможный алгоритм для осуществления уничтожения террористической группы противника в смысле [2], используя в качестве модели предложенный выше взвешенный неориентированный граф. Для этого построим для граф-модели корневую структуру уровней (КСУ) [7] с корнем в узле, имеющем наибольший вес, т.е. отвечающем лидеру (вариант, когда значение максимального веса соответствует нескольким вершинам, рассматривается ниже). Для удобства дальнейшего изложения обозначим этот узел x , КСУ $F(x)$ есть разбиение множества вершин x графа:

$$F(x) = \{L_0(x), L_1(x), \dots, L_{l(x)}(x)\},$$

такое, что $L_0(x) = \{x\}$, $L_1(x) = Adj(L_0(x))$, $L_i(x) = Adj(L_{i-1}(x) - L_{i-2}(x))$, $i=2,3,\dots,l(x)$,

где $Adj(x)$ – множество узлов графа, не принадлежащих $L_{i-1}(x)$, на смежных хотя бы с одним узлом из $L_{i-1}(x)$. Эксцентриситет узла x [1] по отношению к структуре уровней называется длиной $F(x)$, а ширина $\omega(x)$ структуры $F(x)$ и определяется так: $\omega(x) = \max \{ |L_i(x)| \mid 0 \leq i \leq l(x) \}$.

Для графа, представленного на рис.2, корневая структура уровней, описанная выше, будет иметь вид, представленный на рис.3. Все «командные» цепи – это очевидно простые цепи графа, исходящие из нулевого уровня корневой структуры и заканчивающиеся либо вершиной, степень которой равна 1, либо вершиной, лежащей в последнем уровне КСУ; если V_k, V_m – две последовательные вершины такой цепи, то номер уровня в который попала вершина V_m . Узлы, попавшие в один уровень структуры, определяют ту совокупность, удаление которой приведет к распаду графа на смежные компоненты за счет разрыва всех цепей связи, т.е. к уничтожению террористической группы.

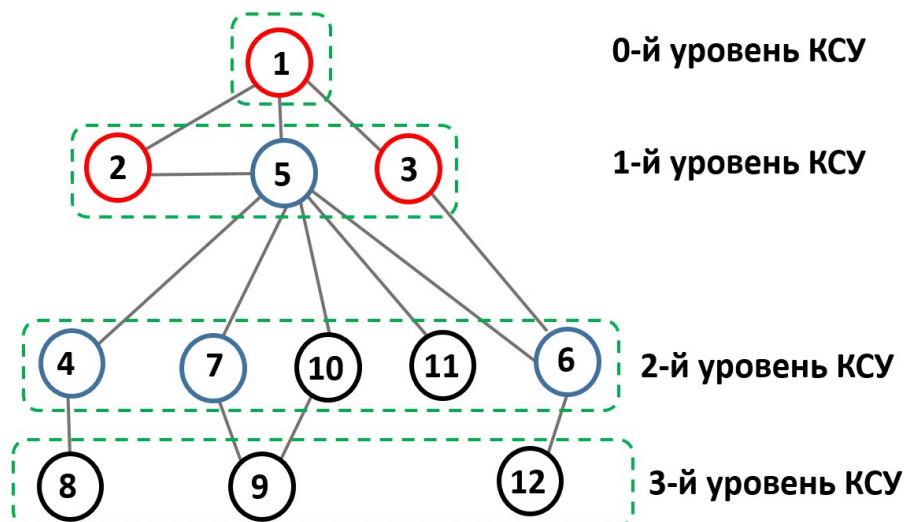


Рис.3 Корневая структура уровней

Заметим, что при этом не потребовалось явное определение множества «командных» цепей. Способ построения КСУ приведет к тому, что лидеры будут «отрезаны» от непосредственных исполнителей, что лишит возможности организованных активных действий данную террористическую группировку.

Пусть имеется несколько вершин с максимальным весом. Тогда при построении КСУ роль «корня» будет играть не один узел: все вершины графа с максимальными весовыми значениями, отвечающие лидерам противника, помещаются на нулевой уровень структуры. Остальные шаги для выделения разделяющего множества графа остаются без изменения.

Рассмотрим вариант нанесения удара по террористической структуре, когда во главу угла ставится уничтожение (блокирование) минимального (или просто малого) количества людей. Переходя к графовой интерпретации, задача заключается в поиске минимального разделяющего множества графа, (или распределяющего множества, содержащего малое количество вершин). Для того, очевидно, потребуется длинная и узкая корневая структура, в которой целесообразно исключить узлы из того уровня, где их количество минимально. Самая длинная КСУ отвечает корню, являющемуся периферийным узлом графа. Однако поиск периферийных узлов является дорогостоящим предприятием, требуя для своего осуществления, как правило, более, чем $\approx(|X| |E|)$ арифметических операций, где X – это количество вершин, а E – количество ребер графа, и для графа большой размерности может оказаться достаточно громоздким в вычислительном смысле [7]. Учитывая это, будем использовать КСУ с корнем в псевдопериферийном узле [7], алгоритм поиска которого представлен в следующем виде:

Шаг 1. (инициализация) Выбрать произвольный узел r .

Шаг 2. Построение структуры уровней) Построить структуру уровней с корнем в r :

$$F(r) = \{L_0(r), L_1(r), \dots, L_{l(r)}(r)\}.$$

Шаг 3. (стягивание последнего уровня). Выбрать в $L_{l(r)}(r)$ узел x с минимальной степенью.

Шаг 4. (построение структуры уровней). Построить $F(x) = \{L_0(x), L_1(x), \dots, L_{l(x)}(x)\}$. Если $l(x) > l(r)$, положить $r \leftarrow x$, перейти на шаг 3.

Шаг 5. Узел x – псевдопериферийный.

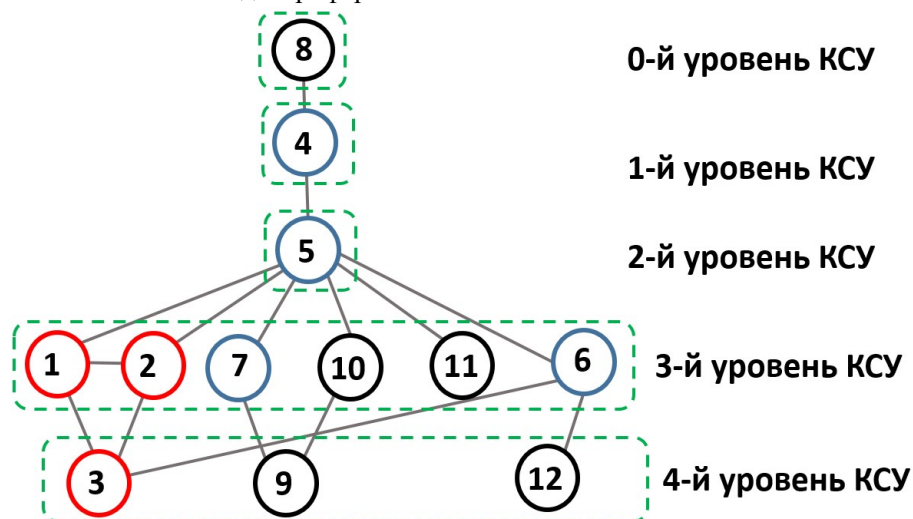


Рис.4 Структура уровней с корнем в псевдопериферийном узле

Исходя из предположенного алгоритма, используем корневую структуру на рис.3 как начальную для поиска псевдопериферийного узла. В последнем уровне выбором узел с минимальной степенью $x=12$. Проведя действия, предусмотренные алгоритмом, полученным в качестве псевдопериферийного узла, узел $x=8$, корневая структура с корнем в этом узле приведена на рис.4.

Шаг 1. Найти псевдопериферийный узел x в графе противника.

Шаг 2. В корневой структуре уровней $F(x) = \{L_0(x), L_1(x), \dots, L_{l(x)}(x)\}$ найти уровень $L_i(x)$, содержащий наименьшее число графа.

а) Если $L_i(x)$ определяется однозначно, то положить $I = i$, перейти на шаг 3.

б) Пусть $L_i(x)$ определяется неоднозначно, т.е. $i \in \text{IND}$, $|\text{IND}| > 1$, где IND – некоторое множество индексов уровней КСУ $F(x)$

Вычислить $E_{ir} [L_i(x)]$ для $\forall i \in \text{IND}$,

Определить $E_{\min} = \min_{i \in \text{IND}} E_{ir} [L_i(x)]$,

$$I = \underset{i \in \text{IND}}{\operatorname{argmin}} E_{ir} [L_i(x)].$$

Шаг 3. $L_i(x)$ – искомым уровнем КСУ, исключенный из нее.

Замечание 2. В предложенном алгоритме используется сравнение энергий террористических групп после удаления их некоторых членов. Воспользовавшись симметричностью матрицы смежности неориентированного графа, для оценки $E_{ir} [L_i(x)]$ можно вычислить норму не всей матрицы смежности, а лишь ее верхней (или нижней) треугольной части, что позволит сократить вычислительную работу по сравнению с первоначальным вариантом практически в два ряда.

Замечание 3. Количество арифметических операций для реализации предложенного алгоритма для графа множеством вершим X определяется как $\approx (|X|^2)$.

Конечно, такой алгоритм не гарантирует отделение лидеров от непосредственных исполнителей, но разбиение на связанные компоненты в любом случае приведет к ослаблению террористической группировки и потребует определенного времени на ее восстановление.

Выводы

1. Показано, что использование при построении моделей различных террористических группировок взвешенных графов является перспективным направлением, дающим возможность учесть иерархию противника без перехода к ориентированному графу и введения в графовую модель различных дополнительных математических объектов.

2. Симметричность матрицы смежности полученной графой модели дает потенциальную возможность выигрыша в количестве арифметических операций при обработке получаемого графа по сравнению с тем вариантом, когда граф оказывается ориентированным.

3. Предложен новый подход к проблеме численной оценки ущерба, наносимого террористической группировке контртеррористическими действиями, основанный на использовании теории матриц и введение весовой энергии террористической группы, на основании которого построен новый алгоритм для определения разделяющего множества графа противника.

Конечно, многие задачи, решаемые на графах, для получения точного решения требуют полного перебора, однако наличие большого числа эвристических алгоритмов, хорошо зарекомендовавших себя на практике при обработке графов, не имеющих отношения к представлению террористических групп, дает возможность рассчитывать на успешное использование некоторых из этих алгоритмов и на графовых моделях.

ЛИТЕРАТУРА

1. Оре О. Теория графов. – М:Наука, 1980.-336с.
2. Майника Э. Алгоритмы оптимизации на сетях и графах. – М:Мир, 1981. – 323с.
3. Krebs V. E. Mapping networks of terrorist cells. – Connections 24(3).- 2001. – Pp. 43-52
4. Brailovskyi M., Khoroshko V. Models of Interaction of a Potentially Dangerous Terrorist Group and the Security Service on a Protected Object. SPCSJ, vol.2, w3, September 2018. – p.1-8.

5. Кобозева А. А., Хорошко В. А. Анализ информационной безопасности – К: изд. ГУНКТ, 2009. – 251с.
6. Brams S.J., Mutlu H., Ramirez S. L. Influence in Terrorist Networks: From Undirected to Directed Graphs. Studies I Conflict & Terrorism. – 2006. – 29. – Pp. 703-718.
7. Джордж А., Лю Дж. Численное решение больших разреженных систем уравнений. Изд. 2-е. – М: Мир, 2014. – 333с.

Методы защиты от современных векторов кибер атак в странах Европы Security methods against modern cyber attack vectors in countries of Europe

Георгий Иашвили¹, Максим Явич¹, Сергей Гнатюк² Андрей Фесенко³

1. Кавказский Университет, Тбилиси, Грузия

2. Национальный Авиационный Университет, Киев, Украина

3. Киевский национальный университет им. Тараса Шевченко, Киев, Украина

АННОТАЦИЯ. Для получения данных организаций, сервисов и пользователей кибер преступники прибегают к изощренным и продвинутым методам атак. К концу 2017 года европейским центром по кибер преступности – ЕСС была выработана стратегия по борьбе с мошенничеством и преступностью в сети. В данной статье мы проанализировали тенденцию улучшения безопасности пользователя на примере эксперимента с фишинг тренажерами. Знание, и массивное представление распространённых видов кибер атак может послужить одним из ключевых моментов в борьбе с преступностью в интернете, столь популярной на территории Европы.

ABSTRACT. Cyber criminals use sophisticated and advanced attack methods in order to obtain the data of organizations, services and users. By the end of 2017, the European Center for Cyber Crime - ECC has developed a strategy to deal with fraud and crime in the network. In this article, we have analyzed the tendency of improving user's security using the example of an experiment with phishing simulators. Knowledge, and a massive representation of common types of cyber attacks can be one of the key moments in the fight against crime on the Internet, that is rather popular in Europe.

КЛЮЧЕВЫЕ СЛОВА: кибер атаки, атаки на Европу, кибер преступность

KEYWORDS: cyber attacks, attacks on Europe, cyber crime

Для получения данных организаций, сервисов и пользователей кибер преступники прибегают к более изощренным и продвинутым методам атак. Так в 2017 году было скомпрометировано более 2 миллиардов записей, а к началу 2018 года их количество возросло до 4.5 миллиардов. По данным мирового финансового форума одними из самых распространённых методов кибер атак в странах Европы являются: продвинутые наборы фишинга; атаки удаленного доступа; атаки с помощью смартфонов; использование уязвимостей умных домов и интернета вещей; использование искусственного интеллекта; вымогатели – ransomware [2].

Продвинутые наборы фишинга

На сегодняшний день фишинг остается самым распространённым и действенным методом атак. Несмотря на сравнительно небольшой цикл жизни фишинг сайтов (в лучшем случае 2-3 дня), охват целевой аудитории достаточно большой. Для создания фишинг схемы не требуется особых технических знаний, и даже средний пользователь сети сможет выстроить механизм атаки. С

выходом более мощных и продвинутых инструментов фишинг становится все более опасным орудием в руках злоумышленников. На просторах dark web сегодня можно встретить огромное количество наборов фишинг атак.

Атаки удаленного доступа

Количество удаленных атак растет с каждым годом, и данные атаки становятся все более изощренными. Одним из основных типов атак удаленного доступа в 2018 году был cryptojacking, нацеленный на удаленный майнинг с использованием машины жертвы. Еще одной популярной целью атак удаленного доступа является устройства с открытыми портами, которыми и пользуются злоумышленники. К данному виду устройств относятся IP камеры, сетевые устройства, и другая периферия.

Атаки с помощью смартфонов

Наиболее распространенные векторы атак на смартфоны связаны с небезопасным использованием интернет ресурсов. Владельцы портативных устройств подвержены фишингу и атакам с помощью вредоносного программного обеспечения. По данным RSA, более 60% мошенничества в интернете совершается с помощью мобильных платформ, в то время, как 80% противозаконных действий с помощью мобильных устройств достигается с помощью мобильных приложений вместо мобильных веб-браузеров [3]. Большое количество пользователей производит финансовые операции при помощи мобильных устройств в незащищенных сетях. В результате чего их данные утекают в сеть и зачастую размещаются на специальных торговых площадках в dark web. Наряду с этим, за последние годы все чаще фиксируются DDoS атаки, произведенные с помощью смартфонов.

Использование уязвимостей умных домов и интернета вещей

Согласно прогнозам экспертов из Gartner, к 2020 году количество устройств интернета вещей достигнет отметки в 7 миллиардов единиц. Подавляющее количество пользователей IoT не видят в них потенциальной угрозы в силу того, что у многих устройств попросту нет пользовательского интерфейса. Данный факт может привести к проблемам с пониманием того, какие данные устройство собирает и какими управляет.

Использование программ / скриптов вымогателей – ransomware

Раньше объектами программного обеспечения вымогателей были компьютеры. За последние годы данный тип угрозы претерпел изменения и на сегодняшний день используется для атак на всевозможные устройства. Суть ransomware заключается в шифрации файлов жертвы на компьютере либо другом умном устройстве с целью получения денежных средств в обмен на ключи для расшифровки файлов.

Согласно проведенному ENISA – ом анализу, в 2018 году происхождение и методы примерно 21% проведенных атак так и не были установлены. В Threat Landscape Report – е сообщается, что пользователи и организации из стран Европы подвергались большому количеству до тех пор неизвестным атакам со стороны злоумышленников, что подтверждает приведенная статистика (рис. 1). В список так же вошли атаки типа DDoS и Brute force, но уже с явным отставанием, так как большинство современных механизмов защиты с легкостью отражают подобные атаки [4].

Векторы атак на страны Европы в 2018 году

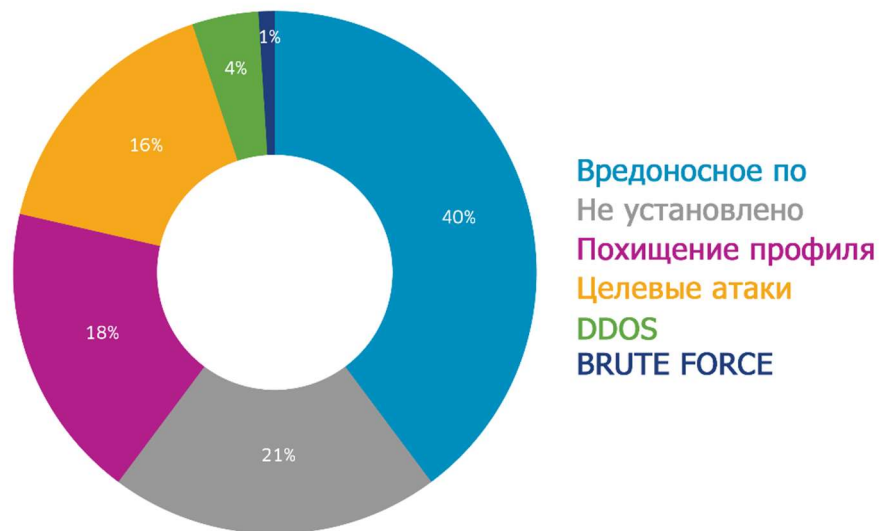


Рис.1

Такое количество неизвестных атак обусловлены очень частым выходом новых векторов атак, а также всевозможными вредоносными программами / скриптами, которые пока не изучены, следовательно, защитные механизмы против них и сопутствующая информация выходят с определенной задержкой.

Следующие по списку похищения профилей и целевые атаки, напрямую связанные с разнообразными веб ресурсами, будь то социальная сеть, либо информационный сайт. Довольно прибыльный вид атак, так как получив нужные данные, злоумышленники продают их на специальных ресурсах.

Но на первом месте несмотря на известные векторы атак остается вредоносное программное обеспечение. Злоумышленники под разными предлогами подкладывают жертвам всевозможные скрипты и программы, с помощью которых в результате производятся атаки разных категорий.

Методы борьбы с кибер преступностью

К концу 2017 года европейским центром по кибер преступности – ЕСС была выработана стратегия по борьбе с мошенничеством и преступностью в сети. Каждый квартал агентство публикует данные и тенденции развития в направлении кибер безопасности. Система ЕСС основана на трех основных компонентах: кибер стратегия; судебная экспертиза; спец операции.

Борьба с кибер преступностью напоминает вечную погоню. Только получается выявить новый вектор атак и организовать соответствующий защитный механизм против него, как на рынке появляется что-то совершенно новое, поражающее десятки, а то и сотни тысяч пользователей. Особо актуальной целью хакеров являются страны Европы. Из-за большого масштаба и развития технологий, объектов для атаки с каждым годом становится все больше.

Международными организациями предпринимаются определенные меры для предотвращения глобальных кибер атак. Для более четкого контроля данных и уменьшения утечек данных пользователей из стран Европы, в мае 2018 года был установлен общий регламент по защите данных GDPR — General Data Protection Regulation. Согласно GDPR посетителя веб сайта необходимо в максимально понятной форме оповещать о любой попытке получения и обработки его персональных данных этим ресурсом.

Основным и более действенным методом борьбы с ориентированными на пользователя кибер атаками является повышение уровня знаний в отрасли кибер безопасности у сотрудников организаций. Зачастую работники даже не подозревают о существовании того или иного вида атак, что служит причиной утечек ценных данных.

Тренировка сотрудников должна стать неотъемлемой частью механизма защиты организации от кибер атак. По заявлению IT Governace в 2017 – 2018 годах процент успешных атак на пользователей – сотрудников компаний существенно возрос по сравнению с предыдущими годами. В марте нынешнего года knowbe4 опубликовали статистику результатов годового фишинг тренинга сотрудников различных крупных и малых организаций в странах Европы. Параллельно проводился анализ успешно прошедших атак на этих пользователей (рис. 2).

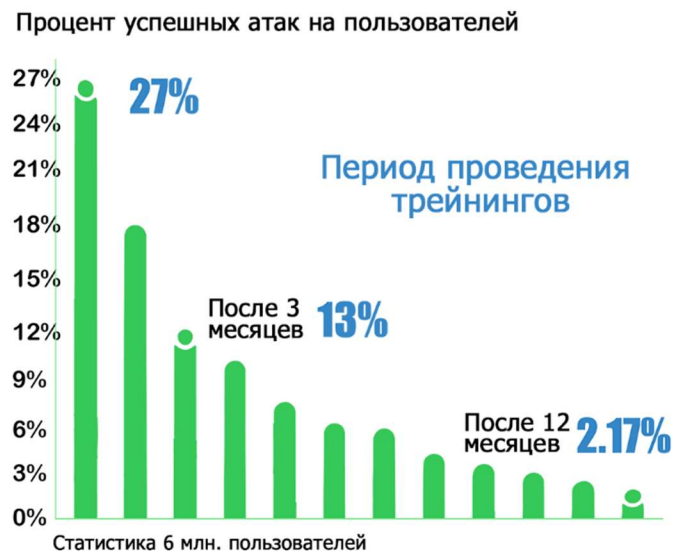


Рис. 2

Для анализа и проведения атак использовались различные фишинг симуляторы, с помощью которых отсылались письма с вредоносными ссылками, контентом и прикрепленными файлами. По результатам эксперимента, за один год тренинга процент попадания на фишинг уловки сократился с 27% до 2.17%, что безусловно положительно сказалось на общем развитии пользователей, и количестве утечек в сеть конфиденциальной информации.

Мы считаем, что тренинги с участием работников крупных, средних, а также малых организации по основным направлениям кибер безопасности и разновидностям кибер атак положительно скажутся на общей картине кибер безопасности в странах Европы. В данной статье мы разобрали тенденцию улучшения безопасности пользователя на примере эксперимента с фишинг тренажерами. Подобные тренинги необходимо проводить регулярно, охватывая все больше тематик и современных векторов атак. Лучшее оружие для пользователя интернета - это знание, и массивное представление распространённых видов кибер атак может послужить одним из ключевых моментов в борьбе с преступностью в интернете, столь популярной на территории Европы.

Использованная литература:

- [1] Center for strategic and international studies - Significant Cyber Incidents, May 2019
<https://www.csis.org/programs/technology-policy-program/significant-cyber-incidents>
- [2] Einaras von Gravrock - Here are the biggest cybercrime trends of 2019, March 2019 -
<https://www.weforum.org/agenda/2019/03/here-are-the-biggest-cybercrime-trends-of-2019/>
- [3] RSA - CURRENT STATE OF CYBERCRIME, 2018
<https://www.rsa.com/content/dam/premium/en/white-paper/2018-current-state-of-cybercrime.pdf>
- [4] Radware - Mobile Security Threats on The Rise as Hackers Can Launch DDoS Attacks on Their Mobile Phones, 2016 <https://security.radware.com/ddos-threats-attacks/cyber-attacks-in-the-palm-of-your-hand/>

სტალინისეული კიბერ დანაშაული

Stalin's Cyber Crime

Natalia Patarkatsasvhili , Ani Dekanosidze
Ivane Javakhishvili Tbilisi State University, Journalism Sophomore

ABSTRACT. Cyber crime as the term is invented in 21st century, but world history introduces us a lot of cyber criminals, among them the biggest is Stalin, man who made biggest crime without internet and computers in the 20th century. So, we think making parallels between two different centuries is very new and interesting. Based on the facts, we discuss Stalin's „cyber” crime's whole scheme.

ანოტაცია: კიბერ დანაშაული ოცდამეერთე საუკუნის მოვლენაა, თუმცა ისტორია მრავალ კიბერ დანაშაულს გვაცნობს, რომელთა შორისაც ყველაზე დიდი სტალინია, ადამიანი, რომელმაც ინტერნეტისა და კომპიუტერის გარეშე შეძლო მე-20 საუკუნის ყველაზე დიდი კიბერ დანაშაულის ჩადენა. ამდენად ვფიქრობთ, რომ ორ სხვადასხვა საუკუნეს შორის პარალელის გაკლება ძალიან საინტერესო და ინოვაციურია. ჩვენ ფაქტებზე დაყრდნობით განვიხილავთ სტალინის „კიბერ” დანაშაულის სრულ სქემას.

საკვანძო სიტყვები: კიბერ დანაშაული, სტალინი, კიბერ ქურდობა, ვირუსული პროგრამები, იდენტიფიკაციის ქურდობა, ინტერნეტი.

სტალინისეული „კიბერ” დანაშაული

საქართველოს სისხლის სამართლის კოდექსის თანახმად კიბერ დანაშაულად მიჩნეულია ნებისმიერი მართლსაწინააღმდეგო ქმედება, რომელიც კომპიუტერული სისტემის გამოყენებით არის ჩადენილი. ერთი შეხედვით ტერმინი „კიბერ დანაშაული” მხოლოდ ოცდამეერთე საუკუნის პრობლემად და „მონაპოვრად” გვევლინება, თუმცა ისტორია მრავალ კიბერ დანაშაულს გვაცნობს, რომელთაგან უდიდესი დიდი ბელადი, საბჭოთა კავშირის იმპერიის მმართველი, პიარმექანიზმის დიდი ოსტატი- იოსებ ჯუღაშვილი, სტალინი გახლდათ. კიბერ დანაშაულებათა ციკლი კომპიუტერის, ინტერნეტის და ყოველგვარი ტექნოლოგიის გარეშე, სადაც ყველა დღევანდელ ტერმინს თავისი სტალინური ანალოგი ჰქონდა- ესაა ხანა, როდესაც ადამიანთა მორალური, ფიზიკური, ეკონომიკური და ფსიქოლოგიური გაბანკოტება და ანულირება ჩვეულებრივი სისტემური მოქმედება იყო. გაზეთები, პოსტერები, წიგნები, სიმღერები, ლექსები, პოემები- თითოეული მათგანი ინდივიდის ბრბოდ ქცევსკენ იყო

მიმართული, ჩვენ ამას მე-20 საუკუნის კიბერ ქურდობა დავარქვით- ერთი ადამიანის მიერ მისაკუთრებული 15 რესპუბლიკის მოსახლეობა, მოპარული და მითვისებული სიცოცხლეები.

სწორედ სტალინი გვევლინება ოცდამეერთე საუკუნის ტერმინის „ვირუსული პროგრამების“ ავტორად, დღეს თუ ადამიანის პირადი ინფორმაციის მოსაპოვებლად ინტერნეტი და კომპიუტერია საჭირო, სტალინი ამას პარტიის მონების საშუალებით აკეთებდა. მას არა მხოლოდ წვდომა ჰქონდა თითოეული ადამიანის პირად ცხოვრებასა თუ სიცოცხლეზე, არამედ ხშირად თავადაც ქმნიდა მათ. სუფთა ქალაქებში აწერდა ხელს ისე, რომ წარმოდგენაც არ გქონდა რას დაწერდნენ ფურცლის ზედა ნაწილში, ერთ დღეს ექიმი, შესაძლოა მეორე დღეს მოლაღატე და სამშობლოს გამყიდველი ყოფილიყავი.

მსოფლიოს უდიდესი დამნაშავე, საკუთარი თავის უკვდავსაყოფად ყველა ხერხსა და გზას იყენებდა. მე-20 საუკუნის სათვალთვალ კამერები- კაგებეს აგენტები შენს ყველა ნაბიჯს ადევნებდნენ თვალყურს. წითელი სახელმწიფოს ბელადმა იცოდა თუ რას ფიქრობდი ძილში, უფრო მეტიც, შენს ცხოვრებას, შენს აზრებსაც კი აკონტროლებდა.

კიბერ დანაშაული ყოველდღე უფრო დიდ მასშტაბებს იძენდა და ნელ-ნელა დიდ „წმენდას“ უახლოვდებოდა. ყველა ის ადამიანი, რომელთანაც სტალინი შეაღწია, გზაჯვარედინზე იდგა, მის ტვინში შენახული ყველა ინფორმაცია მის განაჩენზე მიუთითებდა, ვირუსის- ანუ ანტისაბჭოური აზრის შემთვევაში, ისე ქრებოდი როგორც პატარა ფაილი მილიონში. ამ პატარა ფაილების რაოდენობამ კი რამდენიმე მილიონს გადააჭარბა, მილიონობით ადამიანი სტალინის კიბერ დანაშაულის მსხვერპლი გახდა. გაქრობის ტალღამ მწვერვალს 1937 წელს მიაღწია. რა ხდებოდა ჯალათების რეზიდენციაში და რა საშუალებებით ტყვინდნენ „მოსიარულე ვირუსებს“? როგორ სრულდებოდა დასახვრეტა სიის შევსება და როგორ აღასრულებდნენ რამდენიმე წუთში გამოტანილ განაჩენს? ბოლშევიკური ტერორის გეგმა საკმაოდ მარტივი იყო- უნდა დასჯილიყო უმკაცრესი ფორმით ყველა, ვისაც გააჩნდა განსხვავებული აზრი, ვისი შეხედულებაც წინააღმდეგობაში მოდიოდა პარტიის იდეოლოგიასთან, განადგურებულიყო ყველა, ფარული თუ აშკარა მოწინააღმდეგე დიდი იმპერიისა, რომელსაც შეეძლო სხვაც სასიკვდილოდ დაევირუსებინა. ხელისუფლების მიზანი ცხადი იყო- სრულად გაეკონტროლებინა ყველა და მოეკლა თავისუფლად მოაზროვნე ადამიანები-დღევანდელი სიტყვებით, ეს იყო მე-20 საუკუნის იდენტიფიკაციის ქურდობა. ამ მიზანს კი რეპრესიების საშუალებით- ანუ ანტივირუსებით ახორციელებდნენ. ამ მძიმე, შემზარავი და მეორეს მხრივ ნათელის მომფენი ინფორმაციის მიღების საშუალებას შს-ის არქივში დაცული მასალები გვაძლევს, 1954 წლის „თვითმხილველთა დაკითხვის ოქმები“, რომლებშიც უშუალოდ რეჟიმის მოხელეები ყვებიან, სისასტიკეზე რომელიც ციხის ბნელ საკნებში მუდმივად მძინვარებდა და უამრავი უდანაშაულოს სიცოცხლის მოსპობის მიზნით გახდა. ისინი სტალინისეული კიბერ გადაკიდების მსხვერპლნი გახდნენ, ოღონდ ინტერენტის გამოყენების გარეშე, ეფექტურად და სწრაფად. აქტებში ვკითხულობთ სიკვდილმისჯილთა სიტყვებს „გაუმარჯოს ამხანაგ სტალინს“, რა ირონიაა და ამავე დროს რამხელა დასტური დიდი ბელადის მიზნის აღსრულებისა, შენი სიკვდილის ხელის მომწერი შენი უკანასკნელი სიტყვის

ადრესატი ხდება, ისიც ხოტბის შესხმის კონტექსტში. სწორედ აქ, ამ მონაკვეთში, კარგად ვხედავთ სტალინის როგორც სისტემურად მოქმედი დამნაშავეს, ინტერენტისა და კომპიუტერული ტექნოლოგიების გარეშე, „კიბერ საბოტაჟის“ ავტორის მიღწეულ შედეგებს.

რეჟიმის აღზრდილი მოხელეები იძულებით თუ ნებით რეჟიმისავე მოტრფილები რომ იყვნენ ამას უამრავი ფაქტი ადასტურებს, ამიტომ ძნელად თუ ვნახავთ ფურცელს სადაც გამონჭრთნილ ჩინოვნიკებს ეჭვი შეეჭონდეთ რეჟიმისეულ სამართალში- სტალინმა ყველა კვალი გააქრო და მისი ფაილი სრულიად გაასუფთავა ზედ კი მისი სახელი - დიდი ბელადი რეალურად სისხლიანი, თუმცა იმპერიის მცხოვრებთათვის ოქროს ასოებით დააწერა.

როგორც წესი, ეპოქის დახასიათებისას აუცილებელი და გადამწყვეტია კონტექსტის გააზრება და შემდეგ კრიტიკა თუ ანალიზი. ჩვენ თანამედროვე მსოფლიოს მონაპოვარი - კომპიუტერული ტექნოლოგიების საშუალებით მიღწევადი ქმედება, ამ შემთხვევაში კი დანაშაული, მე-20 საუკუნეს დაუკავშირეთ. ჩვენი, და არა მხოლოდ ჩვენი, ქვეყნისთვის უმძიმეს პერიოდს- საბჭოთა კავშირს. რეალურად ყველა ის ტერმინი, რომელიც დღეს კიბერ დანაშაულს უკავშირდება, შეცვლილი სახით სტალინის პერიოდში არსებობდა, მართალია ინტერენტისა და ტექნიკის გარეშე, თუმცა ყველაფერი ახალი, ხომ კარგად დავინყებულები ძველია, ხოდა სტალინმა იდეა მოგვცა, რომლის სრულყოფა, განვითარება და ტექნოლოგიის საშუალებით განხროცილება უკვე ჩვენ, კიბერ სივრცის, ოცდამეერთე საუკუნის „მონაპოვრებმა“ შევძელით.

მნიშვნელოვანია წარსულის ცოდნა აწმყოში რეალობისა და მომავლის აღსაქმნელ-განსაჭვრეტად, შესაძლოა ეს კავშირი აფსურდად ჩანდეს, რეალურად კი იმაზე უფრო ღრმა და გრძელი ჯაჭვი არსებობს, ვიდრე ერთი შეხედვით ჩანს. მე-20 საუკუნის ყველაზე დიდი „კიბერ“ დამნაშავე „ჩვენი“ დიდი ბელადია, სწორედ ერთ-ერთი უდიდესი წინაპარი და ფუძემდებელი დღევანდელი კიბერ დამნაშავეებისა თუ თავად კიბერ დანაშაულისა.

ბიბლიოგრაფია

1. შსს-ის არქივში დაცული დოკუმენტური მასალა
2. https://police.ge/files/proeqtebi_reporma%20photos/organizebuli-danashauli/kiberdanashauli-informacia-moqalaqeebistvis.pdf

მაღალი რიგის პრიმიტიული მატრიცების გენერაცია სხვადასხვა სიმძლავრის აბელის მულტიპლიკაციური ჯგუფების ელემენტებით

Generation of high order primitive matrix elements with elements of abelian multiplicative groups

რ. მეგრელიშვილი¹, მ. ჯინჯიაძე²

¹თბილისის ივ. ჯავახიშვილის სახ. სახელმწიფო უნივერსიტეტი, თბილისი, საქართველო;

ელ-ფოსტა: richard.megrelishvili@tsu.ge

²აკ. წერეთლის სახელმწიფო უნივერსიტეტი, ქუთაისი, საქართველო

ელ-ფოსტა: mjinji@yahoo.com

ანოტაცია - ნაშრომში განხილულია ორიგინალური მატრიცული ცალმხრივი ფუნქცია და მისი შესაბამისი მაღალი რიგის მატრიცული მულტიპლიკაციური სასრული კომუტაციური ჯგუფის გენერაციის განზოგადებული მეთოდი. განხორციელებულია ველის პრიმიტიული ელემენტების აგების ჩასმა-გაფართოების მეთოდის ზოგადი ხერხი სხვადასხვა სიმძლავრის მქონე მატრიცული ჯგუფების ელემენტებით.

ABSTRACT. In this paper is considered the original one-way function matrix and it's corresponding generation method of high level matrix multiplicative finite commutative group. The general method of the insertion-enlarging method of building the primitive elements of the field is derived with elements of the matrix groups with different power.

საკვანძო სიტყვები: მატრიცული ცალმხრივი ფუნქცია, აბელის სასრული ველი, ასიმეტრიული კრიპტოგრაფია, მაღალი რიგის მატრიცული სასრული ველი, პრიმიტიული მატრიცული ელემენტი.

KEYWORDS: matrix one-way function; Abel finite field; asymmetric cryptography; high level matrix multiplicative finite field; primitive matrix element;

ცალმხრივი მატრიცული ფუნქცია

კრიპტოგრაფიული გასაღების გაცვლის დიფი-ჰელმანის ცნობილი მეთოდის ერთ-ერთ მოდიფიკაციას წარმოადგენს გასაღების გაცვლის მატრიცული ალგორითმები, რომელთა

მუშაობის საფუძველს ქმნის მაღალი რიგის ციკლური მულტიპლიკაციური მატრიცული ჯგუფები $GF(2)$ ველზე.

დავუშვათ, P მატრიცა წარმოადგენს ციკლური მატრიცული ჯგუფის პრიმიტიულ ელემენტს. ხოლო $\langle P \rangle$ ამ მატრიცის მიერ წარმოქმნილი მულტიპლიკაციური ჯგუფია, სიმძლავრით $2^n - 1$, სადაც n წარმოადგენს P კვადრატული მატრიცის ზომას.

საერთო გასაღების შემუშავების მატრიცული ალგორითმი შემდეგი სახისაა:

- გამგზავნი მხარე მიმღებ მხარეს ღია არხით უგზავნის $u_1 = vP_1$ ვექტორს, სადაც $P_1 \in \langle P \rangle$ გამგზავნის მიერ შერჩეული საიდუმლო მატრიცაა, ხოლო $v \in V_n$ ვექტორი საყოველთაოდ ცნობილია (V_n - ვექტორული სივრცეა $GF(2)$ ველზე);

- მიმღები მხარე თავის მხრივ ირჩევს $P_2 \in \langle P \rangle$ საიდუმლო მატრიცას და გამგზავნ მხარეს უგზავნის $u_2 = vP_2$ ვექტორს;

- გამგზავნი გამოთვლის $k_1 = u_2P_1$ ვექტორს;

- მიმღები გამოთვლის $k_2 = u_1P_2$, სადაც k_1 და k_2 - საიდუმლო გასაღებებია;

ცხადია, $k_1 = k_2 = k$, რადგანაც $k = vP_1P_2 = vP_2P_1$, $\langle P \rangle$ ჯგუფის კომუტაციურობის გამო. ვთქვათ, $v = (v_1, v_2, v_3, \dots, v_n) \in V_n$ და $u = (u_1, u_2, u_3, \dots, u_n) \in V_n$ არასაიდუმლო ვექტორებია ზემოთმოყვანილი ალგორითმიდან, ხოლო

$$P_1 = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in \langle P \rangle$$

საიდუმლო მატრიცაა. მაშინ, ალგორითმის თანახმად

$$vP_1 = \begin{pmatrix} v_1a_{11} + v_2a_{21} + \dots + v_na_{n1} \\ v_1a_{12} + v_2a_{22} + \dots + v_na_{n2} \\ \vdots \\ v_1a_{n1} + v_2a_{n2} + \dots + v_na_{nn} \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \quad (1)$$

მიღებულ წრფივ განტოლებათა სისტემაში უცნობების რაოდენობა განტოლებების რაოდენობის კვადრატია. ცხადია, სისტემის ამოხსნა რეალურ დროში შეუძლებელია, თუკი მატრიცის ზომა საკმარისად დიდია. ეს იმდენად მნიშვნელოვანი გარემოებაა, რომ თავისთავად აუცილებელს ხდის მაღალი სიმძლავრის მქონე, ახელის მულტიპლიკაციური მატრიცული ჯგუფის გენერაციას, რომლის პრიმიტიული ელემენტი მაღალი რიგის კვადრატული მატრიცა იქნება.

სასრული მატრიცული ჯგუფები

განვიხილოთ $(1 + \alpha)^j$, სადაც $j = 0, 1, 2, \dots$, ხოლო α წარმოადგენს პრიმიტიული პოლინომის ფესვს $GF(2^n)$ ველში მოდულით $p(x)$.

$$\begin{aligned}
 (1 + \alpha)^0 &= 1 && 1 \\
 (1 + \alpha)^1 &= 1 + \alpha && 11 \\
 (1 + \alpha)^2 &= 1 + \alpha^2 && 101 \\
 (1 + \alpha)^3 &= 1 + \alpha + \alpha^2 + \alpha^3 && 1111 \\
 (1 + \alpha)^4 &= 1 + \alpha^4 && 10001 \\
 (1 + \alpha)^5 &= 1 + \alpha + \alpha^4 + \alpha^5 && 110011
 \end{aligned} \tag{1}$$

მიღებული პოლინომების კოეფიციენტები ქმნის სტრუქტურას, რომელიც სერპინსკის სამკუთხედის სახელითაა ცნობილი.

სერპინსკის სტრუქტურა შეიცავს მრავალ ქვესტრუქტურას, რომლებიც მულტიპლიკაციური ჯგუფების გენერატორად (მანარმოებელ მატრიცად) გამოდგება, ანუ პრიმიტიულ ელემენტებს წარმოადგენენ. ასეთია, მაგალითად,

$$P_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \quad P_5 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}, \quad P_7 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \tag{2}$$

და სხვა მრავალი. მათი ნატურალური ხარისხები ქმნის აბელის მულტიპლიკაციურ ციკლურ ჯგუფებს.

ადვილად შეიძლება დავრწმუნდეთ, რომ P_3, P_5, P_7 მატრიცების ახარისხებით მიღებული

$$P_3^k, P_5^k, P_7^k, k = 1, 2, \dots, 2^k - 1 \tag{3}$$

სიმრავლეები აბელის მულტიპლიკაციურ ციკლურ ჯგუფებს წარმოადგენენ.

მაგ.:

$$\begin{aligned}
 P_3^1 &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, P_3^2 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, P_3^3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, P_3^4 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \\
 P_3^5 &= \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, P_3^6 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, P_3^7 = P_3^0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
 \end{aligned} \tag{4}$$

ჩვენს მიერ P_3 მატრიცის, როგორც საბაზო სტრუქტურის, ჩასმა-გაფართოების მეთოდით მიღებული იყო მეორე რიგის გაფართოება ****:

$$P_{3^2}(i, j) = \begin{pmatrix} P_3^i & P_3^j & P_3^j \\ P_3^j & 0 & 0 \\ P_3^j & P_3^j & 0 \end{pmatrix}, \text{ სადა } i, j=0..6. \quad (5)$$

P_3 მატრიცას ეწოდება საბაზო სტრუქტურა. P_3^i და P_3^j მატრიცებს - შესაბამისად პირველი და მეორე მაფართოებელი მატრიცები.

$P_3^k, k = 1, 2, \dots, 2^3 - 1$ სიმრავლეს ეწოდება $F(P_{3^2}(i, i + 1))$ ჯგუფის წინარე ჯგუფი.

ჩვენს მიერ დამტკიცებული იყო შემდეგი წინადადების ჭეშმარიტება ****:

P_3 მატრიცის ნებისმიერი მეორე რიგის $(i, i + 1)$ გაფართოება - $P_{3^2}(i, i + 1), i = 0..5,$

წარმოადგენს პრიმიტიულ ელემენტს და წარმოქმნის აბელის მულტიპლიკაციურ სასრულ ჯგუფს $F(P_{3^2}(i, i + 1))$, რომლის სიმძლავრეა $2^{3^2} - 1$.

მაგალითად, პრიმიტიულია $P_{3^2}(0,1)$ მატრიცა, ხოლო $[P_{3^2}(0,1)]^{2^{2 \cdot 3^1} + 2^{3^1} + 1}$ მატრიცა დიაგონალურ მატრიცას წარმოადგენს:

$$[P_{3^2}(0,1)]^{2^{2 \cdot 3^1} + 2^{3^1} + 1} = \begin{pmatrix} P_3^3 & 0 & 0 \\ 0 & P_3^3 & 0 \\ 0 & 0 & P_3^3 \end{pmatrix} \quad (6)$$

ცხადია, დიაგონალური მატრიცის ნებისმიერი ხარისხი $([P_{3^2}(0,1)]^{2^{2 \cdot 3^1} + 2^{3^1} + 1})^i, i = 1, 2, \dots$ ისევ დიაგონალური მატრიცა იქნება, და რადგან $F(P_{3^2}(0,1))$ სიმრავლე სასრული ჯგუფია, ამიტომ, როცა $i = 2^{3^1} - 1$, ვღებულობთ

$$\left([P_{3^2}(0,1)]^{2^{2 \cdot 3^2} + 2^{3^2} + 1}\right)^i = \begin{pmatrix} (P_3^3)^i & 0 & 0 \\ 0 & (P_3^3)^i & 0 \\ 0 & 0 & (P_3^3)^i \end{pmatrix} = \begin{pmatrix} P_3^{3i \bmod i} & 0 & 0 \\ 0 & P_3^{3i \bmod i} & 0 \\ 0 & 0 & P_3^{3i \bmod i} \end{pmatrix} \quad (7)$$

რადგან მატრიცულ ოპერაციებს ვანარმოებთ წინარე ჯგუფის მოდულით, შესაბამისად, (7) ერთეულოვანი მატრიცაა. რაც ნიშნავს, რომ $F(P_{3^2}(0,1))$ სიმრავლე სასრული ჯგუფია.

შეგნიშნოთ, რომ საბაზო სტრუქტურულ შეიძლება ავიღოთ (4) სიმრავლის სხვა ნებისმიერ არაერთეულოვანი ელემენტი. არსებობს ამ ელემენტის ისეთი გაფართოება P_3^0 და P_3^1 მაფართოებელი მატრიცებით, რომელიც პრიმიტიული ელემენტია.

მაგალითად, პრიმიტიული იქნება შემდეგი გაფართოებები :

$$\begin{pmatrix} P_3^1 & P_3^1 & 0 \\ 0 & 0 & P_3^1 \\ P_3^0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} P_3^1 & 0 & P_3^0 \\ P_3^1 & P_3^1 & P_3^1 \\ 0 & P_3^1 & P_3^1 \end{pmatrix}, \begin{pmatrix} 0 & P_3^1 & 0 \\ 0 & P_3^1 & P_3^1 \\ P_3^0 & 0 & P_3^1 \end{pmatrix} \quad (8)$$

განვიხილოთ სერპინსკის სამკუთხედის უფრო მაღალი რიგის ქვესტრუქტურა :

$$P_5 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

აღვიღად მონშდება, რომ P_5 პრიმიტიული ელემენტია, რაც ნიშნავს, რომ P_5^k , $k = 1, 2, \dots, 2^5 - 1$ სასრულ მულტიპლიკაციურ კომუტაციურ ჯგუფს წარმოადგენს.

საბაზო სტრუქტურულ ავიღოთ P_3 მატრიცა და გაფართოვოთ იგი P_5^0 და P_5^1 მატრიცებით.

არსებობს P_3 მატრიცის ისეთი გაფართოებები P_5^0 და P_5^1 მატრიცებით, რომლებიც წარმოადგენს პრიმიტიულ ელემენტებს. მაგალითად,

$$\begin{pmatrix} P_5^1 & P_5^1 & P_5^1 \\ P_5^1 & 0 & 0 \\ P_5^0 & P_5^1 & 0 \end{pmatrix}, \begin{pmatrix} P_5^1 & P_5^1 & P_5^1 \\ P_5^1 & 0 & 0 \\ P_5^1 & P_5^0 & 0 \end{pmatrix} \quad (9)$$

მატრიცები პრიმიტიული ელემენტებია. შესაბამისად, $F(P_{3 \times 5^1}(P_5^0, P_5^1))$ სიმრავლე სასრული მულტიპლიკაციური კომუტაციური ჯგუფია.

მართლაც, ადვილი შესამონშებელია, რომ

$$[P_{3 \times 5^1}(P_5^0, P_5^1)]^{2^{2 \cdot 5^1} + 2^{5^1} + 1} = \begin{pmatrix} P_5^2 & 0 & 0 \\ 0 & P_5^2 & 0 \\ 0 & 0 & P_5^2 \end{pmatrix} \quad (10)$$

მატრიცა დიაგონალურია. (7) გარდაქმნის ანალოგიური მსჯელობით მივიღებთ, რომ

$$\left([P_{3 \times 5^1}(P_5^0, P_5^1)]^{2^{2 \cdot 5^1} + 2^{5^1} + 1} \right)^i, i = 1, 2, \dots \text{ მატრიცები დიაგონალურია, ხოლო როცა } i = 2^{5^1} - 1,$$

მივიღებთ

$$\begin{pmatrix} P_5^{2i \bmod i} & 0 & 0 \\ 0 & P_5^{2i \bmod i} & 0 \\ 0 & 0 & P_5^{2i \bmod i} \end{pmatrix} \quad (11)$$

ერთეულოვან მატრიცას. რაც ნიშნავს იმას, რომ, $F(P_{3 \times 5^1}(P_5^0, P_5^1))$ სიმრავლე სასრული მულტიპლიკაციური კომუტაციური ჯგუფია, სიმძლავრით $2^{3 \times 5^1} - 1$.

ახლა განვიხილოთ P_5 პრიმიტიული მატრიცის $k = 2$ რიგის გაფართოებები P_5^i , $i = 1, 2, \dots, 2^5 - 1$ კომუტაციური ჯგუფის ელემენტებით - $P_{5^k}(P_5^0, P_5^1)$, $k = 2$. არსებობს P_5 მატრიცის ისეთი გაფართოება, რომელიც პრიმიტიულ მატრიცას ქმნის. მაგალითად,

$$P_{5^k}(P_5^0, P_5^1) = \begin{pmatrix} P_5^1 & P_5^1 & P_5^1 & P_5^1 & P_5^0 \\ P_5^1 & 0 & 0 & 0 & 0 \\ P_5^1 & P_5^1 & 0 & 0 & 0 \\ P_5^1 & 0 & P_5^1 & 0 & 0 \\ P_5^1 & P_5^1 & P_5^1 & P_5^1 & 0 \end{pmatrix}, k = 2 \quad (12)$$

ჩვენს მიერ შემუშავებული პროგრამული პროდუქტის საშუალებით მონმდება, რომ $(P_{5^k}(P_5^0, P_5^1))^i$ მატრიცა, სადაც $i = 2^{4 \cdot 5^{k-1}} + 2^{3 \cdot 5^{k-1}} + 2^{2 \cdot 5^{k-1}} + 2^{5^{k-1}} + 1$, $k = 2$ დიაგონალურ მატრიცას წარმოადგენს:

$$(P_{5^k}(P_5^0, P_5^1))^{2^{4 \cdot 5^{k-1}} + 2^{3 \cdot 5^{k-1}} + 2^{2 \cdot 5^{k-1}} + 2^{5^{k-1}} + 1} = \begin{pmatrix} P_5^4 & 0 & 0 & 0 & 0 \\ 0 & P_5^4 & 0 & 0 & 0 \\ 0 & 0 & P_5^4 & 0 & 0 \\ 0 & 0 & 0 & P_5^4 & 0 \\ 0 & 0 & 0 & 0 & P_5^4 \end{pmatrix}$$

რომლის ნებისმიერი ნატურალური ხარისხი $(P_{5^k}(P_5^0, P_5^1))^{i \times j}$, ცხადია, დიაგონალური მატრიცაა, რომლის დიაგონალზეც განლაგებულია P_5^i , $i = 1, 2, \dots, 2^5 - 1$ წინარე ჯგუფის ელემენტები. ხოლო, როცა $j = 2^{3^{k-1}} - 1$, მივიღებთ

$$\begin{pmatrix} P_5^{4j \bmod j} & 0 & 0 & 0 & 0 \\ 0 & P_5^{4j \bmod j} & 0 & 0 & 0 \\ 0 & 0 & P_5^{4j \bmod j} & 0 & 0 \\ 0 & 0 & 0 & P_5^{4j \bmod j} & 0 \\ 0 & 0 & 0 & 0 & P_5^{4j \bmod j} \end{pmatrix} = \begin{pmatrix} P_5^0 & 0 & 0 & 0 & 0 \\ 0 & P_5^0 & 0 & 0 & 0 \\ 0 & 0 & P_5^0 & 0 & 0 \\ 0 & 0 & 0 & P_5^0 & 0 \\ 0 & 0 & 0 & 0 & P_5^0 \end{pmatrix}$$

ერთეულოვან მატრიცას.

მიღებული შედეგი იძლევა მაღალი რიგის მატრიცების მულტილიკაციური კომუტაციური სასრული ჯგუფების გენერირების პერსპექტივას.

გამოყენებული ლიტერატურა:

- [1] Mohammad Mehdi Nasrabadi, Ali Gholamian - On A-nilpotent abelian groups - Proceedings - Mathematical Sciences, Springer, November 2014, Volume 124, Issue 4, pp 517–525
- [2] Chiş C, Chiş M and Silberberg G, Abelian groups as autocommutator groups, Arch. Math. (Basel) 90(6) (2008) 490–492

ENSURING EUROPEAN CIVIL AVIATION CYBERSECURITY

Sergiy Gnatyuk¹, Maksim Iavich², Giorgi Iashvili², Andriy Fesenko³

¹National Aviation University, ²Caucasus University, ³Taras Shevchenko Kyiv National University

ABSTRACT. The criticality level of civil aviation (CA) information infrastructure is considerably amplified by high degree of connectivity and interaction between ground and aircraft systems. Malicious interference into mentioned systems puts at threats passengers, crew and ground staff security. Unauthorized access to so-called critical aviation information system (CAIS) is very crucial and it may have serious and tragic consequences. The control aviation security documents declare following requirements to ensure CAIS security against cyberthreats (potential cause of an unwanted incident, which may result in harm to a system, individual or organization – ISO / IEC 27032). Doc 30 declares that measures addressing cyberthreats to CA have been included in the National Civil Aviation Security Programme, the National Quality Control Programme and the National Civil Aviation Security Training Programme. Similar requirements are declared in Annex 17 to Chicago Convention on International Civil Aviation, Doc 8973 as well as in Doc 9985. However, there are still a lot of unsolved problems related to CAIS identifying, its criticality assessment and development of methods to provide its cybersecurity. From this viewpoint in the paper integrated complex approach to provide CA cybersecurity was proposed.

KEYWORDS: cybersecurity, critical information infrastructure, European civil aviation, complex information security system, critical aviation information system

Introduction

Cyberterrorism evolution (since the birth of computer technology in the 1960s) shows that attacks in cyberspace [1] today have a strong political overtones and more evident in cybernetic influence on international level. Only the first half of 2014 may be noted such cyberincidents: hackers broke into the Schengen Information System; powerful DDoS attack focused on 3 biggest NATO sites (<http://ccdcoe.org/>, <http://nato.int/> & <http://nato-pa.int/>) from pseudo-Ukrainian groups named «CyberBerkut» & «Anonymous Ukraine»; cyberattacks on Ukrainian Cabinet of Ministers, Prosecutor General's Office of Ukraine & National Security and Defense Council of Ukraine; 1.3 millions of big communication operator Orange France were victims of cyberattack focused on their personnel data; Russia launched large-scale cyberwar against Ukraine, it is bound to the revolutionary events and carries political overtones to destabilize the situation in the state and the violation of its sovereignty & integrity; powerful DDoS attacks from Russian Federation territory on Ukrainian Central Election Commission in preparation and voting in Presidential election.

Cyberattacks and acts of cyberterrorism are very crucial for critical infrastructure of any state (Fig. 1), because these may have serious and tragic consequences. For instance in Civil Aviation (CA) any unauthorized access to control system calls into shot hundreds and thousands of passengers' lives. CA is moving away from traditional radar systems in favor of more modern digital tools connection – the problem is that new technologies potentially allow attackers get stuck between the pilot and dispatcher. The guidance aviation security documents declare following requirements to ensure cybersecurity [1].

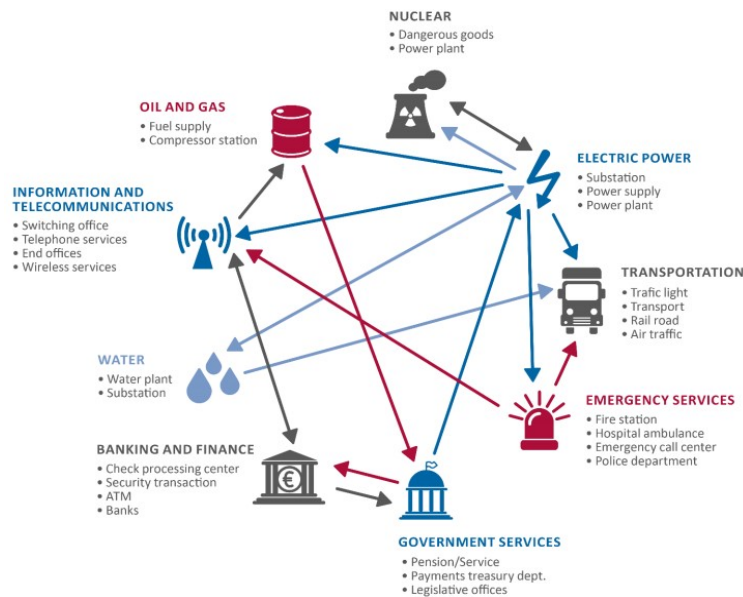


Fig.1. Critical infrastructure in accordance to ENISA

European and world requirements for CA cybersecurity

ECAC Doc 30 [2] declares that measures addressing cyber threats to CA have been included in the National Civil Aviation Security Programme, the National Quality Control Programme and the National Civil Aviation Security Training Programme. A set of security control consists of below measures [2]:

- 1) Implementation of effective measures to protect Critical Aviation Information Systems (CAIS);
- 2) Including the CAIS in their threats assessment processes;
- 3) Separating the CAIS networks from public;
- 4) Responsibility for securing CAIS is allocated by operators to a properly selected, recruited and trained individual;
- 5) Security measures are considered in the design, implementation, operation and disposal of new CAIS;
- 6) Supply chain security measures for hardware and software should be applied to CAIS;
- 7) Remote access to CAIS is only permitted under pre-arranged and secure conditions;
- 8) Cyber attack incidents must be recorded for future evaluation and counter & preventive measures efficiency increasing.

It is also worth noting that the most comprehensive list of measures to mitigate cyberthreats' negative influence on CAIS there is in ICAO Doc 8973 [3]. Among them is noted as follows: 1) Administrative Measures; 2) Virtual (Logic) Control Measures; 3) Physical Controls. Besides this document, also focuses on CAIS: security by design, networks separation & secured remote access for legitimate users, supply chain security & cyber attack incidents records [3]. Despite the examples of the last cyberattacks and requirements of guidance aviation security documents, the main purpose of this paper is offer an integrated complex approach to provide CA cybersecurity.

Basic issues of cyberspace security

Modern threats to information security (cybersecurity) [4] characterizes of asymmetric and flexibility. Cyberattacks has long ceased to be an end in itself and become an effective means to achieve a wide range of purposes; their variety is limited only by the imagination and fantasy of violator. All existed cyberattacks can be differentiated into 3 categories: attacks adversely affecting on confidentiality, integrity and availability on the information. All other types are derived from these. By the way, confidentiality, integrity and availability (so-called CIA-triad) are the main features of information security (and consequently cybersecurity).

In [2] cyberspace was defined as the complex environment resulting from the interaction of people, software and services on the Internet by means of technology devices and networks connected to it, which does not exist in any physical form. Multi-criteria analysis was carried out and the following conceptual

definitions were proposed [4]: *Cyberspace* is virtual space resulting from the interaction of users, software, hardware and network technologies (including Internet) for information (electronic information resources) transformation processes on purpose of ensuring the information needs of society. *Cyberterrorism* is kind of terrorism that is a conscious and purposeful application of information system resources to implement terrorist acts in cyberspace and also to achieve other related purposes in terrorist groups' interests. *Cyberattack* is attempt or realization of security threats in cyberspace aimed at its components (in particular confidentiality, integrity and availability) considering its vulnerabilities.

Complex approach to provide European CA cybersecurity

Considered examples of cyberattacks and acts of cyberterrorism are very crucial for critical infrastructure of any state, because these may have serious and tragic consequences. For instance in CA any unauthorized access to control system calls into shot hundreds and thousands of passengers' lives. CA is moving away from traditional radar systems in favor of more modern digital tools connection – the problem is that new technologies potentially allow attackers get stuck between the pilot and dispatcher.

As well many tools to take control of aircraft in the air were created and tested successfully. During The Fourth HITB Annual Conference (was held in Amsterdam, 2013) the practical demonstration on how to remotely attack and take full control of an aircraft was carried out. The attack performed will follow the classical methodology, divided in discovery, information gathering, exploitation and post-exploitation phases. The complete attack will be accomplished remotely, without needing physical access to the target aircraft at any time, and a testing laboratory will be used to attack virtual airplanes systems. ADS-B (Automatic Dependent Surveillance – Broadcast) and ACARS (Aircraft Communications Addressing and Reporting System) protocols will be used during the discovery and information gather phases, but none of those protocols are the objective of this research, I will just use them to plot and analyze the potential targets. Very basic information on such protocols will be displayed as well as additional references for further reading. The real target of the attacks will be some on-board systems, complex enough to be vulnerable to (almost) common vulnerability research and exploitation techniques. Different post-exploitation vectors will finally be considered in order to gain better aircraft control.

Today we know many examples of attacks to CA throughout the world (e.g. Malaysia, Turkey et al), but most of these cases are not advertised. This is the purpose of hidden blocking vulnerabilities, but appearance of unwanted consequences can make the world community in a loud voice to talk about cyberterrorism in aviation and respond in the short term.

The cybersecurity will be successful only if a comprehensive approach to building a system of CAIS security (Fig.2). That's why Ukrainian complex approach includes a set of organizational and engineering measures that are intended to secure from disclosure, leakage and unauthorized access.

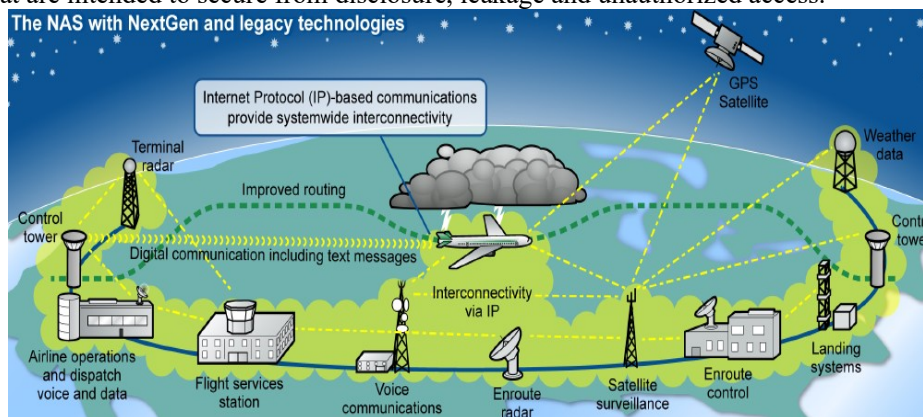


Fig. 2. CA as infrastructure with NextGen technologies

Let's look at stages of *CAIS' complex information security system (CISS)* development in detail (steps from 1st to 13th):

1. *Documents Preparing*. At this stage, the project of documents is prepared that defines the

organizational component of the system (the order project of CISS creation, the condition project of Information Security (IS) service, projects on job descriptions, procedures etc.) that are approved by the administration. It can be also created an IS service or appointed persons who are entrusted to ensure IS and control over them. Responsibility relies on the owner of the system.

2. *Audit.* The following documents are developed: the Certificate of CAIS inspection (contains the description, principles of CAIS construction and architecture); the list of CAIS objects which require protection. During the CAIS inspection should be analyzed and described: the general block diagram and structure (the list and structure of the equipment, technique and software, their communication, features of configuration, architecture and topology, program and hardware-software IS means, mutual arrangement of means etc.); types and characteristics of liaison channels; interactive features of separate components, their mutual influence in private; possible restrictions etc. Such characteristics of the physical environment are a subject matter of the analysis: territorial arrangement of CAIS components (the general plan, the situational plan); availability of territory protection and facility access procedure; the influence of environment factors, protection from the means of technical investigation; availability of communication elements, life-support systems and communications, which have an output for borders of a controllable zone; availability and characteristics of grounding systems; storage conditions of magnetic, paper and other data carriers; availability of the design and operational documentation on components of physical environment.

3. *Threats & Violator Model Development.* Using the information which is presented in the Environment Certificate of Inspection of CAIS operation, the potential threats to the data are defined. There is a research of some possible ways for CAIS data threats realization that is: outflow channels; special purpose channels & unauthorized data access methods. The results of Environmental Inspection of CAIS operation affirm the list of protection objects, as well as potential data threats are defined and the model of threats and violator model are developed. Model of threats (threats model) is the abstract formalized or unformalized methods and means description of threats realization. Violator model is the comprehensive structured characteristic of the perpetrator which is used together with the threats model for development of IS policy. On what information properties violation or CAIS threat is directed: confidentiality violation, integrity violation, data availability violation, surveillance and management of CAIS violation.

4. *Security Policy Development.* Security policy is a set of requirements, rules, restrictions, and recommendations etc. which regulate the data processing order and directed on IS from the certain cyber threats. Security policy offers the following guarantees: a) It is provided an adequacy of IS level to a level of its criticality in CAIS; b) Realization of IS measures profitability; c) In any environment of CAIS operation the assessment and testing are provided; d) Personification of security policy positions is provided, the reporting (registration, audit) for all critical from the security point of view resources to which an access is provided during CAIS operation; e) The personnel and users are provided with the Full Documentation Set according to the IS support; f) Critical from the point of view of IS CAIS technologies (function) have all corresponding support plans of continuous work and its renewal in case of unforeseen situations.

5. *Technical Specification.* The main stages of technical specification formation: a) Classification and description of CAIS resources; b) Development (design) of an information model for existing CAIS, information CAIS flows, interfaces between the user and CAIS etc; c) A list of threats and possible channels of information leakage determination; d) Expert Assessments of expected loss in case of threats; e) Identification of security services; f) Requirements identification for organizational, physical and other protective measures implemented in addition to the complex software and hardware protection; g) Requirements identification for metrological work; h) Models identification that is designed, and technological stand; i) Cost-efficiency assessment of selected assets; j) Making the final decision on the CISS content.

6. *Technical Project.* CISS technical project is developed on founding and in accordance with technical specification on CISS creation. In the process of CISS project (design) development there are proved and designed decisions which give an opportunity to realize technical specification requirements, to provide compatibility and co-operation of different CISS components, and also different measures and technical specification methods. A technical project on CISS creation includes: a) Development of general design decisions necessary for realization of technical specification on CISS requirements; b) The decision

on CISS structure, operation algorithms and conditions of use of defense (security) facilities; c) The decision on CISS architecture and implementation mechanism, defined by a functional structure of IS services; d) Procedure description of technical events on support of CISS development sequence, architecture, tests, the operational environment and CISS documentation according to the set of realization guarantees of security services; e) Development, registration, coordination and the documentation affirmation corresponding to the technical specification size, on CISS; f) Documentation development on IS resources supply and-or technical requirements on their development; g) Preparation and documentation registration on security means deliveries or production containing them in the structure, for CISS complete set (configuration); h) Development, registration and the affirmation of working and operational CISS documentation and, if necessary, its separate component parts.

7. *IS Plan*. At this stage it is required to realize organizational, primary, technical and basic technical measures of restricted access IS, to establish required IS zones, to lead certification of technical equipment of an information activity support, IS means, workplaces (facilities) according to the IS requirements.

8. *Operational Documentation*. At this stage there is a development, registration and the affirmation of working and operational documentation and, if necessary, its separate parts. The working documentation contains detailed decisions on CISS design realization, maintenance of CISS management and interaction of its components, and also the necessary documentation for testing, carrying out of starting-up and adjustment works, carrying out CISS tests.

9. *CISS Implementation*. Implementation of organizational IS measures in CAIS provides: work on administrative documents preparations which regulate an activity of CISS support; compiling of instruction to person who participates in processing or IS in CAIS according to the list specified in the project on CISS; completion of work and the affirmation of documents which are included into IS plan in CAIS except those documents for which the results of the following stages are necessary. Commissioning works, according to the requirements of the preliminary CISS design in CAIS, provide installation, initialization and testing the work ability of CISS. Installation and initialization of CISS, which has the expert conclusion about its compliance with the requirements of normative documents, is carried out in accordance with the procedure specified in the maintenance documentation for this complex.

10. *Preliminary Tests*. The purpose of the preliminary tests is checking the work ability of the CISS and possibility of taking it to the research operation. The CISS work ability and its compliance with technical specification requirements is checked during the tests. Preliminary tests are carried out according to the program and test methods. Developer of the CISS prepares program and test methods and the customer agrees CAIS. Results of the preliminary tests are reiterated in «Protocol Testing», which contains findings regarding the possibility of taking CISS in research operation (exploitation), as well as a list of identified weaknesses, the necessary measures for its removal, and recommended time for doing these tasks.

11. *Research Operation*. During research operation of the CISS: a) Technologies of information processing, a turnover of machine data carriers, management of security means, access differentiations of users to CAIS resources and the automated control over users' actions are examined; b) Employees and IS users get practical skills with the help of technical and hardware-software IS means, study conditions of organizational and administrative documents concerning access differentiation to technical means and information resources; c) Performing (if necessary) the revision of the software, additional tuning and configuration. According to the results of the arbitrary shape work the report on completion of the research is operation drawn up, which includes the conclusion on the possibility (impossibility) of CISS representation on the public examination.

12. *Public Examination*. Public examination of the CISS is a separate step of acceptance tests of the CAIS. Public examination is conducted to determine CISS conformity with the requirements of normative documents on IS and its possibility of introducing CISS consisting of CAIS in the operation (exploitation).

13. *CISS Support*. CISS support contains (accordantly IS plan and operational documentation): warranty & after sales technical service.

Conclusions

Accordingly in this paper, based on guidance aviation security documents and analysis of last attacks in cyberspace, the complex approach to ensure CA cybersecurity was offered. It consists of 13 steps (from

documents preparing to exploitation and support) and its implementation can allow to provide an effective cybersecurity of CAIS as well as European CA.

References

1. ISO/IEC 27032, Information technology – Security techniques – Guidelines for cybersecurity, 2012, 50 p.
2. ECAC Policy Statement in the Field of Civil Aviation Security, 13th edition, 2010, 138 p.
3. Doc 8973, Aviation Security Manual, 10th edition, ICAO, 2017, 808 p.
4. S. Gnatyuk, Critical Aviation Information Systems Cybersecurity, Meeting Security Challenges Through Data Analytics and Decision Support, NATO Science for Peace and Security Series, D: Information and Communication Security. – IOS Press Ebooks, Vol.47, №3, pp. 308-316, 2016.