I. J. Computer Network and Information Security, **2024, 4, 144-153**
Published Online on August 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijcnis.2024.04.10

Targeted Attacks Detection and Security Intruders Identification in the Cyber Space

Zhadyra Avkurova*
Department of AI Technology, NAO Karaganda Industrial University, Temirtau, 101400, Republic Str 30, Kazakhstan
E-mail: zhadyra.avkurova.83@gmail.com
ORCID iD: https://orcid.org/0000-0002-0706-6075
*Corresponding author

Sergiy Gnatyuk
Faculty of Computer Science and Technology, National Aviation University, Kyiv, Ukraine
E-mail: s.gnatyuk@nau.edu.ua
ORCID iD: https://orcid.org/0000-0003-4992-0564

Bayan Abduraimova
Department of Computer Science, L. N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan
E-mail: abduraimova_bk@enu.kz
ORCID iD: https://orcid.org/0000-0003-3913-1895

Kaiyrbek Makulov
Department of Computer Science, Yessenov Univeristy, Aktau, Kazakhstan
E-mail: kaiyrbek.makulov@yu.edu.kz
ORCID iD: https://orcid.org/0000-0002-0826-0371

Received: 12 May 2023; Revised: 15 July 2023; Accepted: 09 October 2023; Published: 08 August 2024

Abstract: The number of new cybersecurity threats and opportunities is increasing over time, as well as the amount of information that is generated, processed, stored and transmitted using ICTs. Particularly sensitive are the objects of critical infrastructure of the state, which include the mining industry, transport, telecommunications, the banking system, etc. From these positions, the development of systems for detecting attacks and identifying intruders (including the critical infrastructure of the state) is an important and relevant scientific task, which determined the tasks of this article. The paper identifies the main factors influencing the choice of the most effective method for calculating the importance coefficients to increase the objectivity and simplicity of expert assessment of security events in cyberspace. Also, a methodology for conducting an experimental study was developed, in which the goals and objectives of the experiment, input and output parameters, the hypothesis and research criteria, the sufficiency of experimental objects and the sequence of necessary actions were determined. The conducted experimental study confirmed the adequacy of the models proposed in the work, as well as the ability of the method and system created on their basis to detect targeted attacks and identify intruders in cyberspace at an early stage, which is not included in the functionality of modern intrusion detection and prevention systems.

Index Terms: Targeted Attacks, Cybersecurity, Detection, Information Infrastructure, Fuzzy Logic, APT, Modelling.

1. Introduction

The introduction of ICT in all spheres of human activity, society and the state help to increase the efficiency of various business processes, but there is an acute problem of information security [1]. The number of new cybersecurity threats and opportunities is increasing over time, as well as the amount of information that is generated, processed, stored and transmitted using ICTs. Particularly sensitive are the objects of critical infrastructure of the state, which include the mining industry, transport, telecommunications, the banking system, etc. It is also worth noting that critical objects are often become a target for intruders, therefore, first of all, the state must ensure a sufficient level of cybersecurity for critical infrastructure objects, which must be clearly defined and protected. To effectively solve the problem of malicious activity of the intruder, his identity should also be identified. An analysis of the behavioral characteristics of an intruder makes it possible to determine at least his belonging to a certain group of persons (profiling), which is very important for further investigation. In this aspect, attack detection systems that can not only detect but also identify cybersecurity intruders are important for science and practice in this industry.

2. Review of Related Papers

In the previous works of the authors [2-4], the procedure for the formation of the DS is described, while one of the stages is the choice of alternatives of linguistic identifiers, carried out by one of the methods for determining the IC, reflecting the advantages of the expert in the assessment. Consider examples of the practical application of some of these methods and determine the criteria for choosing an appropriate method for use in the system being developed [5-7].

Example 1 – RM. For information system resources (ISR), experts should rank the alternatives displayed by security parameters (information confidentiality (IC), pass by passes (PBP), secure password storage (SPS), security incident recording (SIR), personnel selection and testing (PST); access to premises (AP). Table 1 shows the ranks given to each expert. At the same time, for the second expert, the alternatives of PBP, SPS and AP are equivalent, so he determined their rank as the arithmetic mean of ranks 2, 3 and 4: (2+3+4)/3=3. For the third expert, the parameters of IC and PST turned out to be equivalent, so their rank is defined as (3+4)/2=3.5. Calculations are reflected in table. 1.

Table 1. Ranks

Parameters	Experts				
	1	2	3	4	
IC	2	1	3,5	3	
PBP	1	3	2	4	
SPS	3	3	1	2	
SIR	6	5	5	5	
PST	5	6	3,5	6	
AP	4	3	6	1	

Example 2 – DS method. The expert was asked to analyze six parameters (see note 1) and determine their relative advantage. For each parameter, the set of alternatives $A \in \{IC, PBP, SPS, SIR, PST, AP\}$ is determined by the estimate Π_i that displays the number of security parameters more important than the current. The benefit vector looks as follows $\Pi = \{0,3,1,5,4,2\}$

Example 3 – VM method. Let us consider the case of pairwise comparison of five security parameters of ISR. The PW matrix has the form like (1):

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & 2 & 1 \\ 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 2 \\ 0 & 2 & 0 & 1 & 1 \\ 1 & 2 & 0 & 1 & 1 \end{bmatrix} \tag{1}$$

The first iteration is obtained by adding elements of each row of the matrix: $P^1(1) = \sum_{j=1}^{5} a_{1j} = 1 + 0 + 2 + 2 + 1 = 6$, $P^2(1) = 4$, $P^3(1) = 6$, $P^4(1) = 4$, $P^5(1) = 5$. It can be seen that two criteria are in first place (6 points) and two are in third place (4 points). The second iteration is carried out by adding the products of the matrix row element and the corresponding result of the preliminary iteration: $P^1(2) = a_{11} \cdot P^1(1) + a_{12} \cdot P^2(1) + ... + a_{15} \cdot P^5(1) = (1 \times 6) + (0 \times 4) + (2 \times 6) + (2 \times 4) + (1 \times 5) = 31$, $P^2(2) = 22$, $P^3(2) = 28$, $P^4(2) = 17$, $P^5(2) = 23$. Apparently, the first criterion takes the first place and for further iterations the distribution will be unchanged - the first criterion will be in the first place: $P^1(3) = a_{11} \cdot P^1(2) + a_{12} \cdot P^2(2) + \ldots + a_{15} \cdot P^5(2) = (1 \times 31) + (0 \times 22) + (2 \times 28) + (2 \times 17) + (1 \times 23) = 144$, $P^2(2) = 112$, $P^3(3) = 130$, $P^4(3) = 84$, $P^5(3) = 115$. Thus, the results of the third iteration can be substituted into (1) and the IC can be calculated by (2) – (4):

$$P_3^1 = \frac{144}{(144+112+130+84+115}) = \frac{144}{585} = 0{,}24; \tag{2}$$

$$P_3^2 = \frac{112}{585} = 0{,}191; P_3^3 = \frac{130}{585} = 0{,}22; P_3^4 = \frac{84}{585} = 0{,}14; \tag{3}$$

$$P_3^5 = \frac{115}{585} = 0{,}196. \tag{4}$$

Example 4 – SM. A scale of relative importance is used: 0 - it is difficult for an expert to compare, 1 - equal contribution of two alternatives, 3 - experience and judgment give a slight advantage of one alternative to another, 5 -

experience and judgment give a strong advantage of one alternative to another, 7 - one of the alternatives significantly dominates the other, 9 - the advantage of one alternative is obvious; 2,4,6,8 - compromise cases. If, when comparing the first alternative with another, the above number (3) is obtained, then when comparing the second alternative with the first, the reciprocal value (1/3) is obtained. The importance of four ISR security parameters should be assessed (see Note 1). Calculations are reflected in table. 2.

Table 2. Calculations of IC

Parameters	IC	AP	SPS	PBP	$\tilde{\lambda}_i$	λ_i
I	II	III	IV	V	VI	VII
IC	1	3	5	9	3,41	0,581
AP	1/3	1	3	5	1,5	0,255
SPS	1/5	1/3	1	3	0,67	0,114
PBP	1/9	1/5	1/3	1	0,29	0,050
Total					5,87	1

Example 5 – MR method. Let it be necessary to determine prizes for four ISR protection projects *i={1,2,3,4}* of the development companies AA, BB, CC, DD. All projects were sent to four experts *i={1,2,3,4}.* Table 3 shows the ranks of projects that were assigned by each of the experts: $x_1 = \frac{(1+2+1+1)}{4} = \frac{5}{4} = 1{,}25$; $x_2 = \frac{15}{4} = 3{,}75$; $x_3 = \frac{8}{4} = 2$; $x_4 = \frac{12}{4} = 3$. IC are defined as $\lambda_i = \frac{x_i}{N}$, where N - the sum of all ranks (N=10). The results show that the first place is AA, the second is CC, the third is DD, the fourth is BB. The advantages of the method are ease of use and speed of calculation. It is possible to increase the accuracy by introducing additionally the notion of expert weights (qualification coefficient). Calculations are reflected in table 3

Table 3. Ranks of IC projects

Project	Expert				x_i	λ_i
	1	2	3	4		
AA (1)	1	2	1	1	1,25	0,12
BB (4)	4	4	3	4	3,75	0,37
CC (2)	3	1	2	2	2	0,2
DD (3)	2	3	4	3	3	0,3

Example 6 – MC method. Let's determine the optimal among ISR protection projects using example 5. The experts evaluated the projects according to three parameters - cost (B); labor intensity (Tr.) and developer experience (D), the importance of which, is $\lambda_i = [0{,}5 \quad 0{,}2 \quad 0{,}3]$, respectively, on a five-point scale. Estimates are presented in table 4.

$$f(1) = min\ \ [f_1(1)\cdot\lambda_1; f_2(1)\cdot\lambda_2; f_3(1)\cdot\lambda_3] = min\ \ [4\cdot 0{,}5; 3\cdot 0{,}2; 4\cdot 0{,}3] = 0{,}6; f(2) = min\ \ [f_1(2)\cdot\lambda_1; f_2(2)\cdot\lambda_2; f_3(2)\cdot\lambda_3] = min\ \ [3\cdot 0{,}5; 1\cdot 0{,}2; 2\cdot 0{,}3] = 0{,}2; f(3) = min\ \ [0{,}5; 0{,}4; 1{,}2] = 0{,}4; \quad f(4) = min\ \ [1; 0{,}4; 0{,}6] = 0{,}4.$$

The optimal solution is defined as the maximum:

$$U(f_1(x), \dots, f_n(x)) = max\ \ [f(1); f(2); f(3); f(4)] = [0{,}6; 0{,}2; 0{,}4; 0{,}4] = 0{,}6.$$

Thus, the best will be the project proposed by the AA design group.

Table 4. Evaluation of projects according to criteria

Project	Criterion		
	B	Tr.	D
AA	4	3	4
BB	3	1	2
CC	1	2	4
DD	2	2	2

Example 7 – NPM method. Let us determine the best ISR security project using example 6 (see Table 4), but this time we will set the optimal estimates according to the criteria as $x_i^* = [3; \quad 2; \quad 3]$. Let us construct from (4) a metric for *p*=2, and having calculated it for each of the projects, we will determine the closest one to the optimal criteria.

$$L_1(x) = \sqrt{\left(0{,}5\left(\frac{(3-4)}{(3-4)}\right)\right)^2 + \left(0{,}2\left(\frac{(2-3)}{(2-3)}\right)\right)^2 + \left(0{,}3\left(\frac{(3-4)}{(3-4)}\right)\right)^2} \approx 0{,}62;$$

$$L_2(x) = \sqrt{\left(0{,}5\left(\frac{(3-3)}{(3-4)}\right)\right)^2 + \left(0{,}2\left(\frac{(2-1)}{(2-3)}\right)\right)^2 + \left(0{,}3\left(\frac{(3-2)}{(3-4)}\right)\right)^2} \approx 0{,}36;$$

$L_3(x) \approx 1{,}04$; $L_4(x) \approx 0{,}58$. Therefore, under certain conditions, the project of the BP group is optimal.

So, we see that there are many ways to determine the IC, each of which is fundamentally applicable to the formation of the final results of the examination, but in some cases one method should be preferred, while others cannot be applied for various reasons. The most important factor influencing the choice of the IC accounting method is the physical nature of the parameters and the relationship between them. The parameters are determined based on the purpose of the examination. Next, you should determine the degree of relationship between the parameters - dependence or independence - and the nature of the relationship (independence by utility, by advantage, by indifference, etc.)

A significant factor is the complexity of conducting an examination and the complexity of obtaining expert information, which is determined by real languages and the possibilities of conducting an examination. The least time of communication with experts requires ranking and the TRS method; the LS method requires more time for communication with experts (12 times more than the ranking and 2 times more than the ChA method).

Table 5. Comparison of methods

Methods				Form of expression		RSc	Tr.
				In.D	Out.D		
PC	LS			M	L	ScR	M
	EV	VM		M	L	ScR	M
		SM		M	T	ScR	H
		CYM		M	T	ScR	M
		NM		M	T	ScR	H
		YM		M	T	ScR	H
RC	MR			T	L	ScO	L
	TR	RAMF		T	L	ScO	L
		APR	ChA	L	L	ScO	M
			PL	L	L	ScO	M
UFA	UPC	AC	SM	LKL	L	ScR	H
			YM	LKL	L	ScR	H
			CRM	LKL	L	ScR	H
		MCM		M	L	ScR	L
	VF	MCM		L	L	ScR	H
		PCM		L	L	ScR	H
FT	AM	TRS		L	G	ScI	L
		DMA		L	G	ScI	L
	CMM			L	M	ScO	H
		RV		L	G	ScI	H
EPD	IPD	ChC		L	G	ScI	H
		NPM		L	L	ScR	H
		CMM		L	L	ScR	L
	EPD	CGT		L	G	ScI	M
		GT		L	G	ScI	M

The choice of the method for determining the IC is also influenced by the degree of consistency between the assessments of professionals. The degree of consistency primarily depends on the number of experts in the EG and the level of their qualifications. At the same time, it is influenced by the chosen method for estimating IC. Thus, the greatest consistency of experts is provided by the LS, the least - by the direct numerical assessment, while the ranking, for all its simplicity, makes it possible to obtain fairly accurate IC and close to their value obtained by the LS method. The complexity of processing expert data is not the main factor at the current level of development of computer technology. However, the use of complex methods for processing expert information may require the development of special software, which will affect the timing of the examination. Obviously, the simplest methods from this point of view are the rank and point methods. For greater clarity, the results of comparing the methods for determining IC are summarized in table 5, where input (In.D) and output data (Out.D.); matrix (M), tabular (T), linear (L), graphic (G) form of data expression; scales of names (ScN), order (ScO), intervals (ScI), relations (ScO); recommended scale (RSc), labor intensity (Tr.), which is rated as high (H), medium (M) and low (L).

The purpose of the work is to develop and study a software tool for detecting targeted attacks and identifying cybersecurity intruders at critical infrastructure objects.

3. Software Tool Realization

On the basis of the proposed method of experimental research and mathematical models [2-3], the software “Intruder detection and identification system” was developed for simulation modeling. In this program, you can run the procedures for the formation and correction of standards of linguistic changes, tuples and DR, as well as configure the parameters of the simulation. The appendices contain its source code.

The executable software module can be used on any computer that meets the minimum requirements for operation:

- Processor Intel Pentium IV/Xeon 2.4 GHz or higher
- RAM 1024 MB or more
- Hard drive 40 GB or more
- OS - Microsoft Windows.

To work with constant and conditionally constant information with a certain set of values, objects of the "Directory" type are used in the system. To implement the proposed mathematical models, the following objects of the "Directory" type were created: "Tuples", "Parameters", "Linguistic Variables", "Intervals", "Experts", "Intruder Types", "Linguistic Identifiers". You can access these directories from the "References" menu in the interface.

The "Tuples" directory stores tuples used during system operation to detect and identify an intruder and a list of parameters of which he consists.

The "Experts" directory is used for storing information on all experts who participated in the study (the number of experts can be large, and during the experiment it is possible to select an expert with the necessary competence, who formed standards and rules), and the "Intruder Types" directory is necessary for storing a list of intruders identified in the pilot study. The Figure 2b shows the window of the list of the "Types of Intruder" directory.

The "Parameters" directory stores all the parameters involved in tuples. Also, intervals and linguistic variables are indicated in each parameter.

The "Linguistic Variables" and "Intervals" directories are used to store information about the above linguistic variables and intervals. "Linguistic Identifiers" directory serves for stored information on all linguistic identifiers used in the study.

With the help of objects of the "Documents" type, the input of information about the implementation of any operations into the system, as well as their viewing and correction, is organized. In the developed configuration, the following objects of the "Documents" type were created: "Setting Standards", "Setting Rules", "Experiment". You can access these documents from the "Documents" menu in the interface.

Parameter standards are set by an expert in the document "Setting Standards" (for example, see Fig. 3). On the "Expert data" tab, the expert fills in the values of each of the parameter change intervals. The software tool automatically plots the membership function of linguistic change standards.

In the "Setting rules" document, the expert sets the linguistic identifiers used in the system on the "LI" tab and assigns one of them (L, MLH, MHL, H, C) to the corresponding rules on the "Rules" tab as described in […].

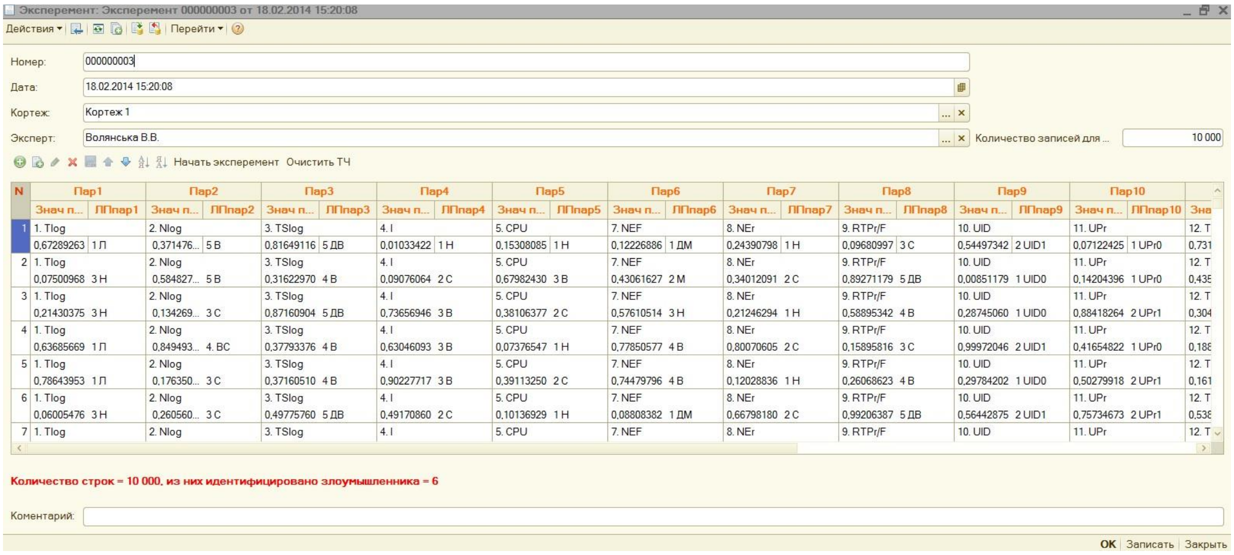
After specifying the standards and linguistic identifiers with the help of the "Experiment" document, it is possible to simulate the state of cyberspace and identify the fact of the violation and the category of the intruder.

In the form of the "Experiment" document element, the number of series of experience is set in the column "Number of records for the experiment", as well as other characteristics (date, tuple, expert) and it is launched (the element form window is shown on Fig. 1).

Reports are used to obtain summary information based on the data entered in the system. In the developed configuration, only one object of the Reports type was created - the Modeling Report. Using this report, you can access a detailed report on the results of the simulation carried out by the software tool. Through the Reports menu in the interface, you can access this report.

4. Experimental Study

The purpose of any experimental research is to identify the qualities of the objects under study, verify the authenticity of hypotheses, as well as a wide and in-depth study of the scientific topics under study. There are many different classifications of experiments depending on the field of science, the purpose of the study, the structure of objects and phenomena, organizational measures, the nature of the interaction of the object and the means of research. Of particular importance in the course of the experiment is the correct development of the experimental methodology - a certain sequence of processes, as a result of which the goal of the study is achieved. Specifically, the correctness of the development of experimental research methodology describes its value.



Эксперемент: Эксперемент 000000003 от 18.02.2014 15:20:08

Номер: 000000003

Дата: 18.02.2014 15:20:08

Кортеж: Кортеж 1

Эксперт: Волянська В.В.

Количество записей для ... 10 000

N	Пар1		Пар2		Пар3		Пар4		Пар5		Пар6		Пар7		Пар8		Пар9		Пар10		
	Знач п...	ЛПпар1	Знач п...	ЛПпар2	Знач п...	ЛПпар3	Знач п...	ЛПпар4	Знач п...	ЛПпар5	Знач п...	ЛПпар6	Знач п...	ЛПпар7	Знач п...	ЛПпар8	Знач п...	ЛПпар9	Знач п...	ЛПпар10	Зна
1	1. Tlog		2. Nlog		3. TSlog		4. I		5. CPU		7. NEF		8. NEr		9. RTPr/F		10. UID		11. UPr		12. T
	0,67289263	1 Л	0,371476...	5 В	0,81649116	5 ДВ	0,01033422	1 Н	0,15308085	1 Н	0,12226886	1 ДМ	0,24390798	1 Н	0,09680997	3 С	0,54497342	2 UID1	0,07122425	1 UPr0	0,731
2	1. Tlog		2. Nlog		3. TSlog		4. I		5. CPU		7. NEF		8. NEr		9. RTPr/F		10. UID		11. UPr		12. T
	0,07500968	3 Н	0,584827...	5 В	0,31622970	4 В	0,09076064	2 С	0,67982430	3 В	0,43061627	2 М	0,34012091	2 С	0,89271179	5 ДВ	0,00851179	1 UID0	0,14204396	1 UPr0	0,435
3	1. Tlog		2. Nlog		3. TSlog		4. I		5. CPU		7. NEF		8. NEr		9. RTPr/F		10. UID		11. UPr		12. T
	0,21430375	3 Н	0,134269...	3 С	0,87160904	5 ДВ	0,73656946	3 В	0,38106377	2 С	0,57610514	3 Н	0,21246294	1 Н	0,58895342	4 В	0,28745060	1 UID0	0,88418264	2 UPr1	0,304
4	1. Tlog		2. Nlog		3. TSlog		4. I		5. CPU		7. NEF		8. NEr		9. RTPr/F		10. UID		11. UPr		12. T
	0,63685669	1 Л	0,849493...	4. ВС	0,37793376	4 В	0,63046093	3 В	0,07376547	1 Н	0,77850577	4 В	0,80070605	2 С	0,15895816	3 С	0,99972046	2 UID1	0,41654822	1 UPr0	0,188
5	1. Tlog		2. Nlog		3. TSlog		4. I		5. CPU		7. NEF		8. NEr		9. RTPr/F		10. UID		11. UPr		12. T
	0,78643953	1 Л	0,176350...	3 С	0,37160510	4 В	0,90227717	3 В	0,39113250	2 С	0,74479796	4 В	0,12028836	1 Н	0,26068623	4 В	0,29784202	1 UID0	0,50279918	2 UPr1	0,161
6	1. Tlog		2. Nlog		3. TSlog		4. I		5. CPU		7. NEF		8. NEr		9. RTPr/F		10. UID		11. UPr		12. T
	0,06005476	3 Н	0,260560...	3 С	0,49775760	5 ДВ	0,49170860	2 С	0,10136929	1 Н	0,08808382	1 ДМ	0,66798180	2 С	0,99206387	5 ДВ	0,56442875	2 UID1	0,75734673	2 UPr1	0,538
7	1. Tlog		2. Nlog		3. TSlog		4. I		5. CPU		7. NEF		8. NEr		9. RTPr/F		10. UID		11. UPr		12. T

Количество строк = 10 000, из них идентифицировано злоумышленника = 6

Коментарий:

Fig.1. The form window of the "Experiment" document element

The first step in conducting an experimental study is to draw up a plan-program of the study:

- Purpose and objectives of the experiment: a) study of the proposed SVIP based on expert methods and fuzzy logic regarding the effectiveness of its work, namely, the correctness of identifying the fact of an IS violation and the process of categorizing the intruder.
- Choice of input and output parameters:
- input parameters - fuzzy Tlog, Nlog, TSlog, I, CPU, NEF, NEr, RTPr/F and crisp parameters UID, AtEF, TrFin, ModF, TrFout, KS); output parameters - the degree of confidence of the expert in the decision to identify the fact of the activity of the intruder (L, MLH, MHL, H, C) and the category of the intruder (D, S, C, H, SB, B).
- The sequence of actions is described in previous articles [2-3], where the method and system for detecting and identifying an intruder are considered.
- Input fuzzy parameters change from 0 to 1, and crisp parameters take a value strictly 0 or 1.
- Tools used: 1C Enterprise environment was used for the study.
- Analysis of simulation results will be presented in the next section.

The second step is carried out after the approval of the methodology - this is the determination of the scope of experimental studies and the necessary means (software, hardware, etc.). The third step is the direct conduct of the experiment, and the final step is the processing of experimental data, the systematization of all numerical data, the verification of convergence into a single system of units, the construction of dependency graphs, tables, diagrams, etc.

In accordance with the developed methodology, experimental studies were carried out. Next, we describe the course of the experiment, as well as the processing and analysis of its results.

In the course of the experiment, with the help of the developed SVIP in cyberspace, in accordance with the specified network and host parameters, 300,000 current states of the cyberspace environment were simulated, of which 207 were characteristic for the actions of certain categories of intruders. Control of all current states of the environment was carried out using 50625 rules, while the total distribution of detections in percentage by category is as follows:

- disinformer (15%),
- spammer (20%),
- cracker (18%),
- hacker (14%),
- spam bot (15%)
- hacker bot (16%).

This distribution (Fig. 2) is 100% expected and corresponds to the simulation conditions. Each category of intruder was uniquely identified by the corresponding group of established rules.

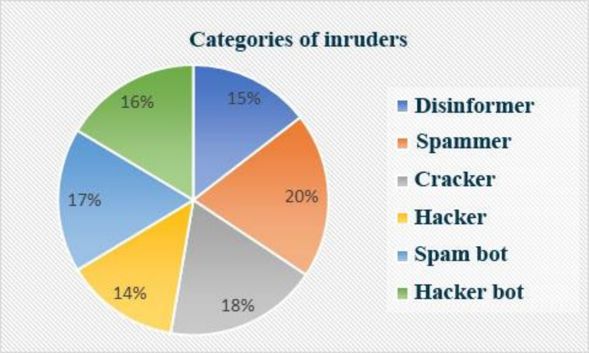


Fig.2. Diagram of the distribution of identified intruders according to their categories

For example, a hacker bot was detected 36 times (16% of the total distribution), while the rules of the ER4 group were activated, of which in 35 cases the wear was carried out with the result of detecting the intruder "High" and, in one case, "Critical".

Detailed results of the experiment are shown in tables 6 and 7. It notes the rule according to which the intruder was identified and the category to which he belongs, the degree of confidence of the expert in the decision (linguistic identifier) and the number of triggers.

Table 6. General report

Tuple		Quantity
Type of Intruder		
L1	The name of the combination is fuzzy	
Tuple 1		300 000
		299 793
Hacker bot		36
Disinfomner		28
Cracker		39
Spam bot		32
Spammer		40
Hacker		32
Total		300 000

5. Discussion

Therefore, the conducted experimental study confirmed the adequacy of the proposed models (their reactions to the simulated actions), as well as the ability of tools created on their basis with extended functionality to detect and identify intruders in fuzzy conditions and thus ensure the effective functioning of CERT/CSIRT at a certain level [8-10].

Let us consider the functioning of the intruder detection and identification system using the example of working with real parameters taken from the sensors developed in the previous work [3]. Let us analyze in detail the values of the controlled parameters (see Table 6) when installing foreign packages, we will process them and evaluate the membership functions in accordance with the developed standards. Tlog login time: 9:15:20 – 9:16:29. The membership function for this parameter falls within the data interval near the point 0.375 - given the developed standard of this linguistic change, this indicator is closest to the value "illegitimate", i.e. $t_{\mathrm{Tlog}} \cong H$. The frequency of login requests parameter to the Nlog system during the password brute force process is 47 login requests per 1 minute, which, according to the standard, corresponds to a value of 0.47, i.e. $t_{\mathrm{Nlog}} \cong BC$. Since, in the interests of convenience, the MUSE and CPU parameters have been combined into one, we will consider the collected data in a complex way. So the share of used RAM has grown from a value of about 19% before the attack to more than 60%. The CPU load, especially at the end of the attack, is close to 100% and is on average in the range of 70-80%, so we get a value close to 0.7, i.e. $t_{\mathrm{CPU}} \cong B$. Regarding the parameter of the number of failures and errors NEr, it should be noted that during the time from 09:12:57 to 09:13:41 the number of failures and errors was recorded 18. According to the developed standards, this indicator is close to the "high" value, i.e. $t_{\mathrm{NEr}} \cong B$. The results also show the appearance of unusual processes, file transfer to the system and file modification, which is evident from a significant increase in disk (the amount of information per record increased from 12 KB during system idle to 24 KB on average during the attack) and network activity (incoming traffic increased from 120 B when idle to 0 600 B on average during an attack). Parameters TSlog - time spent on login; I - intensity of actions; NEF - number of executable files; RTPr/F - the execution time of the process/file was obtained by simulating the "Intruder Detection and Identification System" software tool during the experiment [11-13].

Table 7. Detailed report

Tuple		Quantity	
Type of intruder			
LI	The name of the combination is fuzzy		
Tuple 1		207	
Hacker bot		36	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 1 ДМ, 4. I = 1 Н, 5. CPU = 1 Н, 7. NEF = 1 ДМ, 8. NEr = 1 Н, 9. RTPr/F = 4 В	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 5 ДВ, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 2 М, 8. NEr = 2 С, 9. RTPr/F = 1 ДМ	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 5 ДВ, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 5 ДВ, 8. NEr = 2 С, 9. RTPr/F = 1 ДМ	1	
…	…	…	
5 К	1. Tlog = 3 Н, 2. Nlog = 5 В, 3. TSlog = 4 В, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 1 ДМ, 8. NEr = 3 В, 9. RTPr/F = 1 ДМ	1	
Disinformer		28	
4 В	1. Tlog = 1 Л, 2. Nlog = 4. ВС, 3. TSlog = 5 ДВ, 4. I = 1 Н, 5. CPU = 3 В, 7. NEF = 5 ДВ, 8. NEr = 3 В, 9. RTPr/F = 4 В	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 4. ВС, 3. TSlog = 5 ДВ, 4. I = 2 С, 5. CPU = 3 В, 7. NEF = 2 М, 8. NEr = 3 В, 9. RTPr/F = 5 ДВ	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 4. ВС, 3. TSlog = 5 ДВ, 4. I = 2 С, 5. CPU = 3 В, 7. NEF = 5 ДВ, 8. NEr = 1 Н, 9. RTPr/F = 5 ДВ	1	
…	…	…	
4 В	1. Tlog = 3 Н, 2. Nlog = 5 В, 3. TSlog = 5 ДВ, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 1 ДМ, 8. NEr = 3 В, 9. RTPr/F = 3 С	1	
Cracker		39	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 1 ДМ, 4. I = 1 Н, 5. CPU = 1 Н, 7. NEF = 2 М, 8. NEr = 3 В, 9. RTPr/F = 3 С	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 1 ДМ, 4. I = 1 Н, 5. CPU = 1 Н, 7. NEF = 3 Н, 8. NEr = 3 В, 9. RTPr/F = 3 С	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 4 В, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 1 ДМ, 8. NEr = 3 В, 9. RTPr/F = 2 М	1	
…	…	…	
5 К	1. Tlog = 1 Л, 2. Nlog = 4. ВС, 3. TSlog = 5 ДВ, 4. I = 2 С, 5. CPU = 3 В, 7. NEF = 5 ДВ, 8. NEr = 3 В, 9. RTPr/F = 5 ДВ	1	
Spam bot		32	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 1 ДМ, 4. I = 1 Н, 5. CPU = 1 Н, 7. NEF = 1 ДМ, 8. NEr = 1 Н, 9. RTPr/F = 2 М	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 4 В, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 2 М, 8. NEr = 3 В, 9. RTPr/F = 3 С	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 4. ВС, 3. TSlog = 4 В, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 1 ДМ, 8. NEr = 2 С, 9. RTPr/F = 1 ДМ	1	
…	…	…	
4 В	1. Tlog = 3 Н, 2. Nlog = 5 В, 3. TSlog = 5 ДВ, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 3 Н, 8. NEr = 3 В, 9. RTPr/F = 2 М	1	
Spammer		40	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 1 ДМ, 4. I = 1 Н, 5. CPU = 1 Н, 7. NEF = 3 Н, 8. NEr = 2 С, 9. RTPr/F = 1 ДМ	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 1 ДМ, 4. I = 1 Н, 5. CPU = 1 Н, 7. NEF = 3 Н, 8. NEr = 3 В, 9. RTPr/F = 3 С	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 1 ДМ, 4. I = 1 Н, 5. CPU = 1 Н, 7. NEF = 5 ДВ, 8. NEr = 1 Н, 9. RTPr/F = 1 ДМ	1	
…	…	…	
5 К	1. Tlog = 2 П, 2. Nlog = 5 В, 3. TSlog = 5 ДВ, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 4 В, 8. NEr = 3 В, 9. RTPr/F = 1 ДМ	1	
Hacker		32	
4 В	1. Tlog = 1 Л, 2. Nlog = 1 Н, 3. TSlog = 5 ДВ, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 2 М, 8. NEr = 3 В, 9. RTPr/F = 2 М	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 3 С, 3. TSlog = 4 В, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 2 М, 8. NEr = 3 В, 9. RTPr/F = 2 М	1	
4 В	1. Tlog = 1 Л, 2. Nlog = 4. ВС, 3. TSlog = 4 В, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 2 М, 8. NEr = 3 В, 9. RTPr/F = 2 М	1	
…	…	…	…
4 В	1. Tlog = 3 Н, 2. Nlog = 5 В, 3. TSlog = 4 В, 4. I = 3 В, 5. CPU = 3 В, 7. NEF = 2 М, 8. NEr = 2 С, 9. RTPr/F = 1 ДМ	1	
Total		207	

After compiling the obtained results, we get a data tuple, which leads to the operation of one of the rules ER133 - ER148 (depending on the value of the simulated parameters) and can be identified by the rule ER145 and ER148 with linguistic identifier В (high degree of expert's confidence in the decision).

As we see in the example under consideration, it can be seen from the results that the established decision rules and parameter standards are correct and correspond to real conditions. So when installing foreign packages, which is typical for a cracker, the memory and processor load increases noticeably, disk and network activity increases. These results are fully consistent with the prevailing rules for the cracker.

6. Conclusions

The main factors influencing the choice of the most effective method for calculating the IC in order to increase the objectivity and simplicity of the EO of security events in cyberspace are determined.

A methodology for conducting an experimental study has been developed, which defines the goals and objectives of the experiment, input and output parameters, a hypothesis and research criteria, the sufficiency of experimental objects and the sequence of necessary actions.

The conducted experimental study confirmed the adequacy of the models proposed in the work, as well as the ability of the method and system created on their basis to detect APT attacks and identify intruders in cyberspace at an early stage [14-17], which is not included in the functionality of modern intrusion detection and prevention systems. During the experiments 50625 rules (based on fuzzy logic) were used and various categories of intruders were identified: disinformer (15%), spammer (20%), cracker (18%), hacker (14%), spam bot (15%), hacker bot (16%). The number of rules can be grown as well as the categories of intruders can be changed in different spheres and sectors.

The obtained research results can be used to improve the security and reliability of computer systems and networks as well as security level of critical information infrastructure. Proposed tool can be used by CERT/CSIRT teams.

References

[1] Yu. Danik, R. Hryschuk, S. Gnatyuk, Synergistic effects of information and cybernetic interaction in civil aviation, Aviation, Vol. 20, №3, pp. 137-144, 2016.

[2] Avkurova Z., Gnatyuk S., Abduraimova B., Fedushko S., Syerov Y., Trach O. Models for early web-attacks detection and intruders identification based on fuzzy logic, Procedia Computer Science, 2021, Vol. 198, pp. 694-699.

[3] Avkurova Z., Gnatyuk S., Abduraimova B. Structural and Analytical Models for Early APT-Attacks Detection in Critical Infrastructure, Communications in Computer and Information Science, 2022, Vol. 1635, pp. 455-468.

[4] Gnatyuk S., Berdibayev R., Avkurova Z., Verkhovets O., Bauyrzhan M. Studies on cloud-based cyber incidents detection and identification in critical infrastructure, CEUR Workshop Proceedings, 2021, Vol. 2923, pp. 68-80.

[5] E. A. Burkov, P. I. Paderno, O. E. Siryk, E. A. Lavrov and N. B. Pasko, "Analysis of Impact of Marginal Expert Assessments on Integrated Expert Assessment," 2020 XXIII International Conference on Soft Computing and Measurements (SCM), 2020, pp. 14-17, doi: 10.1109/SCM50615.2020.9198772.

[6] P. D. Reddy and A. Mahajan, "Expert System for Generating Teaching Plan Based on Measurable Learning Objectives and Assessment," 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT), Austin, TX, USA, 2016, pp. 207-208, doi: 10.1109/ICALT.2016.61.

[7] E. Szafranko, "Methodology of an Assessment of Building Construction Variants with the Use of Expert Systems," 2017 Baltic Geodetic Congress (BGC Geomatics), Gdansk, Poland, 2017, pp. 252-256, doi: 10.1109/BGC.Geomatics.2017.13.

[8] Iashvili G., Avkurova Z., Iavich M., Bauyrzhan M., Gagnidze A., Gnatyuk S. Content-Based Machine Learning Approach for Hardware Vulnerabilities Identification System, Lecture Notes on Data Engineering and Communications Technologies, Vol. 83, pp. 117-126, 2021.

[9] Gnatyuk S., Berdibayev R., Smirnova T., Avkurova Z., Iavich M. Cloud-Based Cyber Incidents Response System and Software Tools, Communications in Computer and Information Science, Vol. 1486, pp. 169-184, 2021.

[10] Z. Ni and B. Huang, "Gait-Based Person Identification and Intruder Detection Using mm-Wave Sensing in Multi-Person Scenario," in IEEE Sensors Journal, vol. 22, no. 10, pp. 9713-9723, 15 May15, 2022, doi: 10.1109/JSEN.2022.3165207.

[11] M. Poorani, V. Vaidehi, M. Rajesh, Bharghavi, Balamuralidhar and G. Chandra, "Semantic Intruder Detection System in WSN," ICoAC 2010, Chennai, India, 2010, pp. 26-32, doi: 10.1109/ICOAC.2010.5725357.

[12] K. Rajesh, V. Shaguftha, A. Deepika, S. Rajendran and A. Ramukumar, "Intruder Detection and Adaptive Irrigation System Using IOT," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2021, pp. 712-716, doi: 10.1109/ICESC51422.2021.9532778.

[13] Du Chunlai,hang Jianshun,Ma Li,"Defense on Split-Network Attack in Wireless Sensor Network", International Journal of Wireless and Microwave Technologies, vol.2, no.1, pp.38-44, 2012.

[14] Syed Golam Abid, Muntezar Rabbani, Arpita Sarker, Tasfiq Ahmed Rafi, Dip Nandi, "Comparative Analysis of Threat Detection Techniques in Drone Networks", International Journal of Mathematical Sciences and Computing(IJMSC), Vol.10, No.2, pp. 32-48, 2024.

[15] Savenko, O., Sachenko, A., Lysenko, S., Markowsky, G., & Vasylkiv, N. (2020). Botnet Detection Approach Based on the Distributed Systems. International Journal of Computing, 19(2), 190-198. https://doi.org/10.47839/ijc.19.2.1761

[16] J. Li, T. Li, R. Zhang, D. Wu, H. Yue and Z. Yang, "APM: An Attack Path-based Method for APT Attack Detection on Few-Shot Learning," 2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), Exeter, United Kingdom, 2023, pp. 10-19, doi: 10.1109/TrustCom60117.2023.00025.

[17] C. Sheng and C. Gang, "APT Attack and Detection Technology," 2024 IEEE 6th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Chongqing, China, 2024, pp. 795-801, doi: 10.1109/IMCEC59810.2024.10575432.

Authors' Profiles

Zhadyra Avkurova: Lecturer of Department of AI Technology at NAO Karaganda Industrial University and PhD student at L. N. Gumilyov Eurasian National University, Kazakhstan. Major research interests: information technology, information security, cybersecurity, APT-attacks detection, fuzzy logic, AI/ML-based security systems.

Sergiy Gnatyuk: Acting Vice-Rector for research at National Aviation University (Ukraine), DSc, PhD, Professor. In 2007 he received MSc degree in information security from NAU. He received PhD in Eng degree in cyber-security from NAU in 2011 and DSc in 2017. In 2014 he received Associate Professor degree as well as in 2021 he received Full Professor degree. Major research interests: Cryptography, Quantum Key Distribution, Network & Internet Security, Information Security Incident Management, Cybersecurity & CIIP, AI/ML-based Security Systems.

Bayan Abduraimova: PhD, Associate Professor of the Department of Computer Science at L. N. Gumilyov Eurasian National University, Kazakhstan. Major research interests: Cryptography Systems, Network & Internet Security, Cybersecurity of Critical Infrastructures.

Kaiyrbek Makulov: PhD, Associate Professor, Department of Computer Science, Yessenov Univeristy, Aktau, Kazakhstan. Major research interests: Economic and Information Security, Cybersecurity and AI, Information Systems in Educarion.

How to cite this paper: Zhadyra Avkurova, Sergiy Gnatyuk, Bayan Abduraimova, Kaiyrbek Makulov, "Targeted Attacks Detection and Security Intruders Identification in the Cyber Space", International Journal of Computer Network and Information Security(IJCNIS), Vol.16, No.4, pp.144-153, 2024. DOI:10.5815/ijcnis.2024.04.10