Complex method to calculate objective assessments of information systems protection to improve expert assessments reliability

Complex method to calculate objective assessments of information systems protection to improve expert assessments reliability

A Zh Abdenov[1], V A Trushin[2] and G A Abdenova [3]

[1] Novosibirsk State Technical University, Novosibirsk, Russia
[2] Novosibirsk State Technical University, Novosibirsk, Russia
[3] L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

[1] e-mail: amirlan21@gmail.com
[2] e-mail: rastr89@mail.ru

Abstract. The paper considers the questions of filling the relevant SIEM nodes based on calculations of objective assessments in order to improve the reliability of subjective expert assessments. The proposed methodology is necessary for the most accurate security risk assessment of information systems. This technique is also intended for the purpose of establishing real-time operational information protection in the enterprise information systems. Risk calculations are based on objective estimates of the adverse events implementation probabilities, predictions of the damage magnitude from information security violations. Calculations of objective assessments are necessary to increase the reliability of the proposed expert assessments.

Keywords: risk assessment, information security, information system, the objective assessment of the damage, SIEM nodes, information security violations.

1. Introduction

Development of a system with intelligent services for the protection of confidential information (CI) requires the successful implementation of various activities to protect information resources (IR) in computer systems (CS). These activities involve the solution of a number of tasks, in particular, the creation of a system for monitoring security threats. Monitoring systems implement the current and a posteriori approaches to information protection, and the main goal of its creation is to reduce the number of incidents affecting the IR in CS to a minimum level of risk and to minimize the resulting damage. In this paper, we consider an incident to be any illegal, unresolved, unfavorable events (UE) that occur in the company's information systems (IS).

Today, many enterprises widely use various computer systems to search for customers, through employment of software products with intellectual properties, to conclude contracts and sell goods and services. Each of these companies is engaged in the collection, processing, storage and exchange of business information. Company's business processes generate metadata that is stored in various flexible and effective forms in order to ensure rapid data retrieval, use and movement between various enterprise entities and external partners.

The convenience of using the Internet for information exchange makes not only the information that the company sends via the Network, but also the information resources (IR) in any information system (IS) that the company has at its disposal and which are connected to the global public Network, vulnerable. At the same time, the importance of ensuring the safety of IR is understood by the management of every enterprise, organization, as well as private persons, notebook owners. Therefore, the issues of building an information security (IS) system in terms of protection from the most likely attacker, who has sufficient means for penetrating the information structure of the enterprise, are very relevant. Designed IS systems should be reliable with respect to the level of importance, secrecy and criticality of the protected information. At the same time, the cost of reliable protection of IR in the IS

should not exceed the possible damage from the security violation of the protected information.

One of the most promising and effective areas in the development of a security threat monitoring system is currently considered to be SIEM-systems (Security Information and Event Management), which allows for management of information and security events. The main purpose of the SIEM-system construction and operation is to significantly increase the level of information security (IFS) in the information infrastructure by providing the ability to manipulate security information in real-time and to actively manage security events [1, 2]. It is assumed that the active management of statistical information of the past, present and forecasting nature with respect to security incidents and events is based on automatic mechanisms using accumulated information on the history of the analyzed events and the forecast of future events, as well as on automatic adjustment of the monitoring parameters of events to the current state of the protected system [2, 3].

Discussion of issues arising in the development of SIEM-systems for service information infrastructures is carried out in the MASSIF project (Management of Security Information and Events in Service Infrastructure) of European Union's Seventh Framework Program for Research and Technological Development [1]. We assume that one of the issues discussed within the framework of the MASSIF project should be the questions of filling and continuously improving the individual SIEM system nodes with the ability to solve problems of incident classification, forecasting, filtering, and optimizing the use of IR protection in CS [4-8]. The main initial data used by the SIEM system to solve these tasks are records of various audit logs (logs), logging events in the information infrastructure, called "security events". Data on events reflect such actions by users and programs that may have an impact on the IS.

Next generation SIEM-systems aside from the standard node functions should include extensions such as, the analysis of events, incidents and their consequences, decision-making and visualization of information. We will reveal some mechanisms of functional nodes filling according to the hierarchy levels of the SIEM-system: normalization means the reduction of the formats of log entries collected from various sources to a single internal format, which will then be used for their storage and subsequent processing [2, 3]; filtration of the security events is to adjust the current assessments of the security state of IR in the CS and calculate these assessments for a future time interval [5, 7]; prioritization determines the significance and criticality of security events based on the rules defined in the system [1, 8, 9]; analysis of events, incidents and their consequences include procedures for modeling events, attacks and their consequences, vulnerability analysis, risk assessments, forecasting and filtering of NS assessments and incidents [10, 11]; The Decision Support System (DSS) determines the development of measures to optimize and reconfigure security tools to prevent attacks on the CS [2, 3]; Generation of reports and warnings means the formation, transmission, display and (or) printing of the results of operation [1–4].

In this paper, we will propose and disclose some elements of filling of individual SIEM system nodes with algorithms for solving risk calculation problems on the basis of objective assessments, as well as estimating the predictions and filtering of the number of UE, the amount of damage to IR in the CS [5, 6, 10, 11].

At the same time, there arises the problem of estimating the state of possible channels of information leakage and the corresponding assessment of the damage that may occur in the event of information leakage or any other security breach, and likelihood of causing such damage. To determine the adequacy of the cost of the protection system, it is necessary to compare the extent of damage and the likelihood of its application with the costs of providing protection. Unfortunately, it is very difficult and troublesome to obtain an objective real value of each IR in the IS due to a lot of uncertainties, as well as adequate automated and automatic control and accounting of UE in computer systems. Therefore, only expert assessments are often used.

In solving practical problems of information protection, the quantitative assessment of its vulnerability (inability to deter an attack) is of paramount importance. Unauthorized access (UA) to information in the automated and automatic system (AAS) is possible not only through direct access to databases (object), but also in many other ways that do not require such access. In this case, the intentional actions of intruders are the main danger. The impact of random factors in itself does not lead to unauthorized access, it only contributes to the emergence of channels of unauthorized information

retrieval (CUIR), which can be exploited by an attacker.
In this paper, for the purposes of protecting IR in enterprise IS, we will review various approaches to risk management. When calculating risks in IS, we will base not only on well-known expert [3], but also objective assessments of the probabilities of the number of NS implementations. We will also perform prediction assessments of the damage magnitude from IR security violations.
A risk-based approach to assessing damage from attacks by violators and selecting measures to minimize it has been termed risk management. Risk management refers to a full range of successive processes that comply with existing international standards and practices for enterprise risk management [4]. In existing risk management techniques, their identification is carried out by various methods, such as team "brainstorming", drawing up sequences of processes and event trees, analyzing the system architecture, operational modeling, scenario analysis, HAZOP studies [5, 6]. At the present stage, one of the methods of risk management is well tested on expert assessments [6]. Following the approach of Kumamoto and Henley [9], as well as the ideas reflected in [5], we adopt a narrower practical direction of the formal definition of parameters that characterizes the risks of IR security breach. First, it is necessary to classify the types of security infringement attacks. Secondly, it is necessary to accumulate observational data on each type of attack that characterizes the number of violations of each type of attack and the cost estimate of damage from IR security violations.
Methodical assessment of damage to IS caused by UE depends on the risk assessment of UE probability in IS. Evaluation of the objective probabilities of the UE occurrence is one of the important tasks in the algorithm for calculating risk assessments. Note also that the use of these objective assessments to improve the effectiveness of the calculation of risk assessments in enterprise IS is an important task in the methodology for calculating assessments of potential damage from attacks by violators.

2. The analysis and classification of possible UAs to company's information in different zones based on territory and probability

The following methods and methods of committing computer crimes are known: the introduction of unauthorized data; manipulation of unauthorized data; illegal use of files; creation of unauthorized files; overriding of internal control mechanisms; unauthorized destruction and modification of data on exit; unauthorized manipulation of computer programs or documentation; unauthorized manipulation of data processing; manipulation of errors, system failures; unauthorized use of passwords and codes; unauthorized transmission and interception of communications; theft of computer equipment, software; intentional destruction or damage of equipment, software or data, etc.
Territorially potentially possible UAs to company information can take place in different zones [11]:
– external uncontrolled zone of the company;
– zone within the company's controlled territory;
– AAS zone;
– AAS resource zone;
– the database zone.

Furthermore, UAs require following events to take place simultaneously in order to receive information:

– the attacker must gain access to the corresponding zone;
– the necessary CUIR should be available in the accessed zone;
– the manifested CUIR should be accessible to an attacker of the appropriate category (here, it means that the attacker must have the appropriate level of education, skills and equipment to be able to take advantage of CUIR);
– the CUIR at the time of access by an attacker should have protected information.

2.1. External uncontrolled zone.

2.1.1. Calculation of the UE objective probability occurrences

Consider first zone: external uncontrolled zone. All computer crimes can be divided into three classes:

the interception of information; unauthorized access; «data manipulation».
Consider the first class. It:
A) electromagnetic interception, for example, registration of radiation generated by a processor, printer, monitor;
B) direct interception, for example, direct connection to data transmission channels.
Consider the second class:
A) illegal access to the line of the legitimate user;
B) The kind of crime that is called «boarding» is when «computer pirates» get into other people's information systems by guessing their password.
Consider the third class («data manipulation»):
A) type of crime – code substitution, for example, with different code variations;
B) type «Trojan horse» – a secret introduction to someone else's program of such commands that gives attacker ability to execute new, unplanned by the user, software functions;
C) computer viruses – act on the principle, for example, erase all the data of this program, go to the next one and do the same. They have the property of moving through communication networks from one system to another, spreading as a viral disease;
D) the kind of «logical and temporary bomb» – the effect of falsely entered commands on programs under certain conditions and time, etc.
Now we will make a list of the significant types of UEs arising in the IS, leading to a deterioration in the system performance of the computer system (CS). Let this list be a set of various UEs $\{O_1, O_2, \ldots, O_m\}$. Let us distinguish from this set some significant subset of UEs, leading to a perceptible violation of the safety of IR in the CS. We denote this subset as $O = \{O_{i_1}, O_{i_2}, \ldots, O_{i_m}\}$, for example O_{i_1} - the number of UA with respect to the launch failures of individual CS nodes; O_{i_2} - the number of UEs with respect to the incorrect information applied to a specific information process and data processing, etc.
After building a subset O we proceed to the subset's elements properties analysis on the basis of the quantitative indicators of the UEs and the magnitude of the damage that has occurred in the past. Let $O = \{O_{i_1}, O_{i_2}, \ldots, O_{i_m}\}$ – the set of all significant UE, leading to a decrease in the system efficiency of the CS. The mathematical expectation of the damage caused by i's UE in time ΔT (For example, 1 month, 1 half year, 1 year, etc.) can be represented by the formula:

$$e(O_i, \Delta T) = M[e(O_i) \cdot f_i], i = \overline{1, m}, \quad (1)$$

where $e(O_i)$ – a random amount of already occurred UA damage in case of a single UA event; f_i – random amount of UE of type i in time ΔT; m – the total number of all types of already occurred UEs.
If UEs don't have consequences, damage from each UE is independent, then

$$e(O_i, \Delta T) = M[e(O_i)] \cdot M[f_i], \quad i = \overline{1, m}, \quad (2)$$

the damage for the whole set of significant UE will be determined by using the formula

$$E(O, \Delta T) = \sum_{i=1}^{m} M[e(O_i)] \cdot M[f_i]. \quad (3)$$

Algorithm 1

Step 1.1. For simplicity, we consider only one type of UE, for example, we set a specific value i = 1. Further, the quantitative indicators of the UE, for example, μ years we will compile in the table 1:

Table 1. Quantitative indicators of UA that occurred during the last μ years by months (or several quarters, half-year or more rare occurrences of UA).

t	1	2	3	4	5	6	7	8	9	10	11	12
1	$f_1^{(1)}$	$f_2^{(1)}$	$f_3^{(1)}$	$f_4^{(1)}$	$f_5^{(1)}$	$f_6^{(1)}$	$f_7^{(1)}$	$f_{18}^{(1)}$	$f_9^{(1)}$	$f_{10}^{(1)}$	$f_{11}^{(1)}$	$f_{12}^{(1)}$
2	$f_1^{(2)}$	$f_2^{(2)}$	$f_3^{(2)}$	$f_4^{(2)}$	$f_5^{(2)}$	$f_6^{(2)}$	$f_7^{(2)}$	$f_{18}^{(2)}$	$f_9^{(2)}$	$f_{10}^{(2)}$	$f_{11}^{(2)}$	$f_{12}^{(2)}$
⋮	…	…	…	…	…	…	…	…	…	…	…	…
μ	$f_1^{(\mu)}$	$f_2^{(\mu)}$	$f_3^{(\mu)}$	$f_4^{(\mu)}$	$f_5^{(\mu)}$	$f_6^{(\mu)}$	$f_7^{(\mu)}$	$f_8^{(\mu)}$	$f_9^{(\mu)}$	$f_{10}^{(\mu)}$	$f_{11}^{(\mu)}$	$f_{12}^{(\mu)}$

Step 1.2. From the data in Table 1, using formula (4) we can calculate one line of columns averages for monthly (quarterly, semiannual, annual relatively rare occurrences of UA) quantitative values of UA (see Table 2):

$$f_t^{ycp} = \sum_{i=1}^{\mu} f_t^{(i)} / \mu, \ \ t = \overline{1, 12}. \tag{4}$$

Table 2. The average line quantitative occurred UE.

t	1	2	3	4	5	6	7	8	9	10	11	12
f_t^{ycp}	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}

Similar tables desirable to construct in relation to other types of UE that have already occurred within certain period of time, for example one month.

Step 1.3. With respect to the averaged data (see Table 2) it is possible to construct a discrete linear stationary stochastic model in the form of state-space (SS) [9, 10] (according to the procedure described in [9]) of the form:

$$x(t+1) = a \cdot x(t) + b \cdot u(t) + w(t), \quad x(0) = \bar{x}_0, \tag{5}$$

$$f^{ycp}(t+1) = x(t+1) + v(t+1), \quad t = \overline{0, N-1}. \tag{6}$$

where $x(t)$ is the true number of the UA that took place during the month t (could be also a few quarters, a few semesters, as long as a few years for the rare type of UA); $u(t)$ is the externally observable control action in the form of fixed type of UE at the time t ; $w(t)$ is the white Gaussian unobserved effects at a time t with mathematical expectation of zero and unknown variance Q ; $x(0)$ is the number of UE at the beginning time $t = 0$ with mathematical expectation $\bar{x}_0$ and unknown variance $P(0)$; a, b are the unknown coefficients in the model dynamics (5); t is the month of the year (or number of quarters, semesters or years); $N = 12$ is the the number of months in a year; $f^{ycp}(t)$ is the observed random number of UE in a month (Data from the company's record book); $v(t)$ is the white Gaussian sequence of observation errors on the number UE within defined period of time, for example, a month with zero mean and unknown variance R .

In this step, it is necessary to evaluate all of the variances associated with noise of the dynamics model $\hat{Q}$, initial state value of the noise $\hat{P}(0)$, with the noise of the measuring system $\hat{R}$ according to the

formulas set forth in [12].

Step 1.4. Assessments for the coefficients in the model dynamics (5) can be calculated based on the method of least squares (OLS) [9, 10].

Step 1.5. The constructed model (5), (6) will provide the most reliable estimates of the number UE (in real time, using Kalman filter) with respect to each month as a filter estimates [10] for a subsequent, for example, (μ +1) year. The resulting filter estimates should be rounded to the nearest whole number.

Step 1.6. Filtering assessments will calculate the objective probability assessment of UE occurrences. For example, we propose the following procedure for calculating the probability for a particular type of UE.

Suppose we are interested in the probability of the UA, for example, in each of the previous months of the μ-th year. In order to calculate this number we need to count total sum (for the average amount of) of UE filtering assessments during the whole μ of the year ($F^{(\mu)}$), then filtration assessment of UE quantity during each month ($f^{(\mu)}(t)$) is divided by the total sum of UE filtration values of the quantity UE ($F^{(\mu)}$) for one μ is year determined by the formula:

$$p^{(\mu)}(t) = f^{(\mu)}(t)/F^{(\mu)}, \quad t = \overline{1,12}, \tag{7}$$

where $p_t^{(\mu)} = p^{(\mu)}(t)$ is the objective probability of a particular type of UE for each month μ – year and of all 12 months (see. Table 3).

Table 3. Objective probabilities of particular type of UE occurrences within the μ – th year

t	1	2	3	4	5	6	7	8	9	10	11	12
$p_t^{(\mu)}$	$p_1^{(\mu)}$	$p_2^{(\mu)}$	$p_3^{(\mu)}$	$p_4^{(\mu)}$	$p_5^{(\mu)}$	$p_6^{(\mu)}$	$p_7^{(\mu)}$	$p_8^{(\mu)}$	$p_9^{(\mu)}$	$p_{10}^{(\mu)}$	$p_{11}^{(\mu)}$	$p_{12}^{(\mu)}$

In addition, for μ –th year we will have:

$$\sum_{t=1}^{12} p^{(\mu)}(t) = 1, \quad \mu = 4, 5, \ldots.$$

The efficiency of the algorithm 1 is illustrated in the test example in [12].

2.1.2. An objective valuation of the damage prediction of security breaches IR

Algorithm 2

Step 2.1. Let $O = \{O_i, i = \overline{1,m}\}$ – many types of UE, leading to the security breaches of IR. In section 2.1.1. this work has been proposed assessment procedure for calculating monthly (or quarterly or semi-annual, etc.) objective probability of the number of certain type of attacks on IR in the company's IS. Suppose that company's information security (IFS) department possesses statistics regarding the monthly damage assessment, which corresponds to a monthly amount of particular i is the type of information security breaches, i.e. data Table 1 values correspond to the values of Table 4.

Table 4. Quantitative monthly (or quarterly or half-yearly, etc.) indicators damage from IFS violations depending on i –th type attack

t	1	2	3	4	5	6	7	8	9	10	11	12
1	$S_1^{(1)}$	$S_2^{(1)}$	$S_3^{(1)}$	$S_4^{(1)}$	$S_5^{(1)}$	$S_6^{(1)}$	$S_7^{(1)}$	$S_8^{(1)}$	$S_9^{(1)}$	$S_{10}^{(1)}$	$S_{11}^{(1)}$	$S_{12}^{(1)}$
⋮	…	…	…	…	…	…	…	…	…	…	…	…

μ	$S_1^{(\mu)}$	$S_2^{(\mu)}$	$S_3^{(\mu)}$	$S_4^{(\mu)}$	$S_5^{(\mu)}$	$S_6^{(\mu)}$	$S_7^{(\mu)}$	$S_8^{(\mu)}$	$S_9^{(\mu)}$	$S_{10}^{(\mu)}$	$S_{11}^{(\mu)}$	$S_{12}^{(\mu)}$

Table 4 $\left\{S_t^{(i)},\ t=\overline{1,12}, i=\overline{1,\mu}\right\}$ – outputs that characterize the monthly (quarterly, semi-annual) amount of damage caused by the UE in the IS of the company during i –th sampling period.

Step 2.2. Note that company's performance losses not always directly proportional to the number of UE which took place. Nevertheless, based on the data of Table 4 we can construct a linear discrete model in the form of SS, which will correspond to the averages of data columns in Table 4. The elements of the averaged lines are calculated using equation (8) (see Table 5):

$$s_t^{(y)} = (\sum_{i=1}^{\mu} S_t^{(i)})/\mu,\ t=\overline{1,12}. \quad (8)$$

Table 5. The average monthly line (quarterly, semiannual, etc. with respect to a time interval) quantitative value of damage caused to the company's IR.

t	1	2	3	4	5	6	7	8	9	10	11	12
$s_t^{(y)}$	$s_1^{(y)}$	$s_2^{(y)}$	$s_3^{(y)}$	$s_4^{(y)}$	$s_5^{(y)}$	$s_6^{(y)}$	$s_7^{(y)}$	$s_8^{(y)}$	$s_9^{(y)}$	$s_{10}^{(y)}$	$s_{11}^{(y)}$	$s_{12}^{(y)}$

Step 2.3. Based on the data line $\left\{s_t^{(y)},\ t=\overline{1,12}\right\}$ the algorithm described in [10, 11], we can construct a linear stochastic model in the SS form of a stationary type

$$s(t+1)=\hat{c}\cdot s(t)+\hat{d}\cdot+w(t),\quad s(0)=s_0,\quad t=\overline{0,11}, \quad (9)$$

$$s^y(t+1)=s(t+1)+v(t+1),\quad t=\overline{0,11}. \quad (10)$$

In this case, first on the basis of data in Table 5 let's calculate line noise variance estimation model view (9), (10), in particular assessment of variances evaluation Q, R, $P(0)$ [12].

Step 2.4. Furthermore, let's calculate dynamics model coefficients (9) using OLS [9, 10].

Step 2.5. Suppose we have data of quantitative observations of damage caused to the company's IR in $(\mu+1)$ year (see. Table 6) and on the basis of the Kalman filter equations [10] we obtain line filter estimates (see. Table 7).

Table 6. Monthly (quarterly, half-yearly on several years) quantitative measures of damage from IFS violations depending on i -th type of attack in $(\mu+1)$ year.

t	1	2	3	4	5	6	7	8	9	10	11	12
$S_t^{(\mu)}$	$S_1^{(\mu)}$	$S_2^{(\mu)}$	$S_3^{(\mu)}$	$S_4^{(\mu)}$	$S_5^{(\mu)}$	$S_6^{(\mu)}$	$S_7^{(\mu)}$	$S_8^{(\mu)}$	$S_9^{(\mu)}$	$S_{10}^{(\mu)}$	$S_{11}^{(\mu)}$	$S_{12}^{(\mu)}$

Using the equations of the Kalman filter and the data in Table 6, we obtain the sequence estimates of filtering $\left\{\hat{s}(t\mid t), t=\overline{1,12}\right\}$, relatively to more reliable data of monthly quantitative damage (see. Table 7).

Table 7. Monthly (or quarterly or half-yearly, etc. relative to a few years) quantitative filtration damage assessments in $(\mu+1)$ -th year.

t	1	2	3	4	5	6	7	8	9	10	11	12
$\hat{s}(t\mid t)$	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}

Step 2.6. Using rounded to the nearest whole number data filtering estimates on quantitative indicators, UE instances during μ year monthly (or quarterly or half-yearly, etc. relative to a few years), and the data in Table 7 with respect to the filtering estimates of quantitative indicators of the damage inflicted on the company's IR in μ year, you can get the average damage caused by a single UE $\{e(t),\ t=\overline{1,12}\}$. To do this, let's divide the data table row 7 divided into corresponding, rounded to the nearest whole number of UE instances. Calculated data can be summarized in Table 8.

Table 8. The average damage from a single UE instance.

t	1	2	3	4	5	6	7	8	9	10	11	12
$e(t)$	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	e_{12}

Step 2.7. Predicting the UE quantity of the (f_i^{pre}) i-th type by using appropriate model in the form of SS and corresponding averaged damage from the single UE according to Table 8, we can estimate the prediction of the damage to the company in t - the month (or quarter or half-year, and so on a few years) $(\mu+1)$ (current) year.
Operability of the 2 algorithm is illustrated in the test example in [12].

2.1.3. Controls the security group

Security measures (security controls), used in the company for IR and IS protection, can be divided into three main groups: technical, operational and management [6]. Groups, in turn, are divided into families. Listed these security measures in the standard are listed in [13].
The group management control includes security measures for IS, which focus on risk management and information security. The group consists of five families of controls [6, 13]. The group operating controls include security measures for IS, which is primarily implemented and executed by people. The group consists of nine families of controls [6, 13]. The group of technical controls include security measures for the IS, which are primarily implemented and executed through the steps in hardware, software, or firmware components of the system. The group consists of four sets of controls [6, 13].
Experts who conduct assessment may come from different professional training, such as: technical, financial, engineering, and management, with their own individual perceptions, attitudes and motivations in determining the amount of the realized loss from UE. Therefore, before calculating the amount of risk in the company's IS it is vital for all experts to get familiar with all objective assessments and predictions regarding the filtering of monthly (or quarterly or semi-annual, etc. relative to a few years) UE quantities and the damage caused by UE which allow experts to offer the most realistic expert estimates.
Procedure risk assessment based on expert judgment is comprised of six steps [6].
Now consider the risk calculation methodology according to the procedure set forth above.

2.2. Methods of calculating

2.2.1. Example 1. Methods of calculating the objective quantity prediction evaluations UE for external uncontrolled zone

Presumably, in the case of uncontrolled outer zone relative to the UE quantity of UA in IR CS, even in six months, in the range 0-5 of UE. In order to construct a discrete linear stochastic model in the form of SS with two unknown coefficients, you must have 4–5 times more observations than the number of unknown coefficients. In this regard, we must stretch the whole time quantization interval with respect to the collection of a sufficient number of UE, for four or five years, provided that the sample size of UE is no less than eight or ten (N = 10). Under these assumptions, the time interval between the

observations will take at least six months.

Algorithm 1.
1. Suppose that one implementation of a time series (TS) observation for 5 years is $N=10$. In order to improve the data quality with respect to the UE quantitative analysis, we take the number of bands equal to instances, for example, $k=6$ and compile these data in Table 9 [14]: All data in Table 9 are modeled through the use of a uniform distribution law using the restriction on the number UE no more than 5, and using the procedure of rounding to a whole number.

Table 9. Modeled 6 TS values with sample volume $N=10$ in relation to quantitative indicators of UE.

No	1st year		2nd year		3rd year		4th year		5th year	
	1	2	3	4	5	6	7	8	9	10
$y_1(i)$	3	2	4	2	2	4	2	2	5	0
$y_2(i)$	2	2	0	2	5	2	4	2	2	2
$y_3(i)$	2	2	4	3	4	5	2	4	4	4
$y_4(i)$	4	2	1	2	2	4	2	0	0	2
$y_5(i)$	2	2	1	4	1	2	1	1	2	1
$y_6(i)$	2	2	2	2	4	2	1	3	2	2

2. For $k=6$ instances calculate the average quantities of UE quantities in the columns. Then we get average TS that corresponds to the number of UE for every six months during 5-year time period [14]. The data are summarized in Table 10.

Table 10. Averaged TS, with sample volume $N=10$, characterizing quantitative indicators of UE.

No	1st year		2nd year		3rd year		4th year		5th year	
	1	2	3	4	5	6	7	8	9	10
$\bar{y}(i)$	2.500	2.0000	2.0000	2.5000	3.0000	3.1667	2.0000	2.0000	2.5000	1.8333

3. Building a linear discrete stationary stochastic model in the SS form based on the averaged data. Model coefficients are calculated by OLS acquired the following values: a =0.2233; b =1.7958 [10].
4. On the basis of the recurrence formulas given in the paper [12], we calculate dynamics model noise variance Q, the initial state $P(1)$ and the noise of the measurement system R based on an average TS data $\{\bar{y}(i), \overline{1, N}\}$. After calculation, variances took the following values:
Q = 0.0982; P (1) = 0.0982; R = 0.1013.

5. Calculated variables in the preceding two paragraphs have allowed based on the Kalman filter equations to calculate the estimates and predictions, respectively, with respect to evaluation of the filtration, the amount of UE with anticipation intervals for half a year, during the preceding year,

Table 11. The calculated assessment predictions and the corresponding filtering estimates from the previous year.

No.	1st year		2nd year		3rd year		4th year		5th year	
	1	2	3	4	5	6th	7th	8	9	10
$\hat{y}(i \mid i-1)$	2.50	2.354	2.282	2.274	2.329	2.391	2.417	2.289	2.274	2.329
$\hat{y}(i \mid i)$	2.50	2.176	2.14	2,388	2.667	2.783	2.207	2.143	2,388	2.079
$z(i)$	2.50	2.000	2.000	2.500	3,000	3.167	2.000	2.000	2.500	1.833

Figure 1 shows plots reflecting changes in the number of UE based on TS values in relation to real (z), predicted (XP) and filtration ratings (XF)

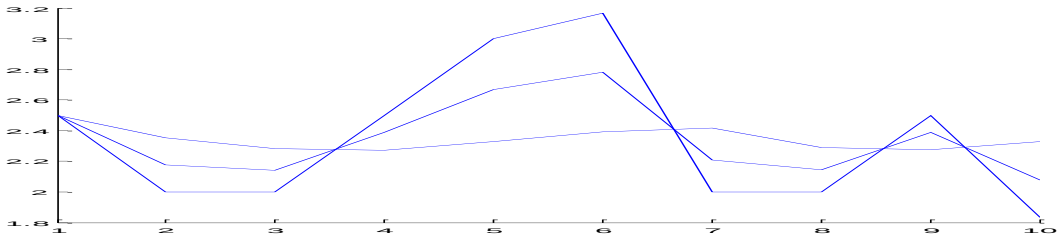

Figure 1. Graphs constructed based on real (z), the predicted (XP) and filtration (XF) estimates.

6. Calculate filtration assessments rounded to the nearest whole number. The calculation results are summarized in Table 12

Table 12. The calculated assessment predictions and the corresponding filtering estimates from the previous year.

No.	1st year		2nd year		3rd year		4th year		5th year	
	1	2	3	4	5	6th	7th	8	9	10
$\hat{y}(i \mid i)$	2.50	2.176	2.14	2,388	2.667	2.783	2.207	2.143	2,388	2.079
xfc	3	2	2	2	3	3	2	2	2	2
pf	0.13	0.087	0.087	0.087	0.13	0.13	0.087	0.087	0.087	0.087

where xfc - n – vector of filtering estimates rounded to the nearest whole number (see. Table 12).

7. Calculate the objective probability estimates regarding UE instances in the previous year based on the algorithm 1, step 1.6 pf - n – vector of objective semi-annual probabilities (see the last line of Table 12).

8. We use filtering estimates to predict the number of incidents occurred in the first half of the $(\mu+1)$-th year. For this purpose we calculate the equation coefficients of linear approximation μ-th occurence on the basis of filtration estimates in the form of TS with the sample size $N=10$: XF = [2.5000 2.1755 2.1396 2.3878 2.6674 2.7825 2.2067 2.1430 2.3882 2.0790]. As a result of the calculations, we obtain the following assessments: a=0.2385; b = 1.7631.

9. Based on the Kalman filter equations, we obtain the following prediction estimate 2.3541. Further on the basis of the next observation $z^{(i+1)}(1)=5$ it is possible to calculate

the filtration estimate $XF^{(i+1)}(1) = 3.6887.$ Next, we can calculate the prediction estimate for the next six months: $XP^{(i+1)}(2) = 2.6195$. Based on the corresponding observation $z^{(i+1)}(2) = 2$ we can calculate the filtration rating $XF^{(i+1)}(2) = 2.307$ and etc.

2.2.2. Example 2. The method of calculating the objective prediction valuation of damage magnitude from security breaches IR

Algorithm 2

1. Let $O = \{O_i, i = \overline{1, m}\}$ - be many types of UE, leading to a violation of the security of the IR, of which we choose one significant type. In clause 2.1.1. This work proposed a procedure for calculating the estimation of the semi-annual objective probability of the number of violations of a certain type of attack on the IR in the enterprise's IS. Suppose that the company's IFS department has statistics on a semi-annual average damage estimate, depending on the occurrence of a single incident, i.e. the values of the data in Table 1 correspond to the values of the data in Table 4, provided that the average loss from the implementation of a single damage is $\bar{s} = 21$ c.u. (conventional units). Then table 4 will correspond to the table with the following values.

Table 13. Modeled $k = 6$ TS with sample size $N = 10$ in the form of data related to quantitative indicators of damage corresponding to the number of UE from Table 9.

No	1st year		2nd year		3rd year		4th year		5th year	
	1	2	3	4	5	6	7	8	9	10
$S_1^{(i)}$	63	42	84	42	42	84	42	42	105	0
$S_2^{(i)}$	42	42	0	42	105	42	84	42	42	42
$S_3^{(i)}$	42	42	84	63	84	105	42	84	84	84
$S_4^{(i)}$	84	42	21	42	42	84	42	0	0	42
$S_5^{(i)}$	42	42	21	84	21	42	21	21	42	21
$S_6^{(i)}$	42	42	42	42	84	42	21	63	42	42

Table 14

The averaged 6-month line relative to 5 years and 6 occurrences of the quantitative indicators of damage inflicted by the enterprise's IR based on the data in Table 13

t	1	2	3	4	5	6	7	8	9	10
$s_i^{(y)}$	70,83	42	42	70,83	63	66,5	42	42	70,83	42

2. Based on the data string $\{s_i^{(y)},\ i = \overline{1, 10}\}$ (see Table 14), using the algorithms described in [7, 12], we can construct a discrete linear stochastic stationary model in the form of SS view (9), (10).

We calculate the coefficients of the dynamics model (9): $\hat{c} = 0.1736, \hat{d} = 45.0051.$

3. Based on the data line of Table 14, with the sample size $N = 10$, estimates of the noise variances of the model of the form (9), (10) are calculated. The calculations gave the following variance estimates: Q = 88.4390; R = 56.7539; P (0) = 88.4390.

4. Suppose that we have the initial data of observations of quantitative indicators of damage inflicted by the company's IR in $(\mu + 1)$ year consistently for two semesters with the initial condition in the

form of a filtration estimate XF (1 | 1) = XF (1) = z (1) = 42 and observations for the two half-year z (2), z (3) and on the basis of the Kalman filter equations. Based on the Kalman filter equations, we first obtain a semi-annual estimate of the damage prediction XP (2|1) = 52.2963 and, taking into account the actual damage monitoring z (2) = 45, we calculate the corresponding damage filtration estimate XF (2 | 2) = 47.8006.

For the second half of the year, we take the initial condition XF (2 | 2) = 47.8006. Using this initial condition, we can calculate the prediction estimate for the second half of the year XP (3|2) = 53.3033. Next, having received at the time $t=3$, observation z (3) = 54, we can calculate the filtration estimate XF (3|3) = 53.7326, which, as we note, corrects the prediction estimate based on objective observation.

Table 15. Half-yearly quantitative indicators of damage from violations of the IFS, depending on i -th type of attack in $(\mu+1)$ year.

t	1	2	3	4	5	6	7	8	9	10	11	12
$S_t^{(\mu)}$	$S_1^{(\mu)}$	$S_2^{(\mu)}$	$S_3^{(\mu)}$	$S_4^{(\mu)}$	$S_5^{(\mu)}$	$S_6^{(\mu)}$	$S_7^{(\mu)}$	$S_8^{(\mu)}$	$S_9^{(\mu)}$	$S_{10}^{(\mu)}$	$S_{11}^{(\mu)}$	$S_{12}^{(\mu)}$

5. Using the Kalman filter equations and the data in Table 5, we obtain a sequence of filter estimates $\{\hat{s}(t\,|\,t), t=\overline{1,12}\}$, with respect to semiannual more reliable quantitative indices of the damage (Table 15).

Table 16. Semiannual quantitative indicators of filtering assessment of damage in $(\mu+1)$ year.

t	1	2	3	4	5	6	7	8	9	10	11	12
$\hat{s}(t\,|\,t)$	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}

6. Using rounded to the nearest whole number filtering data relatively to the quantitative indicators of occurred UE within $(\mu+1)$ year by month, and the data of Table 7 regarding the filtration estimates as quantitative indicators of damage inflicted on the enterprise's IR in $(\mu+1)$ year, it is possible to obtain the average damage caused by a single instance of UE $\{e(t),\ t=\overline{1,12}\}$. To do this, it is necessary to divide lines of Table 7 by corresponding rounded to the whole numbers of UE occurrences. The calculated data can be summarized in Table 17.

Table 17. The averaged damage inflicted by a single incident of UE.

t	1	2	3	4	5	6	7	8	9	10	11	12
$e(t)$	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	e_{12}

7. Predicting the number of UE (f_i^{pre}) i -th type with the help of the corresponding model in the form of a PS and the corresponding average damage according to the data of Table 17, one can obtain an estimate of the prediction of the amount of damage that will be done to the enterprise in t -th half year in relation to $(\mu+2)$ implementation.

The efficiency of the methodology according to Algorithm 2 is illustrated in the test example, as well as in [12].

2.3. Zone-controlled territory AAS zone, AAS resource zone, the database zone.

In this section, we will consider all the above-mentioned zones of the controlled territory, AAS zones, AAS resource zones, and the database zones. As already mentioned above computer crimes can be divided into three classes: the interception of information; unauthorized access; "data manipulation". All classes can be considered separately.
With respect to all the last four zones, using the methodology described in paragraphs 2.1.1. – 2.1.3., it is possible, based on the relevant observational data, to calculate estimates of predictions and assessments of UE quantity filterings, as well as appropriate estimates of damage that may be incurred, for example, by a particular company.
Now, let's assume that there are three ISs of the conditional company, on which we assess the risks: IS-1, IS-2, IS-3. After reviewing the results of objective assessments, the experts discovered certain vulnerabilities and threats related to different control families. (The initial data and numerical calculations are shown in [15]). It is required to carry out expert evaluations to assimilate the above-described methodology.
After ranking the IS according to the level of risk in descending order, the following calculated data were obtained (see Table 18).

Table 18. Ranks of information systems.

Information System	IS-2	IS-3	IS-1
Level of Risk	0,224	0,217	0,213

Table 18 shows that the IS-2 receives the highest level of risk. This means that the probability of implementing threats and damage to this system is greater than for other systems. Therefore, the leaders of the organization, first of all, need to pay attention to the security of IS-2.

2.4. Assessment of the vulnerability of information processed in AAS

In view of the above methodology for calculating the number of UEs and the corresponding assessments of the damage inflicted to IR in the IS, it is possible to derive a formula for assessing the vulnerability of information processed in AAS. To do this, you can enter, for example [11], the following indicators: $P^{(d)}{}_{ikl}$ - probability of access of the criminal of the k-th category to the l-th zone; $P^{(k)}{}_{ijl}$ - probability of the presence (manifestation) of the j-th CUIR in the l-th zone of the i-th component of the AAS; $P^{(n)}{}_{ijkL}$ - probability of access of the violator of the k-th category to the j-th CUIR in the i-th zone of the i-th component of the AAS, provided that the intruder accesses the l-zone; $P^{(n)}{}_{ijl}$ - probability of access of the violator of the k-th category to the j-th CUIR in the i-th zone of the i-th component of the AAC, provided that the intruder accesses the l-th zone;

$$P_{ijkl} = P^{(d)}{}_{ikl} * P^{(k)}{}_{ijl} * P^{(n)}{}_{ijkl} * P^{(n)}{}_{ijl}, \tag{11}$$

- The probability of UA in one component of AAS by a single intruder of one category via one type of CUIR, let's call the basic indicator of the vulnerability of information (from the point of view of the UA).
Taking into account (1) the expression for the basic indicator will look like:

$$P^{(k)}{}_{ijk} = 1 - \prod_{l=1}^{5}[1 - P_{ijkl}] = 1 - \prod_{l=1}^{5}[P^{(d)}{}_{ikl} P^{(k)}{}_{ijl} P^{(n)}{}_{ijkl} P^{(n)}{}_{ijl}]. \tag{12}$$

The basic vulnerability indicators calculated in this way are of limited practical importance. To solve the problems associated with the development and operation of information security systems, use generalized vulnerability indicators, generalized by an index (I, j, k) or by their combination. For example, {K *} is a subset of interest to us from the set of potential violators. Or {l *}, {J *} - is a subset of the components of AAS and CUIR. Then in general, the overall vulnerability index will be defined as:

$$P_{(I^*)(J^*)(K^*)} = 1 - \prod_{\forall i}[1 - P^{(b)}_{ijkl}] \prod_{\forall j}[1 - P^{(b)}_{ijkl}] \prod_{\forall k}[1 - P^{(b)}_{ijkl}]. \tag{13}$$

It is also necessary to consider the method of calculating the vulnerability index taking into account

the time interval at which the vulnerability is assessed. In this case, it should be borne in mind that the longer the time interval, the more opportunities for the intruder for malicious actions and the greater the probability of a change in the state of the AAS. It is possible to determine such time intervals (not reducible at a point) at which the processes associated with a breach of security are homogeneous. Let's call these intervals small. Such a small interval, in turn, can be divided into very small intervals, the vulnerability of information on each of which is determined independently of the others. The dependence on each of the allocated intervals will be the same because of the homogeneity of the ongoing vulnerability processes. Then, through P_t^m we denote the vulnerability index of interest to us at a point (on a very small interval), and through the same P indicator on a small interval. We obtain the following formula:

$$P = 1 - \prod_{i=1}^{n_t} [1 - P_t^m] \tag{14}$$

where t - is a variable index of very small intervals, which represent a small interval; n_t -the total number of very small intervals. This approach is valid if the conditions for violation of information security remain unchanged throughout the interval under consideration. In real AAS, these conditions can vary.

3. Summary and conclusions

The current practice of applied technical or economic research indicates that in order to achieve success with regard to the safety of IR in the company's IS, the researcher should be well versed in three areas: 1) technical and economic theory; 2) mathematical modeling, i.e. the art of formalizing the formulation of the problem, which consists in the ability to translate a problem from the language of problem-oriented to the language of abstract mathematical models; 3) the corresponding software. Therefore, in this work, a systematic presentation of mathematical methods and models of analysis of safety measures was given and was aimed at studying theoretical approaches and was aimed at developing practical skills in their development and application to the calculation of risks with respect to the information systems under investigation.

Risk management is based on data that must be captured, accumulated, analyzed, stored, processed for the purpose of assessing the potential damage from user errors and attacks on IRs in the company's IS, the choice of measures to minimize it, the calculation of prediction estimates and the filtering of all possible parameters and indicators associated with the IFS. In this paper, in particular, methods have been proposed that allow obtaining estimates of the objective probability in the possibility of UE occurrence, estimating the objective cost of damage from IR security breaches in the company's IS prediction and filtering of the damage value estimation corresponding to the quantitative indicators of the UE. All the basic calculations of the IFS indicators in the company's IS used the capabilities of a discrete linear stochastic stationary model in the form of a SS and the Kalman filter equations to obtain more reliable estimates of the state of the object under study.

The work concentrates those methods and models that are most often used for analyzing and developing solutions for the rational choice and distribution of security measures relative to IR, taking into account the influence of safety measures on other controls, etc. In the process of risk assessment, the interdependencies between security measures are taken into account, which makes it possible to prioritize implementation of security measures and develop an adequate risk management strategy.

References

[1] Kotenko I V Sayenko I B 2012 SIEM-systems for management of information and security events *Information protection. INSIDE* 5 pp 54–65
[2] Miller D R Harris Sh Harper A A Van-Dyke S 2011 Black Ch. Security Information and Event Management (SIEM) Implementation McGrawHill Companies p 430
[3] ISO / IEC 27005: 2008. Information technology. Security techniques. Information security risk management
[4] 2006 *Risk management: Implementation principles and inventories for risk management* (ENISA)

[5] Zapechnikov S V 2010 *Information Technology Security* 1 pp 21–27
[6] Chi-Chun Lo Wan-Jia Chen 2011 *Expert Systems with Applications* 39 pp 248–257
[7] Abdenov A Zh Zarkumova-Reichel R N 2015 *Issues of Information Protection* 1 pp 64–70
[8] Kumamoto H Henley E 1996 *Probabilistic risk assessment and management for engineers and scientists* (IEEE, New York) p 620
[9] Zarkumova-Reichel R N Abdenov A Zh 2012 *Forecasting the number of incidents in the enterprise's information security system using a dynamic model* Fundamental Research 6(2) pp 429–434
[10] Sinitsyn I N 2006 *Filters of Kalman and Pugachev: Textbook* (Moscow: Logos) p 640
[11] Meshcheryakov R V Praskurin G A 2004 *Theoretical basis of computer security Razdel 2* (Tomsk: Tomsk State University of Control Systems and Radioelectronics) pp 97–148
[12] Abdenov A Zh Trushin V A Abdenova G A Inozemtseva Yu A 2016 *Artificial intelligence and decision-making* 3 pp 87–99
[12] NIST SP 800-30 2012 *Guide for conducting Risk* (Broadway: Assessments National Institute of Standards and Technology)
[13] Klimova E G Platov G A Kilanova R V 2014 *Computational technologies* 19 pp 27–37
[14] Abdenov A Zh Belkin S A Zarkumova-Reichel R N 2014 *Methodics of risk assessment for information systems on the basis of expert assessments (*Novosibirsk: NSTU) p 71