A Markov model for optimization of information security remedies

A A Kasenov[1], E F Kustov[1], A A Magazev[1] and V F Tsyrulnik[1]

[1]Omsk State Technical University, pr. Mira, 11, 644050, Omsk, Russia

E-mail: magazev@omgtu.ru

Abstract. The problem of selecting an optimal collection of security remedies against a specified set of cyber threats is relevant as far as there is a wide range of different cyber security solutions in modern IT industry. The aim of this work is to formulate an optimization problem for selecting information security remedies using a Markov cyber threat model and to analyze the possibility of solving the problem by the method of sequential analysis of variants. The set of solutions for standard approaches is limited by admissible indicators of the economic damage, and the corresponding restriction by means of a functional and temporal characteristic of the model called its average lifetime is defined. The explicit analytical formula for the average lifetime of an information system expressed in terms of the original parameters of the model is obtained, these parameters being the probabilities of threat occurrence and probabilities of their eliminations by security remedies. The possibility of solving our optimization problem by the method of sequential analysis of variants is analyzed. The program in C++ is developed to experimentally compare the effectiveness of this method compared with the "brute-force" method.

1. Introduction

A continuous thorough analysis of the current cyber threats and vulnerabilities is required for effective solution of information security problems. It is necessary for the timely response to alarm events in information security systems often functioned in a given cyber threat space. Generally, the cyber threat analysis evaluates their occurrence probabilities (for a certain period of time) and the damage to information [1].

In the IT industry there are a considerable number of different solutions to ensure cyber security. Moreover, several different security remedies produced by various manufacturers can be used to eliminate the same cyber threat. Typically, these remedies vary widely in cost and have various abilities for preventing cyber threats. Thus, the problem of selecting some optimal set of remedies is relevant.

Mathematically, the problem statement of selecting an optimal subset from a given set of security remedies can be represented in various models. A review of modern popular approaches to the optimal subsets selection is found in the work [2]. A group of approaches based on theoretical models for the evaluation of investments in information security is highlighted in [3, 4, 5], and a series of approaches based on game theory is presented in [6, 7, 8].

The two problems of optimal selection of security remedies belonging to the class of nonlinear discrete optimization problems are formulated in the paper [9]. The authors indicate the classical linear programming methods not being applicable for solving these problems. There

are also difficulties in explicitly defining the objective function and the constraint set. In this paper, we formulate an alternative optimization problem using a Markov cyber threat model suggested in the work [10] and examined in detail in [11, 12]. Unlike the approach described in [9], we determine the constraint set of our optimization problem by functional and temporal characteristics of the model, rather than indicators of the economic damage. We also analyze the possibility of solving the given optimization problem by the method of sequential analysis of variants which takes into account existing features of the problem and allows us to achieve significant gain compared with the "brute-force" method.

2. Problem statement

Let us consider a model describing an information system affected by n independent external threats with probabilities $q_1, q_2, \ldots, q_n$. Herewith, $\sum_{i=1}^{n} q_i < 1$. To simplify the calculations, we take into account the following assumptions.

1. The occurrence of several threats at the same time is impossible.
2. The next threat can arise only in the case of successful eliminating the previous one.

Furthermore, we assume that all events in the system occur at discrete moments of time: $t = 0, 1, 2, \ldots$. In accordance with these assumptions at each t the system is in one of the following states $s_0, s_1, \ldots, s_n, s_f$. In the state s_0, called *security state*, none of the threats are realized. The state s_i, where $i = 1, \ldots, n$, is characterized by the action of the i-th threat. If the system is in the state s_i, then there are two alternatives at the next moment of time:

- the threat is successfully eliminated with the probability r_i and the system comes back in the state s_0;
- the threat leads to the system failure with the probability $\bar{r}_i \equiv 1 - r_i$.

In the last case we say that the system makes the transition to the *final state* s_f. The state diagram of the system is shown in Fig. 1.

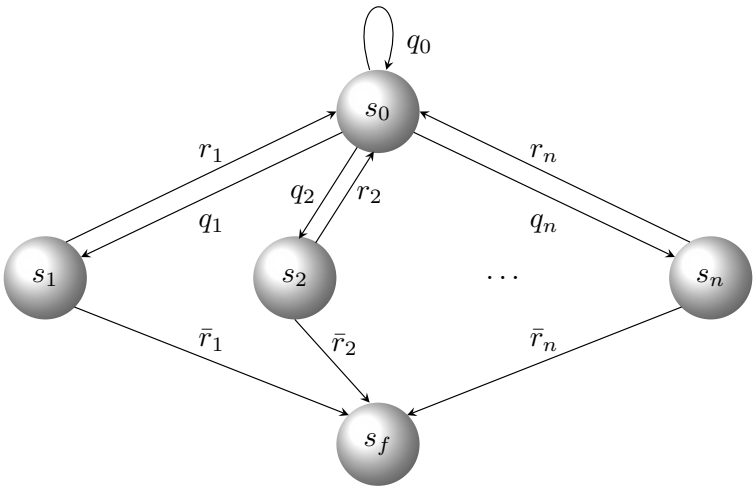

Figure 1. The state diagram of the system.

Here and elsewhere we assume that at the initial moment of time $t = 0$ the system is in the security state s_0.

To analyze the functional and temporal characteristics of the simulated system, it is convenient to introduce the following parameter. We call the number of transitions between the states s_0 and

s_f as the *lifetime* T of the system. Since transitions between states have the stochastic nature, it is clear that the lifetime T is a discrete random variable with some probability distribution. We denote the expected value of the random variable T (the *average lifetime*) as $\overline{T}$. It is obvious that the average lifetime is some function of the original model parameters, i.e.

$$\overline{T} = \overline{T}(\mathbf{q}, \mathbf{r}),$$

where $\mathbf{q} = (q_1, \ldots, q_n)$, $\mathbf{r} = (r_1, \ldots, r_n)$. Further we obtain an explicit analytical expression for this function.

In many practical tasks of designing and using information security systems, there is a problem of selecting an optimal subset from a given set of security remedies. Let us formulate one of such optimizing problems based on the mathematical model under consideration.

Let us assume that we have m various security remedies that eliminate actual information security threats. To describe possible configurations, we assign a Boolean variable x_a: $x_a = 1$ to every a-th security remedy if a-th security remedy is involved, and otherwise $x_a = 0$. Thus, we can describe each specific configuration of security remedies by an m-dimension Boolean vector

$$\mathbf{x} = (x_1, \ldots, x_m) \in \{0, 1\}^m.$$

Obviously, there are 2^m possible configurations.

Let us denote the probability of eliminating i-th threat by a-th security remedy by r_{ia}. Since a few security remedies can eliminate the threat at the same time, in accordance with the addition rule for the probability of a few not mutually exclusive events we have [13]:

$$r_i(\mathbf{x}) = \sum_{b=1}^{m} (-1)^b \sum_{a_1 < a_2 < \cdots < a_b} (r_{ia_1} x_{a_1})(r_{ia_2} x_{a_2}) \ldots (r_{ia_b} x_{a_b}). \tag{1}$$

Hence, the parameters r_i, which define the ability of the system to resist given information threats, are functions of the Boolean variables x_a.

Let us choose some moment of time $T_0 > 0$ and require the average lifetime of the system being not less than the given value: $\overline{T} \geq T_0$. We are interested in such a configuration of the security system, for which the total cost of all security remedies involved is the lowest possible:

$$C(\mathbf{x}) = \sum_{a=1}^{m} c_a x_a \to \min, \tag{2}$$

$$\overline{T}(\mathbf{q}, \mathbf{r}(\mathbf{x})) \geq T_0. \tag{3}$$

Here c_a is the cost of a-th security remedy (in a conventional monetary unit).

Since in the general case the probabilities $r_i(\mathbf{x})$ are non-linear polynomials of the Boolean variables x_a, problem (2), (3) belongs to the class of non-linear discrete optimization problems. It is known that there are no universal effective algorithms for solving this problem. Among low-efficiency methods, the "brute-force" method is primary used. It involves systematic enumeration of all possible m–dimensional Boolean vectors $\mathbf{x}$ satisfying condition (3), and then it checks whether each of them minimizes the objective function $C(\mathbf{x})$. It is clear that the computational complexity of this approach is $O(2^m)$.

The main aim of this work is to explore the possibility of solving optimization problem (2), (3) by the *method of sequential analysis of variants* [14]. Based on sequential construction, analysis, and selection of possible solutions, this method demonstrates good performance for certain classes of optimization problems. First of all, we mean classes of problems with additional properties that allow one to drop some subset of potential solutions. This work includes experimental analysis of the method efficiency applied to solving optimization problem (2), (3).

3. Theory

3.1. Deriving the formula for the average lifetime

The mathematical model described in the previous section allows a natural interpretation in terms of Markov chains. In accordance with the state diagram shown in Fig. 1, the time evolution of the system is a sequence of states. The probabilities of these states at an arbitrary moment in time t are defined by the recurrence relations:

$$p_i(t) = \sum_{j=0}^{n} p_j(t-1)\pi_{ji}.$$

Here $p_i(t)$ is the probability to find the system in the state s_i at time t, π_{ji} is the probability of the system transition from s_j to s_i. The set of all probabilities π_{ji} forms the so-called transition matrix, which in our case has the form

$$\Pi = \begin{pmatrix} q_0 & q_1 & q_2 & \dots & q_n & 0 \\ r_1 & 0 & 0 & \dots & 0 & \bar{r}_1 \\ r_2 & 0 & 0 & \dots & 0 & \bar{r}_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ r_n & 0 & 0 & \dots & 0 & \bar{r}_n \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}, \tag{4}$$

where $q_0 \equiv 1 - \sum_{i=1}^{n} q_i$. The fact that the system is in the security state s_0 at time $t = 0$ leads to the following initial conditions

$$p_0(0) = 1, \quad p_1(0) = p_2(0) = \dots = p_n(0) = p_f(0) = 0. \tag{5}$$

The Markov chain defined by transition matrix (4) and initial conditions (5) is thoroughly analyzed in the works [11, 12]. In particular, the explicit analytical formulae for the probabilities $p_i(t)$ are derived in these works. In this study, we only use the probability of the security state and therefore we write out the formula for $p_0(t)$:

$$p_0(t) = \frac{1}{w}\left(\frac{q_0 + w}{2}\right)^{t+1} - \frac{1}{w}\left(\frac{q_0 - w}{2}\right)^{t+1}. \tag{6}$$

Here the positive parameter w is defined as

$$w^2 = q_0^2 + 4\sum_{i=1}^{n} q_i r_i.$$

Using formula (6), we find the probability distribution for the lifetime T. Let us denote the probability of system transition to the final state s_f in exactly T steps, where $T = 2, 3, \dots$, by $P(T)$. Fig. 1 shows the system being in the state s_f in exactly T steps if and only if at time $T-2$ it was in the security state s_0. Since the probability of this is $p_0(T-2)$, for the probability $P(T)$ we have

$$P(T) = p_0(T-2)\sum_{i=1}^{n} q_i(1 - r_i).$$

Here the expression $\sum_{i=1}^{n} q_i(1-r_i)$ defines the probability of the transition from state s_0 to state s_f. Taking into account (6), the probability distribution for the random variable T has the form

$$P(T) = \frac{1}{w}\sum_{i=1}^{n} q_i(1 - r_i)\left[\left(\frac{q_0 + w}{2}\right)^{T-1} - \left(\frac{q_0 - w}{2}\right)^{T-1}\right]. \tag{7}$$

By definition, the average lifetime of the system is the quantity $\overline{T} = \sum_{T=2}^{\infty} P(T)T$. Substituting (7) in this infinite sum, we obtain

$$\overline{T}(\mathbf{q},\mathbf{r}) = w^{-1}\sum_{i=1}^{n} q_i(1-r_i)\sum_{T=2}^{\infty}\left[T\left(\frac{q_0+w}{2}\right)^{T-1} - T\left(\frac{q_0-w}{2}\right)^{T-1}\right] =$$
$$= w^{-1}\sum_{i=1}^{n} q_i(1-r_i)\left[\frac{2}{q_0+w}\sum_{T=0}^{\infty} T\left(\frac{q_0+w}{2}\right)^{T} - \frac{2}{q_0-w}\sum_{T=0}^{\infty} T\left(\frac{q_0-w}{2}\right)^{T}\right]. \quad (8)$$

By virtue of the inequations $|q_0 \pm w|/2 < 1$, the series on the right side of the expression (8) are convergent [15]:

$$\sum_{T=0}^{\infty} T\left(\frac{q_0 \pm w}{2}\right)^{T} = \frac{2\,(q_0 \pm w)}{(2 - q_0 \mp w)^2}.$$

Then we obtain the following expression for the average lifetime:

$$\overline{T}(\mathbf{q},\mathbf{r}) = \frac{1+\sum_{i=1}^{n} q_i}{\sum_{i=1}^{n} q_i(1-r_i)}. \quad (9)$$

The resulting formula expresses the average lifetime of the system in terms of the original parameters of the model; they are the probabilities of threats $\mathbf{q} = (q_1, \ldots, q_n)$ and the security parameters $\mathbf{r} = (r_1, \ldots, r_n)$.

3.2. The method of sequential analysis of variants

Let us consider optimization problem (2), (3). In accordance with formula (9), this optimization problem can be written as follows::

$$C(\mathbf{x}) = \sum_{a=1}^{m} c_a x_a \to \min, \quad (10)$$

$$\mathbf{x} \in \{0,1\}^m\colon \sum_{i=1}^{n} q_i r_i(\mathbf{x}) \geq \left(1 - \frac{1}{T_0}\right)\sum_{i=1}^{n} q_i - \frac{1}{T_0}, \quad (11)$$

where $r_i(\mathbf{x})$ are given by (1).

As noted above, the most direct way to solve this problem is to use the "brute-force" method. In this approach the average number of iterations equals to 2^m, where m is the number of security remedies used. To reduce the number of iterations and speed up calculations, for solving optimization problem (10), (11) we apply the method of sequential analysis of variants [14].

Let us recall some terminology. Any m-dimensional Boolean vector $\mathbf{x} = (x_1, \ldots, x_m)$ is called a *solution* of optimization problem (10), (11), while a vector of the form $\mathbf{x}_{(p)} = (x_1, \ldots, x_p)$, $p < m$ is called its *partial solution*. A solution $\mathbf{x}$ is called *admissible*, if it satisfies inequality (11). If a partial solution $\mathbf{x}_{(p)}$ can be extended to an admissible solution, it is called *admissible partial solution*.

The main idea of the method of sequential analysis of variants is in constructing partial solutions and dropping the solutions that cannot be extended to optimal ones. The solution of the optimization problem can be represented as moving through a decision tree with nodes being associated with partial solutions and leaf nodes symbolizing full solutions.

The partial solutions that cannot be extended to either acceptable or optimal ones are dropped by so-called *elimination tests* $\sigma = \{\xi_0, \xi_1, \ldots, \xi_k\}$ according to the general rule

$$\sigma(h) = h^{(k+1)},$$

where

$$h^{(j)} = h^{(j-1)} \setminus \xi_{j-1}(h^{(j-1)}), \quad j = 0, 1, \ldots, k+1, \quad h^{(0)} = h.$$

Here h is a collect of partial solutions, and $\xi_j(h)$ denotes the set of partial solutions dropped by the test ξ_j. In the collection σ the following two tests are always presented:

- the test ξ_0 checks that a solution is admissible;
- the test ξ_1 compares admissible solutions by value of the objective function.

The optimization problem (10), (11) has some features that should be taken into account in formulating two more elimination tests.

The test ξ_2 uses the fact that the objective function $C(\mathbf{x})$ is non-decreasing. Its application is reduced to calculating the following evaluation of a partial solution $\mathbf{x}_{(p)} = (x_1, \ldots, x_p)$:

$$\alpha(\mathbf{x}_{(p)}) = C(x_1, \ldots, x_p, 0, \ldots, 0).$$

Denote by C^* an upper bound for the minimum of optimization problem (10), (11). In the first iterations, we set $C^* = +\infty$, and then we equate it to the best value of the objective function on the set of constructed admissible solutions. For an arbitrary set of partial solutions h, the elimination test ξ_2 is given by

$$\xi_2(h) = \{\mathbf{x}_{(p)} \in h \colon \alpha(\mathbf{x}_{(p)}) > C^*\}.$$

Another elimination test ξ_3 analyzes the admissibility of partial solutions. The possibility of its application is related to the specific structure of inequality (11). As it is seen from (2), the given inequality can be rewritten in the form

$$g_1(\mathbf{x}) - g_2(\mathbf{x}) \leq 0, \tag{12}$$

where $g_1(\mathbf{x})$ and $g_2(\mathbf{x})$ are non-decreasing functions. Introducing the evaluation

$$\beta(\mathbf{x}_{(p)}) = g_1(x_1, \ldots, x_p, 0, \ldots, 0) - g_2(x_1, \ldots, x_p, 1, \ldots, 1),$$

we write the test ξ_3 as follows:

$$\xi_3(h) = \{\mathbf{x}_{(p)} \in h \colon \beta(\mathbf{x}_{(p)}) > 0\}.$$

4. Experimental results

Using the theory described in the previous section, we developed a C++ program that solves optimization problem (10), (11) by both the "brute-force" method and the method of sequential analysis of variants. Program input data are the following parameters:

- m is the original number of security remedies that are used;
- n is the number of probabilities of threat realizations;
- $\mathbf{q} = (q_1, \ldots, q_n)$ is the vector of probabilities of cyber threats;
- T_0 is the upper bound for the system average lifetime;
- $\|r_{ia}\|$ is the matrix of probabilities of threat eliminations by security remedies;
- $\mathbf{c} = (c_1, \ldots, c_m)$ is the vector of security remedy costs.

Table 1. The numerical experimental results

The number of security remedies m	The number of iterations, N	
	The "brute–force" method	The method of seq. analysis of var.
5	32	38
6	64	61
7	128	218
8	256	263
9	512	409
10	1024	703
11	2048	632
12	4096	1811
13	8192	1116
14	16384	1166
15	32768	4626
16	65536	2303

The result of the program is the vector $\mathbf{x}^*$ showing the optimal configuration of security remedies and the value of the minimum cost $\mathbf{C}^* = \mathbf{C}(\mathbf{x})$.

To compare the effectiveness of the method of sequential analysis of variants with the "brute-force" method, we performed a series of numerical experiments using our program. In each experiment, we executed the program with various values of parameter m in the range of 5 to 16 and measured the number of iterations N required to find the optimal solution. All other input data were formed randomly using the standard function **rand**() from the C++ library **cstdlib**. To obtain objective results, 10 tests for each value of m were performed (other input parameters were generated randomly), and then the expected value and the standard deviation of N were calculated. Table 1 shows the results of the experiments.

In Fig. 2 we plotted the corresponding dependences of the (average) number of iterations on m for both the "brute-force" method and the method of sequential analysis of variants. As can be seen from Fig. 2, the number of iterations required to solve optimization problem (10), (11) by the method of sequential analysis of variants is less than in the "brute-force" method, moreover, this dependence becomes much stronger with increasing. Based on these results, we come to the conclusion that the method of sequential analysis of variants is more effective in searching an optimal configuration of information security remedies, especially when the number of original remedies is large enough.

5. Conclusion

One of the possible problem statements for the optimal selecting information security remedies based on a Markov cyber threat model is considered. Unlike conventional approaches [2, 9] with the set of solutions being constrained to admissible indicators of economic damage, we define the corresponding constrain by a functional and temporal characteristic of the model called its average lifetime. As far as our optimization problem belongs to the class of nonlinear integer programming problems, there are no universal algorithms for its effective solution. However, the problem specifics (the linearity of objective function (10) and the representation of constrain (11) in form (12)) allows us to apply the method of sequential analysis of variants. Using the developed computer program, we performed a series of numerical experiments confirming the effectiveness of this method compared with the "brute-force" method. In the following studies

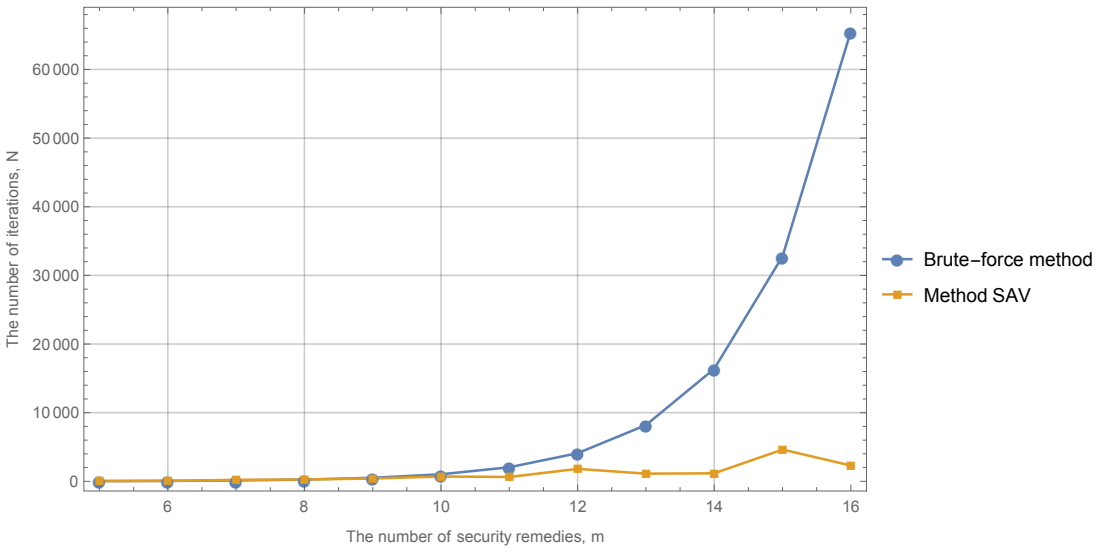

Figure 2. Comparison of the iterations numbers for the "brute-force" method and the method of sequential analysis of variants.

we plan to verify the experimental results using more rigorous theoretical estimates.

Acknowledgements
The reported study was funded by RFBR, project number 19-37-90122.

Bibliography
[1] Hoffman L J 1977 *Modern methods for computer security and privacy* (Prentice-Hall Englewood Cliffs, NJ)
[2] Shirtz D and Elovici Y 2011 *Information Management & Computer Security* **19** 95–112
[3] Cavusoglu H, Mishra B and Raghunathan S 2004 *Communications of the ACM* **47** 87–92
[4] Qiu Q R, Zhang Y F and Han L 2008 An optimization model of product selection in information security technology system *2008 International Conference on Machine Learning and Cybernetics* vol 2 (IEEE) pp 1141–1146
[5] Wang Z and Song H 2008 Towards an optimal information security investment strategy *2008 IEEE International Conference on Networking, Sensing and Control* (IEEE) pp 756–761
[6] Shi J, Lu Y and Xie L 2007 Game theory based optimization of security configuration *2007 International Conference on Computational Intelligence and Security (CIS 2007)* (IEEE) pp 799–803
[7] Fielder A, Panaousis E, Malacaria P, Hankin C and Smeraldi F 2014 Game theory meets information security management *IFIP International Information Security Conference* (Springer) pp 15–29
[8] Bensoussan A, Kantarcioglu M and Hoe S C 2010 A game-theoretical approach for finding optimal strategies in a botnet defense model *International Conference on Decision and Game Theory for Security* (Springer) pp 135–148
[9] Ovchinnikov A, Medvedev N and Bykov A 2007 *Vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta im. N.E. Baumana* (3) 115–121 (in Russian)
[10] Rosenko A 2008 *Izvestiya Yuzhnogo federal'nogo universiteta. Tekhnicheskie nauki* **85**(8) 71–81 (in Russian)
[11] Magazev A and Tsyrulnik V 2018 *Automatic Control and Computer Sciences* **52**(7) 615–624
[12] Magazev A and Tsyrulnik V 2018 Optimizing the selection of information security remedies in terms of a Markov security model *Journal of Physics: Conference Series* vol 1096 (IOP Publishing) p 012160
[13] Feller W 2008 *An introduction to probability theory and its applications* vol 1 (John Wiley & Sons)
[14] Kovalev M 2003 *Discrete optimization (integer programming)* (Editorial URSS, Moskow) (in Russian)
[15] Prudnikov A, Brychkov Y and Marichev O 1986 *Integral and Series. Elementary functions* vol 1 (Gordon and Breach)