UDC 004.056

R.B. Prus, V.A. Shvets, Ph.D. (National Aviation University, Ukraine)

FORMATION OF THE OBJECTIVE FUNCTION IN THE TASKS OF INFORMATION SECURITY MANAGEMENT

The model of search for the optimal defense resource allocation is presented and objective function that minimizes the information loss is chosen. The choice of parameters and dependencies that form the objective function is justified. Estimates of optimal resource allocation are represented. It is conducted sensitivity analysis of mathematical models to components of the objective function.

The main task of information security management is to optimize its indices. Looking for the solution of this problem and building a mathematical model, arise main difficulties when choosing the type of optimality criteria and the objective function, and defining the parameters and dependencies that are part of objective function. The first step in the formation of objective function is the choice of objective indicator. It can be the quantity of extracted information, information security costs, the total loss, which combine losses from leaks and the cost of its defense, cost effectiveness, and the combination of these and other indicators [1-3]. Select objective function as [3]:

$$I(x,y)=\sum_{k=1}^{l} I_k(x,y)=\sum_{k=1}^{l} g_k p_k q_k(x) f_k(x,y), \qquad (1)$$

where x i y - resources of attack and defense, respectively ($\sum_{k=1}^{l} x_k = X$, $\sum_{k=1}^{l} y_k = Y$);

$k=\overline{1,l}$ - object number;

g_k - amount of information on an object;

p_k - probability of attack on k - object;

$q_k(x)$ - probability of attack resource allocation x on k - object;

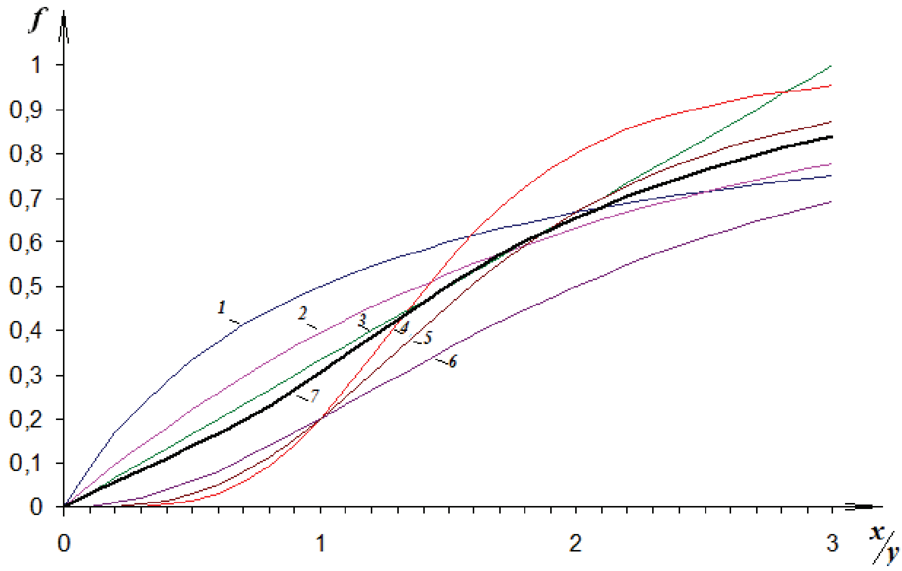
$f_k(x,y)$ - part of extracted information dependence on the ratio x i y.

Variables that are part of the objective function (1), determined on the basis of statistics or through peer review. Our interest in the objective function components due to the fact that we operate under uncertainty. In this case, the crucial question of sensitivity of optimal objective function value $I^0(x,y)=I(x,y^0)$ to change values that fall within its right side.

Begin the research of sensitivity analysis function $I(x,y)$ to dependence $f(x,y)$. For this purpose, consider some dependence $f_t(x,y)$ (t-dependence's number) and follow the changes the optimal resource allocation and relevant information loss when using each dependence $f_t(x,y)$.

Establishing a kind of dependence $f(x,y)$ should note the following considerations: when $x/y \to 0$ $f(x,y)\to 0$, when $x/y \to \infty$ $f(x,y)\to 1$. Consider a few functions that meet this criteria (Figure 1). Also, build averaged function (curve 7), thus seize the Bernoulli-Laplace approach, considering all the options dependencies $f_t(x,y)$ are equiprobable in the lack of statistical information.

Using the Belman's method [4], find the optimal resource allocation $\{y_k^0\}$ for defense and relevant part of extracted information I^0 for system with three objects in the case when an attack is launched on all three objects with coefficients $g_1=0{,}5$, $g_2=0{,}3$, $g_3=0{,}2$.



1 - $f = \dfrac{x/y}{x/y + 1}$;

2 - $f = 1 - e^{-0,5 \cdot \frac{x}{y}}$;

3 - $f = \begin{cases} \frac{1}{3} \cdot \frac{x}{y} & \text{при } \frac{x}{y} \le 3; \\ 1 & \text{при } \frac{x}{y} > 3 \end{cases}$

4 - $f = \dfrac{(x/y)^4}{(x/y)^4 + 4}$;

5 - $f = \dfrac{(x/y)^3}{(x/y)^3 + 4}$;

6 - $f = \dfrac{(x/y)^2}{(x/y)^2 + 4}$;

7 - $\overline{f} = \dfrac{\sum_{i=1}^{6} f_i(x,y)}{6}$

Fig. 1 Dependencies of part extracted information on input resources

In Tab. 1 are given optimal allocations $\{y_k^0\}$ for various functions $f(x,y)$ and different values $Z = \dfrac{X}{Y}$. These allocations correspond to saddle points, i.e. providing compliance equality $\min_j \max_i I(x,y) = \max_i \min_j I(x,y)$, where i and j - options of resource allocation (strategies) attack and defense, respectively.

Table 1

Optimal defense strategies for different objective functions

Number of function $f(x,y)$	Optimal strategies	Z=0,5	Z=1	Z=1,5	Z=2	Z=2,5	Z=3
1	$y_1^0 : y_2^0 : y_3^0$	0,5:0,3:0,2					
	I^0	0,329	0,5	0,599	0,667	0,714	0,75
2	$y_1^0 : y_2^0 : y_3^0$	0,5:0,3:0,2					
	I^0	0,22	0,393	0,527	0,632	0,713	0,777
3	$y_1^0 : y_2^0 : y_3^0$	0,98:0,01:0,01					
	I^0	0,218	0,303	0,383	0,473	0,558	0,643
4	$y_1^0 : y_2^0 : y_3^0$	0,42:0,34:0,24	0,5:0,3:0,2	0,62:0,25:0,13	0,48:0,35:0,17	0,48:0,35:0,17	0,5:0,28:0,22
	I^0	0,167	0,4	0,566	0,795	0,905	0,952
5	$y_1^0 : y_2^0 : y_3^0$	0,43:0,33:0,24	0,5:0,3:0,2	0,59:0,26:0,15	0,52:0,33:0,15	0,51:0,31:0,18	0,49:0,33:0,18
	I^0	0,141	0,333	0,474	0,665	0,794	0,87
6	$y_1^0 : y_2^0 : y_3^0$	0,5:0,3:0,2					
	I^0	0,114	0,25	0,374	0,5	0,609	0,692
7	$y_1^0 : y_2^0 : y_3^0$	0,5:0,33:0,17	0,53:0,3:0,17	0,53:0,32:0,15	0,5:0,3:0,2	0,8:0,1:0,1	0,8:0,1:0,1
	I^0	0,178	0,345	0,504	0,655	0,739	0,796

As seen from the table, using linear function f_3 the best option for defense in case of any values Z is to focus its resources on the most critical first object. If the functions f_1, f_2, f_6 give the same result, namely: is the optimal allocation $\{y_k^0\} = (0{,}5;0{,}3;0{,}2)$, then use the f_4 and f_5 observe some deviations from these options depending on the value *Z*. For example, when *Z*=0,5 best option is the defense resource allocation in the ratio $\{y_k^0\} = (0{,}42;0{,}34;0{,}24)$ (function f_4) and $\{y_k^0\} = (0{,}43;0{,}33;0{,}24)$ (function f_5), when *Z*=1,5 optimal allocation when using f_4 - $\{y_k^0\} = (0{,}62;0{,}25;0{,}13)$ and for function f_5 - $\{y_k^0\} = (0{,}59;0{,}26;0{,}15)$.

Analyze the results taking into account the type of dependence $f_t(x, y)$ (fig. 1). When using functions f_1, f_2, f_6 information loss increases monotonically throughout the interval Z from 0 to 3 and reach values 0,7-0,75. Smooth running of curves apparently is the reason that the optimal defense resource allocation remains the same under any values Z.

Functions f_4 i f_5 have a slightly different nature. When $Z \le 0{,}5$ $f \gtrsim 0$, and the interval from $Z = 0{,}5$ to $Z = 2{,}5$ loss of information dramatically increase to the value $f = \overline{0{,}8;0{,}9}$, then gradually approaching 1. That is why on these three segments optimal defense resource allocation is differing.

Considering the power functions f_4, f_5, f_6, note that the optimal defense resource allocation does not depend on Z in power $n = 2$, and for the case n>2 it varies, reaching a maximum concentration of resources on first object when $Z = 1{,}5$ - the largest slope area of dependence $f_t(x, y)$.

Now try another method of averaging. We will not average function $f(x, y)$, where we find the result, but outcomes, i.e. $\{y_k^0\}$ i I_t^0, and conduct averaging on each function $f_t(x, y)$ and on value Z at that.

To conduct averaging Z, choose some normalized ratios C_Z ($\sum C_Z = 1$), considering that the probability of an attack resource allocation $Z = 2$ is highest. Averaged allocation calculates as follows:

$$(y_k^0)_Z = C_{0,5} \cdot y_k + C_1 \cdot y_k + C_{1,5} \cdot y_k + C_2 \cdot y_k + C_{2,5} \cdot y_k + C_3 \cdot y_k ,$$

$$(y_1^0)_Z = 0{,}05 \cdot 0{,}5 + 0{,}15 \cdot 0{,}53 + 0{,}2 \cdot 0{,}53 + 0{,}25 \cdot 0{,}5 + 0{,}2 \cdot 0{,}8 + 0{,}15 \cdot 0{,}8 \approx 0{,}62 .$$

The results are given in Tab. 2.

Table 2

Optimal resource allocations and relevant information loss by different averaging options

Z	Value of f averaging		Value of $\{y_k^0\}$ averaging	
	$\{y_k^0\}$	I^0	$\{y_k^0\}$	I^0
0,5	0,5:0,33:0,17	0,178	0,55:0,27:0,18	0,210
1	0,53:0,3:0,17	0,345	0,58:0,25:0,17	0,370
1,5	0,53:0,32:0,15	0,504	0,61:0,24:0,15	0,513
2	0,5:0,3:0,2	0,655	0,58:0,27:0,15	0,658
2,5	0,8:0,1:0,1	0,739	0,58:0,26:0,16	0,764
3	0,8:0,1:0,1	0,796	0,58:0,25:0,17	0,825
Value of Z averaging	0,62:0,23:0,15	0,531	0,58:0,26:0,16	0,614

The results in table conclude that averaging value $\{y_k^0\}$, obtain optimal allocation virtually identical for all values Z - $\{y_k^0\} = (0{,}58 : 0{,}25 : 0{,}17)$ (with modifications maximum 0,03), whereas the optimal allocation is found using the averaged function $\overline{f}(x,y)$, in sharp contrast to values $Z = 2{,}5$ and $Z = 3$ towards concentration of a significant part of resources ($y_1 = 0{,}8$) on the first object, where the largest amount of information. However, in this case of averaging the expected information losses less than in the case of averaging $\{y_k^0\}$.

Averaging Z also allows comparing the two approaches. These allocations contain only minor differences, so any of the methods can be used when searching for the optimal resource allocation.

To summarize, we note that using the averaged function $\overline{f}(x,y)$ found the optimal resource allocation is little different from the average optimal allocation, but this method requires fewer calculation procedures. Therefore, to study the impact of other components of the objective function for optimal results be using the averaged function $\overline{f}(x,y)$.

Turn now to the other variables in (1). Considering the values g, p, q, f independent variables, based on Taylor's theorem we have:

$$\nabla I^0(g,p,q,f) = \nabla_g I^0 \partial g + \nabla_p I^0 \partial p + \nabla_q I^0 \partial q + \nabla_f I^0 \partial f$$

and determine sensitivity coefficients of function $I^0(x,y)$ to these variables:

$$\alpha_g = \frac{\partial I^0}{\partial g} = \nabla_g I^0;\ \alpha_p = \frac{\partial I^0}{\partial p} = \nabla_p I^0;\ \alpha_q = \frac{\partial I^0}{\partial q} = \nabla_q I^0;\ \alpha_f = \frac{\partial I^0}{\partial f} = \nabla_f I^0.$$

For example, consider the value g. Coefficient α_g shows how to change the quantity of extracted information with optimal defense resource allocation, if the actual value g different from the adopted. If $\Delta g = 0{,}1$, we have: when $Z = 1$ $\Delta I^0 = 0{,}057$, when $Z = 2$ $\Delta I^0 = 0{,}065$, when $Z = 3$ $\Delta I^0 = 0{,}080$.

Apparently, the sensitivity of the optimal solution increases with $Z = \frac{X}{Y}$. Similar findings we get for other variables. Because the rate of change of partial derivative linear function with respect to independent variables c_s expressed the value of this function at $c_s = 1$ (in our case - the product of all other variables), it is clear that the sensitivity α_{c_s} will be the greater, the more value $I^0(g,p,q,f)|_{c_s=1}$ at the stationary point.

Conclusions

Using the Belman's method and target function (1) is found optimum defense resource allocation under uncertainty. Considering the difficulties of establishing the type of dependence of part extracted information on input resources, different methods of averaging is analyzed – the dependencies themselves $f(x,y)$ and averaging the results $\{y_k^0\}$. As result of research substantiated the expediency of using the averaged function $\overline{f}(x,y)$.

On the basis of analysis sensitivity coefficients of objective function to input data is made conclusion that considered model can be used to find optimal solutions because deviations from the expected loss is permissible under uncertainty.

References

1. *Gordon L.* The Economics of Information Security Investment / *Loeb M.* // ACM Transactions on Information and System Security, November 2002. – Vol. 5, №4. – P. 438-457.

2. *Zadiraka V.K.* Financing costs protection in economic activity / *Oleksiuk O.S, Smolyanyuk R.P, Shtabalyuk P.I* // University proceedings. – 2006. – №3-4 (19-20). – P. 479-490.

3. *Levchenko E.G.* Multilevel resource allocation between elements of information security systems / *Prus R.B, Shvets V.A.* // Bulletin of the Engineering Academy of Ukraine. – 2009. - №2. – P.90-94.

4. *R. Bellman* Dynamic programming. – M.: YYL, 1960. – 400 p.