UDC 004.056.53

Mokhor V.V., Tsurkan V.V.

PROBIT-METHOD FOR INFORMATION SECURITY RISK ASSESSMENT

Анотація:

Приведено обґрунтування постановки задачі розвитку методології кількісної оцінки ризиків безпеки інформації конкретних об'єктів інформаційної діяльності на основі пробіт-аналізу.

Ключові слова: *пробіт-аналіз, пробіт-функція, безпека інформації, ризик, кількісна оцінка ризиків, аналіз ризиків.*

Аннотация:

Приведено обоснование постановки задачи развития методологии количественной оценки рисков безопасности информации конкретных объектов информационной деятельности на основе пробит-анализа.

Ключевые слова: *пробит-анализ, пробит-функция, безопасность информации, риск, количественная оценка рисков, анализ рисков.*

Abstract:

The rationale statement of the problem of quantitative methodology for information security risk assessment of specific sites of information activities on the basis of the probit-method.

Keywords: *probit-method, probit-function, information security risk, quantitative assessment.*

Currently, the presence of information security management systems is becoming one of the key conditions for the strategic development of any organization. According to the standard ISO/IEC 27001:2005 [1], the requirements for the establishing, implementing, operating, monitoring, reviewing, maintaining and improving information security management system must be in the context of information security risk management organization. The main and most difficult stage of risk management is the information security risks analysis.

For the information security risks analysis there are different methods, namely: Austrian IT Security Handbook, AS/NZS4360, BSI 100-3, CRISAM, EBIOS, HB167: 200X, ISF IRAM, ISO 27005:2008, ISO 31000, MAGERIT, MARION, MEHARI , NIST SP800-30, OCTAVE, OSSTMMRAV, SOMAP and others. The choice of a method depends on the requirements of the organization to ensure information security, shall be taken into account the risks and the effectiveness of measures to protect the information.

According to [2], the problem of information security risk analysis is divided into two stages:

- risk identification;
- risk assessment.

In this paper we are interested in the problem of risk assessment, which according to [2] implies:

- determination of the method of risk assessment, and assessment, in turn, can be
    - quality;

  - quantify;
- assessment of the impact of information security incidents;
- characterization of the probability of information security incidents;
- calculation of risk.

In this case, there are four approaches to quantitative risk assessment, differentiated by the investigated areas of its manifestations. In relation to information security appropriate to consider a technocratic approach, based on an analysis of the relative frequencies of occurrence of hazards with the fallout. Methodology for quantitative risk assessment is based on probabilities of initiating events, scenarios to possible incidents and the corresponding probabilities of their realization. In accordance with this methodology, the following methods of quantitative risk assessment:

1. Statistical methods require determination of the probability that a threat to the reporting of information assets within the time interval based on the following requirements:

- objects to the analysis of the statistics to be used, and the objects on which to collect statistics, are equivalent (the requirement of equivalence of objects);
- the conditions under which it is supposed to use the statistics, and the conditions of its collection are equivalent (the equivalence of the requirement);
- sample sizes are sufficient statistics, methods of processing - correct, and sources of information - credible (requiring urgent).

The disadvantages of this method are a group critical of the original data, which is usually either absent or not enough of them to build the correct output.

2. Probabilistic and statistical methods are used to attract additional information on the distribution of damage in case of implementation of the information security risks of the asset. It is assumed that for the operating conditions of organizational and technical system of the company is known for the distribution function of loss of information security incidents. On its basis, the share is defined catastrophic events of all adverse events. Assuming this proportion constant or time series forecasting its value at a given time, we can determine the probability characteristics of catastrophic events.

The accuracy and reliability of the results obtained with the use of probabilistic and statistical methods for determining the quality and the amount of additional information on the distribution of losses.

3. Theoretical and probabilistic methods are used to determine the frequency or probability of the rare information security threats, with significant consequences for which statistics is virtually nonexistent. The basis of this method is based on the laws of initiating events from escalating into extraordinary, the decomposition of the problem, evaluate specific indicators and to determine the frequency of rare adverse events according to the relationship of particular indicators.

Theoretic probabilistic method is labor-intensive, has low accuracy and reliability of the results of the research process, but in the absence of other estimates its use is justified.

4. Expert methods are based on knowledge and experience of experts. These methods are useful in the case where there are no statistics. At the same time, experts are invited to answer questions about the future behavior or information assets, characterized by uncertain parameters or unexplored properties. For interpretation or mathematical processing of expert data, you can use the mathematical apparatus of the theory of fuzzy sets.

The complexity of information security risk analysis expert methods related primarily to the uncertainty characteristics of data sets, based on the experience that shaped the expert and, as a consequence, the lack of guarantees reliable results.

Thus, we can conclude that there are significant limitations in the application of the known methods for quantitative risk assessment of the security of information, and therefore the search for new approaches to tackle the problem of determining the characteristics of the probability (chance) of information security in low statistics, is an actual problem.

In this context, it seems promising to consider the probit-method, the idea of which belongs to the American entomologist Charles Bliss, first described it in an article on the impact of pesticides on the percentage of pest control [3 - 4]. C. Bliss suggested to account for the percentage Pest use probabilistic block - «**prob**ability un**it**» or «probit». First need to introduce the «probit» was motivated by a desire to avoid work with statistical information. Biologists, for whom it was intended, and the method were less familiar with the statistical treatment of the results of the experiment. Currently, this reason has lost its significance, however, the name «probit» and «probit-method» have become common terms, the methodology of the «probit-method» has been developed and is widely used in toxicology, pharmacology, radiobiology, entomology, ecology and other fields both biological and medical research.

The essence of the probit-method is special mapping of *S*-curves similar to the characteristics of the realized losses threats in straight lines, which can then be processed by linear analysis. The inverse transformation is carried out by converting the linear probit-functions in the values of the characteristics of probability.

It is known that the probit-function is a mathematical relationship that links the specific characteristics of the negative impact on an object (in our case - the information asset) to the size of potential losses. Expression for determining the values of the probit-function in the general case, is the following:

$$\Pr(D) = a + b \cdot \ln D + \gamma \cdot \ln \tau,$$

where a, b, γ - factors that characterize the vulnerability of information assets in respect of a specific threat or hazard class; D - negative impact assessment; τ - the time from the beginning to the end of the negative effects.

For areas in which we can neglect the time period relevant threats and this is what typical for most information security threats, use the following format for the probit-function:

$$Pr(D) = a + b \cdot ln D \quad (1)$$

By analogy with the functions of the probability distribution of losses from the sale of scenarios of information security threats with the *S*-curves like it should be assumed that the use of an approach based on the probit-method can be effective.

It should be noted that the solution of problem for the probability characteristics for each pair (threat, vulnerability) is performed in two stages:

1. Determination of value of a probit-function $\Pr(D)$, the right-hand side of (1).

2. The calculation of the probit-function $\Pr(D)$ and the definition of the known values of the probit-function values of probability characteristics.

We begin with the end of the problem, ie the second stage. Let the values a,b,D are known. Then, in accordance with (1), we can calculate the value of the probit-function $\Pr(D)$. As a result, calculated from the known $\Pr(D)$ values of the probability of the threat to the security of information on the following formula:

$$P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\Pr(D)} e^{-\frac{x^2}{2}} dx, \quad (2)$$

where probit-function $\Pr(D)$ acts upper limit of integration, and the expression (2) does not contain empirical coefficients.

In practice, to calculate the integral (2) use a table of the probit-function, the use of which will explain with an example.

Table 1

The values of the probit-function

P, %	0	1	2	3	4	5	6	7	8	9
0		2.67	2.95	3.12	3.25	3.38	3.45	3.52	3.59	3.66
10	3.72	3.77	3.82	3.86	3.92	3.96	4.01	4.05	4.08	4.12
20	4.16	4.19	4.23	4.26	4.29	4.33	4.36	4.39	4.42	4.45
30	4.48	4.50	4.53	4.56	4.59	4.61	4.64	4.67	4.69	4.72
40	4.75	4.77	4.80	4.82	4.85	4.87	4.90	4.92	4.95	4.97
50	5.00	5.03	5.05	5.08	5.10	5.13	5.15	5.18	5.20	5.23
60	5.25	5.28	5.31	5.33	5.36	5.39	5.41	5.44	5.47	5.50
70	5.52	5.55	5.58	5.61	5.64	5.67	5.71	5.74	5.77	5.81
80	5.84	5.88	5.92	5.95	5.99	6.04	6.08	6.13	6.18	6.23
90	6.28	6.34	6.41	6.48	6.55	6.64	6.75	6.88	7.05	7.33
-	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
99	7.33	7.37	7.41	7.46	7.51	7.68	7.65	7.75	7.88	8.09

Suppose you want to find the probability of the information security threats from the known value of probit-functions, $\Pr(D) = 4.53$. This value corresponds to the intersection of the lines of «30%» and column «2%» of the table, in which the probability of the information security threats is 32%. In addition, it is possible that the probit-function takes intermediate values, for example: $\Pr(D)=4.518$, $\Pr(D)\in[4.50;4.53]$. In this case, the probability of the information security threats can beat determined, for example, using a linear interpolation by the following expression

$$P(x)=P(x_0)+\frac{P(x_1)-P(x_0)}{x_1-x_0}\cdot(x-x_0),\ x\in[x_0,\ x_1]. \tag{3}$$

Substituting taken from the table, the values of the probit-function and the corresponding probability in (3), we define $P(4.518)$

$$P(4.518)=P(4.50)+\frac{P(4.53)-P(4.50)}{4.53-4.50}\cdot(4.518-4.50)=31\%+\frac{1\%}{0.03}\cdot 0.018\approx 31.6\%.$$

Thus, the second stage of the solution of the evaluation of probabilities in quantifying information security risk based approach using probit-method can presume to hold.

Now consider the first stage of the problem - defining element values probit-function $\Pr(D)$ the right-hand side of (1). From this expression it is seen that the presence of values $\Pr(D)$ for a specific threat or class of security threats involves determining the values of the coefficients *a, b* and the value of *D*. In this case, directly from (1) that the coefficient *a* reflects a level of fixed costs required to maintain the security given asset, and *b* is a factor «increasing» loss *D*, due to the development of effective scenario update information security threats.

According can specify two alternative ways of determining the values of *D*:

1. Values *D* are defined as the ratio of

$$D=\frac{h}{H}\ ,$$

where h - the value of the losses of responding to the activity of a specific threat against a specific asset information; H - maximum losses, possible due to the implementation of information security threats for the object.

2. Value *D* is determined based on the results of the analysis of parameters of threats and vulnerabilities in the form of a product of a some coefficient λ_h and a coefficient of vulnerability v_t:

$$D=\lambda_h\cdot v_t\,.$$

In this case:

$$\lambda_h=\lambda_0\sum_{i=1}^{M}\delta_i\cdot a_i, \tag{4}$$

$$v_t = v_0 \sum_{j=1}^{N} \varphi_j \cdot a_j \qquad (5)$$

where λ_0, v_0 - a scale factor values are selected on the basis of the conditions:

$$0 < \lambda_h \leq 1,$$

$$0 < v_t \leq 1.$$

From the expressions (4) and (5) that the hazard ratio λ_h and the coefficient of vulnerability v_t defined as the sum M or, as appropriate, N performance a_i or a_j. According to the standard [2], the identification of indicators of threats and vulnerabilities is in the process of risk analysis. In this case, the numerical values of a_i and a_j are set, for example, on a scale based on their ordering and ranking, conducted, for example, the analytic hierarchy process. For smoothing/gain a degree of some indicators of each of them can be placed in the corresponding parameter value δ_i (or φ_j) is selected in the range (0, 1), subject to conditions:

$$\sum_{i=1}^{M} \delta_i = 1$$

or

$$\sum_{j=1}^{N} \varphi_j = 1,$$

respectively. Selecting one of the two ways of determining the values of the quantity *D* depends on the chosen method of risk analysis [2]. If the methodology is focused on priority loss analysis, then you should choose the first option. If the method prioritizes the analysis of threats and vulnerabilities, you should choose the second method.

The greatest uncertainty is related to the definition of the coefficients *a* and b for the sphere of information security. In general, the set of coefficients for each pair of threat/losses are a result of the individual studies. In particular, such studies are carried out, for example, in entomology, ecology, toxicology, pharmacology, radiobiology, and other areas of the biological and medical sciences. Known examples of such studies in hydraulic engineering, structural mechanics, fire safety and other areas other than information security. However, the analysis and identification of features and characteristics, such as information security threats and vulnerabilities inherent in various information assets, a recommendation to the standard [2] and the regulatory requirement of the standard [1]. Currently, the implementation of such studies for each organization and for each of the assets is virtually impossible, both because of the diversity of conditions and the context of their operation, and because of the dynamics of these conditions over time.

Using the same approach, which consists in building a probit-function, creating a unified methodological framework coping with the threat/losses for various information assets and provides a basis for the solution of the quantitative risk assessment in the field of information security.

Thus the development of the methodology of the probit-method as applied to the field of information security can be considered as a separate, independent lines of research.

Literature:

1. ISO/IEC 27001:2005. Information technology. – Security techniques. – Information security management systems. – Requirements.

2. ISO/IEC 27005:2011. Information technology. – Security techniques. – Information Security Risk Management.

3. Bliss C.I. The method of probits / C.I. Bliss // Science. – 1934. – Vol. 79, No. 2037. – P. 38–39.

4. Bliss C.I. The method of probits – a correction / C.I. Bliss // Science. – 1934. – Vol. 79, No. 2053. – P. 409–410.