Definition of basic violators for critically important objects using the information probability method and cluster analysis

Vladimir Kostin
Orenburg State University
Orenburg, Russia
vladimirkostin5@mail.ru

Aleksandr Borovsky
Orenburg State University
Orenburg, Russia
borovski@mail.ru

***Abstract*—One of the approaches to analyzing the relationships between the characteristics of critical objects and typical intruders using the information-probabilistic method is considered. Categories of objects and typical intruders are described by many common heterogeneous characteristics. The probabilistic information method ensured the homogeneity of the entropy potential of the offender training characteristics and the characteristics of the consequences of the offender actions according to the Pearson chi-square criterion and on this basis the characteristics of the offenders and critical objects are summarized in a common information field in single six-point measurement scales. Using the information probabilistic method and the cluster analysis method for the general information field, we obtained the basic type of intruder for each category of objects. The results can be used to determine the requirements for physical protection systems of critical facilities.**

Keywords—information-probability method; entropy; base type of intruder

I. INTRODUCTION

The analysis of the source [1] revealed that each protected object has an attractive potential, which is formed by a hazard potential, in accordance with which the necessary security potential is formed in the form of a physical protection system (PPS). In turn, each typical intruder has the potential for preparedness (danger), which is determined by the degree of his motivation. Thus, many typical intruders have a definite impact on many categorized objects in accordance with their potential capabilities. Typical violators and categorized objects have many characteristics that determine their potential motivation (danger) and attractiveness, which intersect informationally. Obviously, between the categories of critical objects (CIO) and typical intruders (threats) there must be a certain correspondence, which is based on the general commensurability or ratios of the characteristics of these sets (potentials). That is, each category of CIO must correspond to a certain basic type or types of violators.

The author of the article suggests, using the information probabilistic method (IPM) and cluster analysis, to determine the typical basic intruders for each category of CIO, i.e. identify existing compliance. Based on the hazard results of typical violators, an appropriate level of security (safety) of objects is proposed.

This problem is solved to differentiate the necessary requirements when determining the required value of protection of categorized objects from basic typical violators.

Currently, the task of determining the basic threats for various categories of objects is determined mainly by expert methods [2], where there is an element of subjectivity, or based on the theory of fuzzy logic and fuzzy hypergraphs [3]. Recent methods do not allow us to estimate their weight contribution to the formation of the hazard potential of the object. In addition, sometimes when composing fuzzy membership matrices (performing disjunction and conjunction operations) for certain data (if one of the characteristics is significantly small), fuzzy logic methods show doubtfully low results.

II. FORMULATION OF THE PROPLEM

On the basis of processing the combined general information field of the characteristics of violators and categorized objects with an information probabilistic method and cluster analysis, it is necessary to determine each category of CIO of the corresponding typical intruder, that is, it is necessary to determine the degree of potential impact of the i-type intruder on the j-th category of CIO, to solve the problem of determining ratios and on this basis to propose the necessary amount of security for each category of objects. The source [4] discusses the finding of correspondences on images presented in the form of graph models, but this problem has a limitation — the number of vertices in the image graphs must be the same. The source [5] provides an in-depth analysis of clustering methods. A technique for clustering big data is proposed. The issues of graph clustering are considered, while the weight of the signs of cluster analysis is not taken into account.

III. PROBLEM SOLVING

To solve this problem, it is necessary to form a common information field in a unified scale for measuring the characteristics of violators and CIO.

To assess the potential danger of the object from the actions of violators, six private types of losses were introduced [4]:

- political (determined by a decrease in all levels of authority of the authorities and general instability);
- human (consists in the loss of people's lives and health);
- financial (consists in the loss of material values);
- economic (take into account the costs of relocating people from the accident zone and the related compensation payments);
- environmental (loss of natural resources leading to environmental degradation in the region);
- informational (losses consisting in the loss of advanced technologies, confidential information and artistic values).

For each particular type of loss, one of the six scales of potential losses in the event of an emergency (ES) was

determined in the form of a six-point hazard scale, which are presented in Table 1.

The results of the hazard assessment (attractiveness) for the seven categories of CIO in case of emergencies on a six-point scale, obtained in article [7], are shown in table 2.

TABLE I. THE SCALE OF POTENTIAL LOSSES IN CASE OF EMERGENCIES

Indicators	**Types of Emergencies**					
	The local character	**of the municipal character**	**of the intermunicipal character**	**of the regional character**	**of the interregional character**	**of the federal character**
Injured people	no more than 10	no more than 50	no more than 50	more than 50, but no more than 500	over 50, but no more than 500	over 500
Size of property damage (million rubles)	no more than 0.1	no more than 5	no more than 5	more than 5, but not more than 500	more than 5, but not more than 500	more than 500
The scale of the partial losses on the six-point scale [13]	1	2	3	4	5	6

TABLE II. CHARACTERISTICS OF THE CONSEQUENCES OF EMERGENCIES AT FACILITIES ON A SIX-POINT SCALE

Private types of loss of objects	**The scale of losses of categories of objects**						
	1-cat	**2-cat**	**3-cat**	**4-cat**	**5-cat**	**6-cat**	**7-cat**
Political	5	4	4	3	2	2	1
Human	5	5	4	3	2	2	1
Financial	5	5	4	3	3	2	1
Economic	6	5	4	3	3	2	1
Environmental	6	5	4	4	3	2	2
Informational	6	5	5	4	3	2	2

Characteristics of typical violators are determined by order of the Minister of Industry and Energy of the Russian Federation N150 [8] and Government Resolution N875 [9], which are summarized in table 3.

TABLE III. CHARACTERISTICS OF TYPICAL VIOLATORS

Characteristics of violators	**Type of violator**					
	X_1	**X_2**	**X_3**	**X_4**	**X_5**	**X_6**
Number	5 – 20	3 – 5	1	1	1	1
The goal	of terror. Act	of terror. Act	of terror. Act	Theft	Theft	Theft, ter. Act
Consequences of the actions of the offender	federal, regional, territorial	Beyond the boundaries of the facility	Within the boundaries of the facility	Within the boundaries of the facility	Within the facility	Within the facility
The level of awareness	is the general level	average level of awareness	low level of awareness	low level of awareness	high level of awareness	high level of awareness
Melee and firearms weapons equipment	high probability	high probability	high probability	low probability	low probability	Armed
Level of training to overcome barriers, willingness to engage in battle	High level of training	High level of preparation	High level of preparation	Low level of preparation	Low level of preparation	medium level of preparation

To solve this problem, we describe typical intruders and objects in a single scale of measurement of characteristics. Based on the data of Table 3 and expert evaluations of specialists, a transition was made from qualitative to quantitative characteristics of violators, which are summarized in table 4.

TABLE IV. QUANTITATIVE CHARACTERISTICS OF TYPICAL VIOLATORS

A type violator	**Characteristics of violators**					
	Numbers	**Objective of action**	**Consequences of actions**	**Level of information awareness**	**Cold steel, firearms (technical equipment)**	**Level of physical fitness**
X_1	11	10	0.8780	0.7	0.9	1
X_2	4	9	0.5546	0.6	0.8	0.9
X_3	1	8	0.1731	0.4	0.7	0.8
X_4	1	2	0.0067	0.3	0.3	0.3
X_5	1	2	0.1158	0.9	0.3	0.3
X_6	1	5	0.1731	1	1	0.6

For a comparative assessment of the danger potentials of violators, an IWM was used, which allows one to reduce the particular characteristics of the considered violators (objects) to a complex potential in the form of an entropy index. The essence of the method is as follows.

The decision to determine the entropy indicator is associated with the definition of a vector that quantitatively displays the hazard potential of an emergency (intruder, object).We will present the information display of a specific situation in the form of the following scheme: there are dangerous emergencies being compared (violators, objects); each emergency is associated with a set of characteristics that determine its potential hazard. In this case, the situation of assessing the potential of each emergency in a detailed form is characterized by table 5.

TABLE V. MODIFIED MORPHOLOGICAL MATRIX

Name of characteristics	**Many dangerous emergencies (violators, objects)**				
	$\{A_1\}$	…	$\{A_i\}$	…	$\{A_n\}$
X_1	X_{11}	…	X_{1i}	…	X_{1n}
X_j	X_{j1}	…	X_{ji}	…	X_{jn}
…	…	…	…	…	…
X_m	X_{m1}	…	X_{mi}	…	X_{mn}

The weight of parameters in the formation of the estimated potential is characterized by a quantitative measure of the degree of confidence in the situation, objectively existing uncertainty and is identified with the probability distribution P_{ji}. Characteristics $\{X_{ji}\}$ are uniquely set in various physical scales. Therefore, to bring the characteristics $\{X_{ji}\}$ to a single common scale, we use the natural normalization, carried out relative to the extreme values $\{X_{ji}\}$ of the components as without changing the ingredient to the opposite:

$$r_{ji} = x_{ji}/x_{max\ j} \quad (1)$$

and with the change of ingredient to the opposite:

$$r_{ji} = x_{min\ j}/x_{\ ji} \quad (2)$$

with display in $x_{ji} \rightarrow r[0,1]$. Dependencies (1) and (2) ensured the mapping of the sample space shown in Table 5 to another having the power of the continuum shown in table 6.

TABLE VI. MODIFIED MORPHOLOGICAL MATRIX

Name of characteristics	**Many dangerous emergencies (violators, objects)**				
	$\{A_1\}$	…	$\{A_i\}$	…	$\{A_n\}$
X_1	r_{11}	…	r_{1i}	…	r_{1n}
X_j	r_{j1}	…	r_{ji}	…	r_{jn}
…	…	…	…	…	…
X_m	r_{m1}	…	r_{mi}	…	r_{mn}

Elements r_{ji} (continuum power spaces) in a single scale will be identified with elementary events. Moreover, the measure defined on r_{ji} the normalized corresponds to the probability *p(r)*, which is identified with the concept of the integral potential of a given complex of elementary events. The meaning of this measure is to appropriately interpret the concept of probability. Moreover, probability, as a category of dialectics, combines both the measure of the objective possibility of an event and the degree of subjective confidence in the occurrence of events. In order to formalize the problem of choosing a solution, we identify many alternatives with the event space *{A}*, and many attributes with events {x}. Then the connection between all the characteristics that form the estimated potential of the emergency is carried out through the normalized measure defined on these characteristics, which is identified with probability *p(a)*. The probability distribution *p(a)* provides an estimate of entropy H_i.

Information in probabilistic-statistical theory acts as a removable, diminished uncertainty. Therefore, further construction of the method is associated with the study of the laws of transformation of the information field of the Cartesian product of two sets (emergency situations and their characteristics) into quantitative components of information. For this purpose, we introduce concepts such as a priori, a posteriori, and conditional probabilities into the logic diagram, apply the Bayesian theorem and the formula for the total probability, and introduce the concept of the conditional *p* probability of the j-th characteristic in the formation of the estimated potential, provided that the events that form the estimated potential have occurred.

To obtain the dependence of determining the quantity *p(r)*, which is a normalized measure on elementary events *{r}*, we use the fact that the concept of the estimated potential of a given complex of elementary events can be identified with the membership function, which associates each *r* with a real number in the interval [0,1]. Moreover, without violating the generality of reasoning, the desired dependence of the membership function is represented in the form:

$$p_{ji}(r) = r_{ji} / \sum_{i=1}^{n} r_{ji} \qquad (3)$$

One of the methods for calculating the probability of manifestation of the j-th characteristic of the compared emergency situations (violators, objects) to the formation of the estimated potential is based on the input V. Khomenyuk concept of the potential probability distribution, which is determined by the formula:

$$\hat{p}_j(r) = \sum_{i=1}^{n} r_{ji} / \sum_{j=1}^{m} \sum_{i=1}^{n} r_{ji} \qquad (4)$$

The principle of the potential probability distribution is based on the fact that it is preferable to choose with greater probability those characteristics of the system whose properties have a large contribution to the total value of the estimated potential. At the same time, we note that for the principle of potential probability distribution, V.V. Khomenyuk, a priori information about the state of characteristics is based on the principle of insufficient knowledge.

However, the weight of various characteristics in the formation of the estimated potential is different. Obtaining estimates of the a priori distribution is related to the order relation p_j, which was studied in detail in the works of Fishborn [6]. For a simple linear relationship, the order of the Fishborn estimates of a priori probabilities form a decreasing arithmetic progression of the form:

$$\check{p}_j = 2 * (m - j + 1)/m * (m + 1)$$

Introducing an a priori probability based on Fishborn estimates, we, setting the "input" to the method, take into account the different weight of the characteristics in the formation of the estimated potential. Then, using the principle of potential distribution (4) and the provisions of the Bayesian theorem, we obtain a logically justified "exit" from the model in the form of posterior conditional probabilities in the form:

$$p_j = \sum_{i=1}^{n} r_{ji} * \check{p}_j / \sum_{j=1}^{m} \sum_{i=1}^{n} \check{p}_j \qquad (5)$$

Thus, the benefit of introducing a priori probability is that it provides an attachment of the information necessary for analysis.

After introducing the a priori probability into the decision-making method and calculating the a posteriori value of the conditional probability p_j, we proceed to the next modeling stage, which is associated with obtaining probabilistic estimates of the influence of the j-th characteristic of the i-th emergency on the hazard potential. To this end, we use the Bayesian theorem, in which we are talking about reversing the order of statements in conditional probability, that is, in the notation we have adopted $p_{ji}(r)$ and are related p_j. Then the probability $p(a)$ in the considered information situation is determined by the dependence:

$$p_{ji}(a) = p_{ji}(r) \cdot p_j / \sum_{i=1}^{n} \quad \sum_{j=1}^{m} p_{ji}(r) \cdot p_j \qquad (6)$$

Then the significance level of the emergency (intruder, object) is estimated through the uncertainty function:

$$H_i(p) = -\sum_{j=1}^{m} p_{ji}(a)\, lg\, p_{ji}\,(a)\,.$$

For the data in Table 4, the entropy potential of each type of intruder was evaluated using an IVM. The results are presented in Figure 1.

Fig. 1. Entropic training potential of violators.

In Table 7, the characteristics of the damage caused by typical violators of the CIO on the same six-point scale were formed in such a way that the entropy assessment of the potentials of preparation (danger) of violators coincided with the entropy assessment of the potentials of the consequences of the target implementation of violators. That is, the entropy estimates in Figures 1 and 2 are uniform according to the Pearson chi-square criterion.

TABLE VII. ASSESSMENT OF THE CONSEQUENCES OF THE TARGET IMPLEMENTATION OF VIOLATORS WITH A SIX-POINT SCALE

Private types of losses from the actions of violators	**The scale of losses from the type of violator**					
	X_1	**X_2**	**X_3**	**X_4**	**X_5**	**X_6**
Political	6	5	4	1	2	3
Human	6	5	4	1	2	3
Financial	3	2	2	3	5	4
Economic	6	5	4	2	2	3
Environmental	6	5	4	1	3	2
Informational	3	2	1	2	5	5

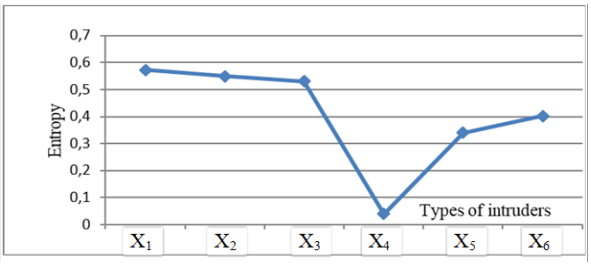

Fig. 2. Entropy potentials of the target implementation of violators.

TABLE VIII. CHARACTERISTICS OF CATEGORIES OF CIO AND TYPICAL VIOLATORS ON THE ENTROPY SCALE

Private types of losses	**Typical violators and categories of objects**												
	X_1	**X_2**	**X_3**	**X_4**	**X_5**	**X_6**	**1-c**	**2-c**	**3-c**	**4-c**	**5-c**	**6-c**	**7-c**
Political	.878	.621	.555	.007	.116	.173	.621	.555	.173	.173	.116	.116	.007
Human	.878	.621	.555	.007	.116	.173	.621	.555	.555	.173	.116	.116	.007
Financial	.173	.116	.116	.173	.621	.555	.621	.621	.555	.173	.116	.116	.007
Economic	.878	.621	.555	.116	.116	.173	.878	.621	.555	.173	.116	.116	.007
Environmental	.878	.621	.555	.007	.173	.116	.878	.621	.555	.173	.116	.116	.116
Information	.173	.116	.007	.116	.621	.621	.878	.621	.555	.173	.173	.116	.116
Entropy potential	.633	.497	.446	.160	.368	.375	.733	.629	.547	.269	.219	.210	.122

On this basis, information on violators and CWS (tables 2 and 7) should be combined into one information field. As a result, we get table 8 with a common information field in a single measurement scale. In table 8, we replace the six-point scale with the entropy scale of the danger of emergency consequences, which was determined using the IWM to the data of table 1 [9] (H is the corresponding entropy value of damage):

1 - local (damage within the territory of the facility) - H = 0.007;

2 - local (damage within the territory of the settlement) - H = 0.116;

3 - territorial (damage to the territory of the constituent entity of the Russian Federation) - H = 0.173;

4 - regional (damage on the scale of two constituent entities of the Russian Federation) - H = 0.555;

5 - state (damage to the limits of more than two constituent entities of the Russian Federation) -H= 0.621;

6 - interstate (damage extends beyond the borders of the Russian Federation) - H = 0.878.

The transition from the six-point scale to the entropy hazard potentials is justified by the requirement to increase the reliability of the damage assessment scale [1].

Applying ICM to the data in Table 8, we obtained the entropy potentials of typical intruders and categorized objects (shown in the bottom row of Table.8 and in fig. 3).

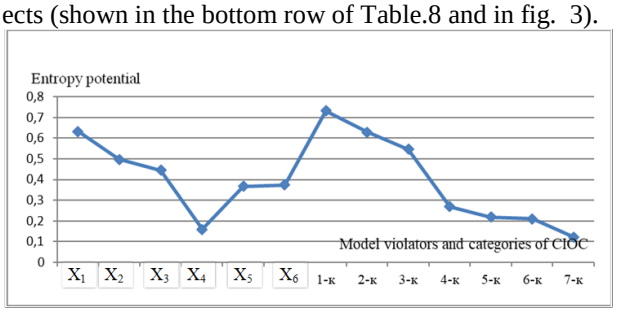

Fig. 3. Graph of the entropy potential of the CIOC and typical violators.

Having solved the problem of combining homogeneous potentials into clusters by the agglomerative method using Statistika 10.0 SPP, we obtained the results of combining typical intruders and categorized KBO into clusters - categories (Table 9). An analysis of Table 9 and Figure 3 shows that the first type of intruder is lower in potential capabilities than an object of the first category, so he needs to combine with internal intruders X5 or X6.

TABLE IX. THE TABLE OF CORRESPONDENCE OF THE BASIC VIOLATORS AND CATEGORIES OF OBJECTS

Typical breaker	**Category CIO**	**Entropy danger H**	**probability of a safe condition**
X_1+(X_5,X_6)	1- st category	0.733	0.98
X_1	2 - nd category	0.633	0.96
X_2	3 - category	0.497	0.93
X_3	3, 4– category	0.446	0.85
X_4	4, 5 - category	0.375	0.60
X_5	4, 5– category	0.368	0.64
X_6	5, 6, 7 - category	0.160	0.68

It should be noted that the preparedness potential is consistent with the hazard potential of the consequences of the actions of violators and with their capabilities to overcome the PPS of objects, i.e. each potential of the intruder can be assigned the corresponding protection potential of the object (PPS potential) - for example, the probability of a safe state of the object. That is, it is required to determine the necessary probability of the safe state of the object depending on the hazard potential of the base intruder for the corresponding hazard category of the object. Obviously, there should be a correspondence between the danger potential of a typical intruder and the degree of protection from his actions, i.e. the nature of the change in the dependencies of the potentials of violators and their counteraction should be similar functions.

If the function of changing the entropy potentials of the type of intruder is related to the required probability of a safe state according to the first type of intruder (the highest estimate is the probability of protection 0.98 - the value is close to the limit), and the weakest type of intruder (the

sensitivity of the detection sensor is 0.7 and the probability of timely arrival - 0.9 gives a value of a safe state of 0.6), i.e. we associate with each type of intruder the required value of the object's security (the probability of a safe state) from its actions. The results of the probabilities of the safe state of the PPS, as similar values to hazards (typical violators), are shown in table 9.

CONCLUSION

The basic typical violators for each category of CIO are defined, which are shown in table 9. The results obtained do not contradict the physical meaning. The results of the probabilities of the safe state of categorized objects (Table 9) can be used to justify the requirements for the effectiveness of the PPS CIO.

REFERENCES

[1] V.N. Kostin, "Assessment of the danger potential of violators based on the informational probabilistic method and the method of principal components," Information technology and computing systems, vol. 3, pp. 74-81, 2016.

[2] A.V. Boyarintsev, A.N. Brazhnik and A.G. Zuyev, "Antiterrorism Issues: Categorization and Vulnerability Analysis of Objects," SPb.: "ISTA–Sistems", 2006, 252 p.

[3] A.S. Borovskiy and A.D. Tarasov, "Automated design and assessment of physical protection systems for potentially hazardous (structurally complex) objects. Part 1. System analysis of the problem of design and assessment of physical protection systems," Samara, Orenburg: SamGUPS, 2012, 155 p.

[4] A.A. Zakharov, A.L. Zhiznyakov and V.S. Titov, "A method for feature matching in images using descriptor structures," Computer Optics, vol. 43, no. 5, pp. 810-817, 2019. DOI: 10.18287/2412-6179-2019-43-5-810-817.

[5] A.A. Agafonov, A.S. Yumaganov and V.V. Myasnikov, "Big data analysis in the geoinformation problem of short-term forecasting of traffic flow parameters based on the method of k nearest neighbors," Computer Optics, vol. 42, no. 6, pp. 1101-1111, 2018. DOI: 10.18287 /2412-6179-2018-42-6-11-11-1-111.

[6] V.N. Kostin, "Assessment of the magnitude of the significance of emergency situations based on the information probabilistic model," Information security problems. Computer systems, vol. 3, pp. 21-32, 2019.

[7] V.N. Kostin and A.K. Ponomarev, "Informational probabilistic method of forming a category of potentially dangerous objects," Bulletin of Computer and Information Technologies, vol. 6, pp. 34-42, 2015.

[8] Order of the Minister of Industry and Energy of the Russian Federation 04.05.2007 No 150 "On the approval of recommendations on anti-terrorist security of industry and energy," 2007, 72 p.

[9] Decree of the Government of the Russian Federation 08/29/2014 No. 875 "On anti-terrorist protection of facilities of the federal service for technical and export control of territorial bodies and subordinate organizations, 2014, 64 p.