Computer Science and Information Systems 19(1):415–433 https://doi.org/10.2298/CSIS210211057K

An Approach for Selecting Countermeasures against Harmful Information based on Uncertainty Management

Igor Kotenko, Igor Saenko, Igor Parashchuk, and Elena Doynikova

St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS),
39, 14th Liniya, 199178, St. Petersburg, Russia
{ivkote, ibsaen, parashchuk, doynikova}@comsec.spb.ru

Abstract. Currently, one of the big problems in the Internet is the counteraction against the spread of harmful information. The paper considers models, algorithms and a common technique for choosing measures to counter harmful information, based on an assessment of the semantic content of information objects under conditions of uncertainty. Methods of processing incomplete, contradictory and fuzzy knowledge are used. Two cases of the algorithm implementation to eliminate the uncertainties in the assessment and categorization of the semantic content of information objects are analyzed. The first case is focused on processing fuzzy data. The second case is based on using an artificial neural network. An experimental evaluation of the proposed technique have shown that the use of both cases makes it possible to eliminate uncertainties of any type and, thereby, to increase the efficiency of choosing measures to counter harmful information.

Keywords: harmful information, assessment, countermeasures, semantic content, information objects, uncertainty.

1. Introduction

At present the Internet and social networks, which can be represented as large sets of interconnected digital network information objects, are becoming one of the most important threats to personal, public and state information security. This determines the need to protect the individual, society and the state from information that spreads through information and telecommunication networks and is capable to harm the health of citizens or motivate them to unlawful behavior. For example, the United States has laws that protect children's Internet and protect children from harmful content posted on the Internet. In the UK, Canada and many other countries, systems are used to block blacklisted sites with harmful content. However, the presence of such systems responsible for blocking harmful content on the Internet and social networks does not mean that the problem of protecting against harmful information has been solved. At the moment, the detection of harmful sites and messages and the formation of black lists is carried out, as a rule, in manual mode.

In scientific and methodological terms, the problem of protection from harmful information has only a small number of scientific and technical solutions. We can say that the methodology for countering harmful information is at the initial stage of

development and implementation. This is fully true for the task of choosing measures to counter harmful information. The solution to this problem should be based on solutions for the development and application of content analysis tools, as well as software and hardware for detecting, evaluating and countering harmful information. At the same time, the concept of harmful (detrimental, dangerous, destructive) information means such information that is prohibited from being distributed on the Internet or social networks by current legislation.

Determination of reliable estimates of digital network content requires that, in order to increase the objectivity of its analysis and make adequate decisions to counter harmful information, data processing is carried out taking into account their uncertainty. This task is of particular relevance for making decisions and choosing specific measures to counter harmful information in real-life conditions. Consequently, models, algorithms and methods for evaluating information objects, as well as choosing means of countering harmful information should underlie the operation of systems for intelligent analytical processing of digital network content. The main purpose of such systems should be the detection, analysis and counteraction of harmful information.

Systems for intelligent analytical processing of network information objects can perform many different functions and consist of many different components. In particular, the components of distributed scanning of information objects, as well as their classification and categorization in accordance with the categories (or types) of harmful information established by law, are mandatory. However, the uncertainty of information available in information objects leads to a significant decrease in the efficiency of these components. Therefore, the component of eliminating the ambiguity of the semantic content of information objects, as well as the component for choosing countermeasures, should also be included among the basic components of such systems. In this regard, the purpose of this work is to clarify the functionality of the uncertainty elimination component and the countermeasures selection component, determine their interrelationships in the analysis of information flows, and develop models, algorithms and methods for selecting measures to counter harmful information based on an assessment of the semantic content of information objects under uncertainty.

The research was firstly presented at International Conference on INnovations in Intelligent SysTems and Applications (INISTA) 2020 [1]. In this paper we detailed and extended the description of countermeasure selection techniques and provided listing of the countermeasure selection algorithm. Besides we have added the second computational experiment for eliminating incompleteness and inconsistency of source data while in the research provided at INISTA 2020 we demonstrated only the first computational experiment that allows eliminating fuzziness.

The paper is organized as follows. Section 2 reviews the related works on selection of countermeasures against harmful information considering uncertainty of observation data. Section 3 provides a general algorithm for uncertainty elimination. Section 4 discusses the methods, models, techniques and algorithms for selection of countermeasures against harmful information. Section 5 describes two computational experiments and obtained results. Section 6 gives conclusions and future research directions.

2. Related Work

Despite the fact that in recent years some solutions on individual components of this kind of protection systems [2-20] have been suggested, they are either at the initial stage of development and implementation, or do not implement the full range of expected capabilities. So, in [2-7] the mechanisms for detecting and counteracting harmful information in network information objects are considered. These papers set out solutions for determining reliable estimates of digital network content. The mechanisms considered in them are based on methods of information classification, methods of intelligent data processing and spam filtering. However, these mechanisms are not focused on working in conditions of semantic uncertainty of information content.
The works [8-10] consider various methods of analyzing social networks to detect and select measures to counter harmful information. In [8], algorithms for searching by event description, identifying users of various networks, and searching by user groups are often used to detect harmful information. Methods for quantitative and qualitative assessment of information impacts in social networks, based on tabular and graphical tools for representing metrics and calculating metrics, are discussed in [9]. In [10] approaches to determine the demographic characteristics of users of social networks are considered. However, since there are other sources of unwanted information in addition to social networks, these approaches cannot be considered universal.

Many works suggest using traffic analysis methods based on the classification of web pages to detect and counteract unwanted information. Thus, in [11], methods based on content analysis of the internal properties of web pages are considered. In [12, 13], it is proposed to use a binary classifier based on identifying groups of internal properties of HTML documents, which is used to train systems for classifying web pages. Methods for training classifiers in order to detect and counter malicious information based on a combination of significant functions of web pages are discussed in [14]. However, all the methods presented in [11-14] are focused on the analysis of web content. Thus, they also cannot cover all sources of harmful information and types of countermeasures.

In some works, it is proposed to implement methods for detecting and countering harmful information based on the results of evaluating the semantic content of information objects using algorithms for classifying web content topics. For example, the papers [15, 16] describe an approach to searching for harmful information based on URL addresses. The advantage of this approach is to reduce the many characteristics of malicious information that need to be analyzed, which entails a reduction in the range of countermeasures. Another popular approach is to use it to analyze links in web content. In [17, 18], based on this approach, a procedure for hierarchical classification of web content is proposed. However, the application of these methods is limited. An interesting method proposed in [19] is the search and extraction of meaningful text from tags with the subsequent application of the classifier to the obtained samples. The same approach in combination with methods of counteracting harmful information is mentioned in [20]. But their application takes a significant amount of time.

Taking into account real conditions in the process of identifying and countering harmful information requires the use of modern approaches, in which the processing and assessment of the properties of harmful information is carried out under conditions of uncertainty. In some works, these approaches are based on methods, models and algorithms for eliminating uncertainty. In [21-24], approaches are considered in which the processing of uncertain information of various types, as well as decision support, are

implemented using artificial neural networks. In [25, 26], it is offered to use fuzzy sets for these purposes. Neural fuzzy networks are proposed to be used for obsolescence of uncertainty in [27]. However, it should be noted that the application of these methods to detect and counteract malicious information is a rather difficult task. On the other hand, a great advantage of these methods is that they allow you to choose measures to counter harmful information based on an assessment of the semantic content of information objects in conditions of uncertainty. These approaches will form the basis of the solutions considered in this paper.

Thus, the analysis of known approaches, methods and solutions for the detection and counteraction of harmful information shows that reliable control of the semantic content of information objects is a complex process that requires the combined use of various mechanisms. However, the task of developing existing methods for choosing measures to counter harmful information is still relevant. Methods for detecting and countering harmful information should be focused on processing poorly formalized (incomplete, inconsistent and fuzzy) data. It is necessary to use additional expert opinions and dynamic (changing) knowledge. The solutions discussed below are focused on the implementation of such approaches.

3. General Algorithm for Eliminating Uncertainty

Information objects are natural language text, media content, embedded parts of other information objects, executable scripts, domain names, IP addresses, etc. Solving the problem of assessing and categorizing the semantic content of information objects is an important stage for detecting harmful information, making decisions and countering harmful information. Elimination of ambiguity (namely, fuzziness, incompleteness and inconsistency) when solving this problem is a necessary condition for its successful solution in real subject areas, when the initial data are exposed to various factors of uncertainty. Such factors, for example, include noise when measuring, modeling, or observing the attributes of the semantic content of information objects. These attributes can be textual, graphic, numeric, logical, ordinal, nominal, etc. Among the various types of uncertainty, the most significant are ambiguity (fuzziness) and insufficiency (incompleteness, inconsistency) of the initial data.

Uncertainty inherent in the initial data of the tasks of detecting, evaluating and making decisions on counteracting harmful information can arise due to the non-stationarity of the information flow, fuzziness, incompleteness and inconsistency in identifying features of harmful information, the dynamics of the security system, the impact of destabilizing (often antagonistic) environmental factors, and also due to the presence of ambiguity of goals and the inconsistency of the tasks of detecting and countering harmful information.

The common algorithm for eliminating the uncertainty of the initial data for the problem of identifying harmful properties of information will be called the algorithm for eliminating uncertainty for assessing and categorizing the semantic content of information objects based on methods of processing fuzzy, incomplete and contradictory knowledge. This algorithm includes the following steps:

1. Input undefined harmful information's features and type of uncertainty (fuzziness or incompleteness and inconsistency).

2. If harmful information's features are fuzzy specified go to step 3 (steps 3-7 allow the experts to specify weight matrices of harmful information's features in advance), otherwise go to step 10.
3. Specification of the fuzzy harmful information's features system and initial membership functions of fuzzy sets.
4. Matching of the expert opinions on adding of the specific information objects' semantic content features to the harmful information's features set.
5. If there is one expert (i=1), go to 9 else go to 6.
6. Specification of the membership functions of fuzzy sets by the next expert (i=i+1).
7. Calculation of the common experts' opinion on adding of the specific information objects' semantic content features to the harmful information's features set (disjunctive sum).
8. If there is the next expert, go to 6 else 9.
9. Final choice of the specific information objects' semantic content features for the harmful information's features set (based on the max of preference function).
10. Generation of weight matrices W_m and W_{Y} for two-layer artificial neural network.
11. Activation of artificial neural network's input layer.
12. Initial setting of neurons of artificial neural network's output layer.
13. Bringing of input layer neurons to the state of second layer neurons.
14. Calculation of states of output layer neurons.
15. If artificial neural network is stable go to step 16 else go to step 13.
16. Sum values of weight coefficients.
17. Final choice of the specific information objects' semantic content features for the harmful information's features set (based on the max value of elements of sum output weight coefficients' vector).
18. Output the results: final harmful information's features considering uncertainty elimination in scope of fuzzy and conflicting knowledge processing.

The algorithm for eliminating uncertainty in the assessment and categorization of the semantic content of information objects is based on the use of the mathematical theory of fuzzy sets [25, 26]. The central link of the algorithm is a decision support mechanism for adding the analyzed fuzzy characteristics of information in digital network content to the set of features of malicious information (steps 3–9 of the algorithm). The criterion for the harmfulness of the processed semantic content in its assessment and classification is the excess of the features of dubious information characteristics in the semantic content of information objects of the threshold value (α-level of the preference function). The analyzed attributes of the semantic content include the presence, quantity, or nomenclature (severity) of questionable informational characteristics. The identification of the fuzzy character of harmful information is implemented based on the opinions of experts. To determine the subjective measure of confidence that this information belongs to a fuzzy set of characteristics of malicious information, membership functions are used. To combine several subjective confidence measures (that is, the opinions of several experts), mathematical operations of addition, union, intersection and disjunctive sum of fuzzy sets are used.

The algorithm also implements the elimination of uncertainty in the assessment and categorization of the semantic content of information objects using artificial neural networks. The neural network mechanism [21-24] for searching and predicting the

relationship between the features of the semantic content of information objects is specified in steps 10-17 of the algorithm. The purpose of this mechanism is reasonable adding of the analyzed incomplete and inconsistency features of information within digital network content to harmful information's features set. The principle of operation of this mechanism is that if there is at least one function that is guaranteed to be included in the set of functions of malicious information, it is possible to generate an input set of functions that takes into account incomplete and incompatible relationships of all functions. After that, an artificial neural network is put into operation. With its help, it is possible to obtain at the output a set of features of malicious information with coefficients (elements) characterizing the weight (severity) of features of malicious information. These coefficients make it possible to evaluate and classify dubious information as harmful.

The output result of the algorithm is a given system of features of the semantic content of information objects, which uniquely determines whether or not a particular information is harmful. Thus, the developed algorithm makes it possible, in conditions of uncertainty, to identify harmful information, as well as to form the initial data for making a decision on counteracting harmful information.

4. Models, Algorithm and Technique for Selecting Countermeasures

This section describes the developed technique for counteracting against harmful information and the related models and algorithm.

The technique is based on the decision-making theory, including multi criteria optimization methods. Input of the technique is as follows: (1) harmful objects; (2) available countermeasures. The main stages of the technique are: (1) generating models of information objects, information system, countermeasures and counteracting process; (2) selecting countermeasures. The output of the technique is the set of selected countermeasures.

We specified the following models considering subjects and objects participating in countermeasure selection process: information system model, information object model, countermeasure model, model of counteracting against harmful information.

An information system *IS*, where counteracting is implemented, is Internet. It incorporates objects *IO* and communication means *IC*. Thus, information system model is specified as follows: $IS = (IO, IC)$. The objects can be physical or informational. The developed model is limited with information objects. Communication is interaction between two or more objects and/or subjects. Communication between the information objects can be determined based on the existing references. In the developed model a subject can be represented as some information object, for example, a profile in social network or on some web site (in this case communication is determined based on the profile's friends and groups), or as a property of information object, for example a counter of visitors for the web page. Communication can be physical or logical. The developed model is limited with logical communications represented as follows: $IC \subseteq IO \times IO$.

We specify information object *IO* as follows:

$$IO = <size, role, hltype, type, state, ioaud, saud>,$$

where

size – a size of information object, it can take values {*s*, *m*, *l*}, where *s* – small, *m* – medium, *l* – large;

role – a role of information object, it can take values {*s*, *r*, *u*}, where *s* – sender, *r* – recipient, *u* – user;

hltype – a type of information object, it can take values {*h*, *n*}, where *h* – harmful objects, *n* – not harmful objects;

type – a more detailed type of information object, it can take values {*ter*, *hea*, *por*, *dru*, *cru*, *none*}, where *ter* – an information object containing public calls for terrorism and extremism; *hea* – the information objects containing information harmful for people's (especially children's) health, and moral and spiritual development; *por* – an information object with pornography propaganda; *dru* – the information objects containing information on ways of development, production and use of drugs and committing suicide, as well as swearing; *cru* – an information object containing direct calls for violence and cruelty (e.g. war), ethnic and religious hatred, or hostility in the content information; *none* – not applicable (if *hltype* is *n*);

state – a state of compromise of information object in case of harmful information impact, it can take values {*compr*, *nonc*}, where *compr* – object is compromised by the harmful information, *nonc* – object is not compromised;

ioaud – audience of the information object that is an array of links on information objects that are linked with the sender by communications and that are recipients of the objects (can be null);

saud – real number (if there is a counter of visitors of information object) or expert assessment of subjects who are the recipients of the object (can be 0).

Information objects can be classified into small objects IO_s, medium objects IO_m and large objects IO_l depending on their size. We consider post, message, comment, media object, etc. as IO_s, web page, group, channel, etc. as IO_m and information system, web site, social network, messenger, etc. as IO_l. These classes are related as follows: $IO_s \subset IO_m \subset IO_l$.

Each information object has a role. The roles are as follows: sender (subset *Rs* of *IO*), recipient (subset *Rr* of *IO*), and user (subset *Ru* of *IO*). *Rs* propagate harmful (or not) information. *Ru* incorporates all information objects that are connected with the considered information objects via *IC* and form the audience *Au* of information object (*ioaud*). Some users represented with information objects are the recipients of information object: $Rr \in Ru$, $Rr \in Au$. A subset *Rr* can be empty. Other users that form the rest part of the *Au* get information object unintentionally.

Information object can be harmful or not. Harmful object if its role is sender (namely *role*=*s* and *hltype*=*h*) affects the audience as follows:

1. Audience compromise state becomes *compr* (this is relevant for *ioaud* and *saud*).
2. Harmful information propagation, i.e. the part of the audience $Rr \in ioaud$ becomes senders *Rs*. While some information objects counted by *saud* can also become senders, it is difficult to trace if they have propagated harmful information.

Thus, harmful object affects information system state as follows: $\{IO^{k_i}, IC^{k_i}\}$ becomes $\{IO^{k_j}, IC^{k_j}\}$, where k_i – previous harmful object number, k_j – current harmful

object number. For *m* known information objects from IO^{kj} (*m*=[0;*M*], where *M* – number of elements in *ioaud* of this harmful object) the following parameters are changed: *role* becomes *s*, *hltype* becomes *h*, *type* becomes *ter*, *hea*, *por*, *dru* or *cru* (depending on considered harmful object type), and *state* becomes *compr*. It should be noticed that number of compromised information objects are $saud_{\Sigma} = |ioaud|_{k_i} + saud_{k_i} + |ioaud|_{k_j} + saud_{k_j}$, while $saud_{k_j} = \sum_{m} saud$.

The countermeasures should eliminate impacts on the audience and stop harmful information propagation. The countermeasures can be taken against:

1. Information object with sender role. Such measures should be taken if an audience of the harmful object is huge and information object is contrary to the laws of the country. In this case the following countermeasures can be taken: removal (or block); a warning.
2. Information objects from the audience of harmful information object. Such measures can be taken in case of low popularity of the sender to stop harmful information propagation or prevent access to harmful information. In this case the countermeasure of informing type can be taken. Thus, to stop harmful information propagation a warning about an illegality of content and responsibility for its distribution can be provided; to prevent access to harmful information a warning that content is for the appropriate age category can be provided.

We specify countermeasure *rm* from *RM* (set of countermeasures) as follows:

$$rm=<rm\_class, rm\_type, rm\_cost, rm\_ef>,$$

where

rm_class – class of countermeasure (barrier, disguise, informing or enforcement);
rm_type – size of the information object (small, medium, or large);
rm_cost – countermeasure cost;
*rm_rol*e – role of information object (sender or recipient);
rm_ef – efficiency of the countermeasure;
rm_cd – collateral damage from the countermeasure implementation.

Countermeasure cost is represented by the weight that depends on the countermeasure intrinsic cost, cost of implementation and maintenance, considering complexity of implementation and maintenance and required resources.

Efficiency of the countermeasure is represented as weight that depends on the ratio of recipients that won't be compromised in case of countermeasure implementation to the common number of recipients that would be compromised otherwise.

Collateral damage is represented as weight that depends on additional losses in case of countermeasure implementation, for example financial losses in case of web site blocking.

The countermeasure model is used to specify counteracting model. The countermeasure affects information system state: {*IO*, *IC*} become $\{IO^l, IC^l\}$, where *l* – countermeasure number. For *j* information objects from IO^l (*j*=[0;*N*], where *N* –number of elements in *ioaud* of the harmful object that were affected by the countermeasure), an information object is deleted or its following parameters are modified: *role* become *r* or *u*, *hltype* become *n*, *type* become *none*, and *state become nonc*. For *d* connections from IC^l (*d*=[0;*D*], where *D* – number of links between harmful object and connected objects

that were affected by the countermeasure), an information connection is deleted. Besides, $saud_{\Sigma}$ decreased.

The specified models are used to formalize the algorithm of counteracting against harmful information. We set the following requirements to the counteraction algorithm: (1) it should consider size of harmful information; (2) it should consider harmful information audience (size and age); (3) it should select the countermeasures that provide maximum efficiency and have minimum cost.

In scope of the counteraction algorithm development the input data for countermeasure selection (that are the output data of the previous stages) are as follows: size of information object (*size* parameter of object's model), role of information object (*role* parameter of object's model), high level type of information object (*hltype* parameter of object's model), and detailed type of information object (*type* parameter of object's model). Thus, we have information to satisfy the first requirement.

To satisfy the rest two requirements to the counteraction algorithm some additional information is required. Thus, to consider harmful information's audience (information objects that are linked with the sender *ioaud* and not linked recipients of the object *saud*) additional harmful information propagation algorithm is developed. It is based on search of linked objects and changing of their *state* to *compromised*. Size and age of the harmful information's audience is calculated considering these compromised objects and their traffic (using counters).

Countermeasure efficiency (that is specified in the countermeasure model as *rm_ef*) is calculated as ratio of recipients that won't be compromised in case of countermeasure implementation, both number of information objects from *ioaud* with *state compr* and *saud*, to the common number of recipients that would be compromised otherwise. While countermeasure cost (*rm_cost* in the countermeasure model) is specified by the experts manually. Besides, in counteraction algorithm class of countermeasure (that is selected depending on the harmful information type) and size of the information object are considered. The pseudocode of the counteraction algorithm is provided below:

```
1. Input io from IO where state=compr, hltype=h, role=s.
2. Input io class, io type.
3. Calculate direct ioaud for io.
4. Determine saud size, saud age for io.
5. Calculate propagated ioaud for io.
6. Determine propagated saud size, saud age for io.
7. Determine cms for io considering role, class, type, propagated
ioaud, propagated saud size, propagated saud age, rm_class,
rm_type.
8. For each countermeasure c from cms:
8.1.     Determine rm_cost.
8.2.     Determine rm_ef.
9. Select scms from cms with min rm_cost and max rm_ef.
10. Output scms.
```

Step 7 of the algorithm above is implemented based on the set of rules. The following classes of countermeasures can be outlined:

1. Barrier, namely, filtering of information objects and blocking of information sources using software. This measure can be implemented by the information object that has sender role (e.g. filtering of messages on the web site) and recipient role (e.g. parental control software, filtering options within operation system) as well.

2. Disguise (or distraction) can be implemented on the part of sender by adding distracting content, e.g. message or picture.
3. Informing, should be implemented on the sender part to motivate the recipients to avoid information object. For example, it can be implemented using warning message about illegality of the content, or age category of content.
4. Enforcement are the measures implemented as the result of laws, such as deleting of information or user blocking that can be implemented on the sender side, or web-site blocking that can be implemented on the domain management level.

On step 7 of the algorithm the rule-based technique for countermeasures list determination is used. It outputs the set of possible countermeasures *cms* considering *role* of information object (sender *s*, recipient *r*, or user *u*), *size* of information object (small *s*, medium *m*, large *l*), *type* of information object (public calls for terrorism and extremism *ter*, information harmful for people's health *hea*, pornography propaganda *por*, information on ways of development, production and use of drugs and committing suicide *dru*, direct calls for violence and cruelty *cru*, or *none*), and total audience size *ioaud* and *saud* and age, and *rm_class* (barrier, disguise, informing or enforcement) and *rm_type* (*size* of the information object *small*, *medium*, or *large*). Examples of rules for step 7:

- "if *role* = *s* and *size* = *s* and *type* = *ter* and total audience size<3000 and age>18 select countermeasures where *rm_class=disguise* or *informing* and *rm_type=small*";
- "if *role* = *u* and *size* = *s* and *type* = *hea* and total audience size<3000 and age>18 select countermeasures where *rm_class=barrier* or *informing* and *rm_type=small*".

If more than one countermeasure *c* is selected, on steps 8-9 of the proposed algorithm the multicriteria optimization is used.

5. Experiments and Discussion

In order to test the feasibility of implementing the proposed algorithms for eliminating uncertainties and choosing countermeasures, a computational experiment was carried out to refine the features of malicious information based on the mathematics of fuzzy sets. Below we consider the operation of a branch of this algorithm, which uses methods for processing fuzzy knowledge (namely, calculating a disjunctive sum).

There is initial set of fuzzy specified harmful information's features. The expert opinions are specified. That is initial membership functions of fuzzy sets that characterize preliminary, fuzzy specified harmful information set, for example:

$$\tilde{\tilde{X}} = [\Delta\tilde{\tilde{x}}_{chil}|\mu(\Delta\bar{x}_{\text{chil}});\ \Delta\tilde{\tilde{x}}_{\text{terr}}|\mu(\Delta\bar{x}_{\text{terr}});\Delta\tilde{\tilde{x}}_{\text{porn}}|\mu(\Delta\bar{x}_{\text{porn}});\ \Delta\tilde{\tilde{x}}_{\text{drug}}|\mu(\Delta\bar{x}_{\text{drug}});\ \Delta\tilde{\tilde{x}}_{\text{war}}|\mu(\Delta\bar{x}_{\text{war}})]^{\text{T}}, \quad (1)$$

where

$\Delta\bar{x}_{\text{chil}}(k)$ – abnormal deviation of the average amount of information harmful for people's (especially children's) health, and moral and spiritual development;

$\Delta\bar{x}_{\text{terr}}(k)$ – abnormal deviation of the average amount of information containing public calls for terrorism and extremism in traffic;

$\Delta\bar{x}_{\text{porn}}(k)$ – abnormal deviation of the average amount of information with pornography propaganda;

$\Delta\bar{x}_{\text{drug}}(k)$ – abnormal deviation of the average amount of information containing data on ways of development, production and use of drugs and committing suicide, as well as swearing; and

$\Delta\bar{x}_{\text{war}}(k)$ – abnormal deviation of the average amount of direct calls for violence and cruelty, ethnic and religious hatred, or hostility in the content information;

μ– membership function of fuzzy set, can take values from 0 to 1.

The disjunctive sum of two fuzzy sets $\widetilde{A}$ and $\widetilde{B}$ characterizing the opinions of the first and second experts, accordingly, is specified using unions and intersections as follows:

$$\tilde{A} \oplus \tilde{B} = (\tilde{A} \cap \bar{\tilde{B}}) \cup (\bar{\tilde{A}} \cap \tilde{B}), \tag{2}$$

where $\bar{\tilde{A}}$ and $\bar{\tilde{B}}$– complements of these fuzzy sets.

The membership function for *j*-th harmful information's feature looks as follows:

$$\forall x_j \in \overline{1,\ldots,5:} \quad \mu_{\tilde{A}\oplus\tilde{B}}(x_j) = \max\{[\min\{\mu_{\tilde{A}}(x_j),\ 1-\mu_{\tilde{B}}(x_j)\}];$$
$$[\min\{1-\mu_{\tilde{A}}(x_j),\ \mu_{\tilde{B}}(x_j)\}]\}.$$

Opinion of the first expert (A) about the assessment and categorization of each feature from the listed above (1) as harmful information's feature can be represented as fuzzy set:

$$\tilde{A} = \{\Delta\tilde{\bar{x}}_{chil}|0{,}3;\ \Delta\tilde{\bar{x}}_{\text{terr}}|0{,}1; \Delta\tilde{\bar{x}}_{porn}|0{,}1;\ \Delta\tilde{\bar{x}}_{drug}|0{,}5;\ \Delta\tilde{\bar{x}}_{war}|0{,}2\}.$$

Opinion of the second expert (B) about the assessment and categorization of each feature from the listed above as harmful information's feature can be represented as similar fuzzy set:

$$\tilde{B} = \{\Delta\tilde{\bar{x}}_{chil}|0{,}7; \Delta\tilde{\bar{x}}_{\text{terr}}|0{,}9; \Delta\tilde{\bar{x}}_{porn}|0{,}4; \Delta\tilde{\bar{x}}_{drug}|0{,}5;\ \Delta\tilde{\bar{x}}_{war}|0{,}4\}.$$

The complements of these fuzzy sets are as follows:

$$\bar{\tilde{A}} = \{\Delta\tilde{\bar{x}}_{chil}|0{,}7; \Delta\tilde{\bar{x}}_{\text{terr}}|0{,}9; \Delta\tilde{\bar{x}}_{porn}|0{,}9; \Delta\tilde{\bar{x}}_{drag}|0{,}5; \Delta\tilde{\bar{x}}_{war}|0{,}8\};$$

$$\bar{\tilde{B}} = \{\Delta\tilde{\bar{x}}_{chil}|0{,}3;\ \Delta\tilde{\bar{x}}_{\text{terr}}|0{,}1; \Delta\tilde{\bar{x}}_{porn}|0{,}6; \Delta\tilde{\bar{x}}_{drug}|0{,}5; \Delta\tilde{\bar{x}}_{war}|0{,}6\}.$$

Intersections of these fuzzy sets are as follows:

$$\tilde{A} \cap \bar{\tilde{B}} = \{\Delta\tilde{\bar{x}}_{chil}|0{,}3;\ \Delta\tilde{\bar{x}}_{\text{terr}}|0{,}1; \Delta\tilde{\bar{x}}_{porn}|0{,}1;\ \Delta\tilde{\bar{x}}_{drug}|0{,}5;\ \Delta\tilde{\bar{x}}_{war}|0{,}2\};$$

$$\bar{\tilde{A}} \cap \tilde{B} = \{\Delta\tilde{\bar{x}}_{chil}|0{,}7;\ \Delta\tilde{\bar{x}}_{\text{terr}}|0{,}9; \Delta\tilde{\bar{x}}_{porn}|0{,}4;\ \Delta\tilde{\bar{x}}_{drug}|0{,}5;\ \Delta\tilde{\bar{x}}_{war}|0{,}4\}.$$

Finally, a union of these fuzzy sets will give the results of disjunctive summation. These results characterize aggregated opinion of two experts about the assessment and categorization of each feature from the listed above as harmful information's feature:

$$\tilde{A} \oplus \tilde{B} = (\tilde{A} \cap \bar{\tilde{B}}) \cup (\bar{\tilde{A}} \cap \tilde{B}) =$$

$$= \{\Delta\tilde{\bar{x}}_{chil}|0{,}7;\ \Delta\tilde{\bar{x}}_{\text{terr}}|0{,}9; \Delta\tilde{\bar{x}}_{porn}|0{,}4; \Delta\tilde{\bar{x}}_{drug}|0{,}5; \Delta\tilde{\bar{x}}_{war}|0{,}4\}.$$

If there are more than two experts, an opinion of the third expert is specified. The aggregated opinion of two previous experts is used as one opinion and all cycle is repeated till there are experts. As the result we get aggregated opinion of experts based on fuzzy knowledge processing.

Let us introduce threshold value of membership function describing a preference of adding the information object' semantic content features to the set of harmful information's features as $\mu^{\text{тр}} \geq 0{,}6$.

Further, for the purpose of the final selection of the features of the semantic content of information objects and their inclusion in the set of characteristics of malicious information, the maximum preference function is used.

The membership function value chart describing a criterion for assessment and categorization in fuzzy conditions is represented in Fig. 1. The results of experimental computations show that the fuzzy sets math allows eliminating this type of input data uncertainty while assessing and categorizing information objects' semantic content using the fuzzy knowledge processing methods.

The state of values of membership functions (for the considered example) should be interpreted as a forecast of the guaranteed preference of adding the specific content feature to the set of harmful information's features.

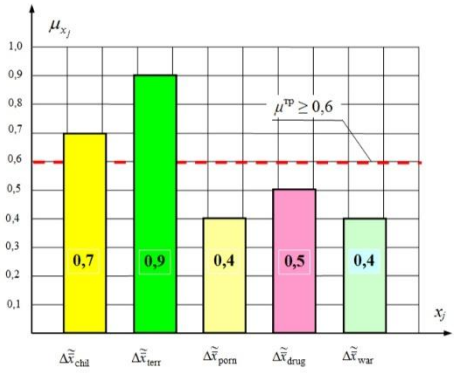

Fig. 1. The results of computational experiment

This state (see Fig. 1) for the *k*-th step of analytical processing of digital network content and for the considered example is characterized by the low preference of adding the following features to the harmful information set:

$\Delta\bar{x}_{\text{porn}}(k)$ – abnormal deviation of the average amount of information with pornography propaganda;

$\Delta\bar{x}_{\text{drug}}(k)$ – abnormal deviation of the average amount of information containing data on ways of development, production and use of drugs and committing suicide, as well as swearing; and

$\Delta\bar{x}_{\text{war}}(k)$ – abnormal deviation of the average amount of direct calls for violence and cruelty, ethnic and religious hatred, or hostility in the content information.

These directions (pornography, drugs and war propaganda) in the content do not exceed the threshold now and they are not harmful. Severity preference is given to the following features as to the most harmful (for the considered example):

$\Delta\bar{x}_{\text{chil}}(k)$ – abnormal deviation of the average amount of information harmful for people's (especially children's) health, and moral and spiritual development;

$\Delta\bar{x}_{\text{terr}}(k)$ – abnormal deviation of the average amount of information containing public calls for terrorism and extremism in traffic.

These are an abnormal deviation of the average amount of information harmful for children's health and average amount of information containing public calls for terrorism. The countermeasures should be implemented against these threats. The calculations were implemented for sample data. They characterize weight of specific feature in the tasks of harmful information detection and counteraction.

Another example of implementation of the proposed algorithms (for uncertainty elimination and countermeasure selection) is the second computational experiment on specification of harmful information features on the basis of mathematical algorithms of artificial neural networks theory.

Let us to consider an example of operation of the second branch of common algorithm for uncertainty elimination. This branch operates using methods of processing of incomplete and conflicting data using artificial neural networks [21-24].

In scope of implementation of this branch of the common algorithm we use the neural network based mathematical procedure for elimination of incompleteness and inconsistency of assessment and categorization of information objects' semantic content features. Two-layer artificial neural network is the basis of the second branch of uncertainty elimination algorithm (steps 10-17 of the algorithm described in Section 3). This branch is developed to search and forecast interconnections between the information objects' semantic content features. As a result, it allows taking the reasonable decision on including (or not) the analyzed incompletely or contradictory specified features of information circulating in digital web content into the set of features of harmful information.

Mathematical essence of the neural network-based branch of the common algorithm for uncertainty elimination (steps 10-17 of the algorithm described in Section 3) is as follows. We determine at least one feature guaranteed to be included in the set of features of harmful information, first. We construct input feature vector $\{\vec{Y}^{i}_{\text{inp}}\}$ using two-layer artificial neural network. This vector $\{\vec{Y}^{i}_{\text{inp}}\}$ considers incomplete and conflicting interconnections of all features (based on opinion of E experts). We get output harmful information's feature vector with coefficients (elements) characterizing weight (severity) of these features based on the results of solving the problem of neural network transformation (steps 10-17 of the common algorithm). The results of these computations allow assessing and categorizing information as harmful on step 18 of the common algorithm (considering incompleteness and inconsistency of input data).

The proposed model for selecting "important" (sensitive) harmful information's features in the conditions of incompleteness and inconsistency allows filtering the subjective values and obtain knowledge empirically based on experts' opinion.

Let empirical data have the form of a protocol:

$\{\vec{Y}^{i}_{\text{inp}}, \quad i = 1, \ \dots \ , \ E\}$,

where vector $\vec{Y}^{i}_{\text{inp}} = (Y^{i}_{\text{inp }1}, Y^{i}_{\text{inp }2}, \ldots, Y^{i}_{\text{inp }P})$ is vector of input features (in terms of artificial neural networks it is input vector *A*) that considers incomplete and conflicting interconnections of all $j = 1, \ldots, P$ harmful information's features according to the opinion of *i*-th expert from the set *E* of experts.

The vector characterizing "importance", for example, for each of 5 (five) previously considered harmful information's features can be common illustrative example:

$\vec{Y}^{1}_{\text{inp}} = (1, 0, 0, 1, -1)$.

This vector is character representation of the expression: "According to the opinion of the first expert "importance" of the harmful information's features is as follows: the first feature $Y_{\text{inp }1}$ (its physical meaning is $\Delta\bar{x}_{\text{chil}}$) and the fourth feature $Y_{\text{inp }4}$ (its physical meaning is $\Delta\bar{x}_{\text{drug}}$) are "important" (sensitive/valuable), the fifth feature $Y_{\text{inp }5}$ (its physical meaning is $\Delta\bar{x}_{\text{war}}$) is not "important" (not sensitive), for the rest features of harmful information (the second and the third) $Y_{\text{inp }2}$ (its physical meaning is $\Delta\bar{x}_{\text{terr}}$) and $Y_{\text{inp }3}$ (its physical meaning is $\Delta\bar{x}_{\text{porn}}(k)$) an opinion of the first expert is absent (it is equal to 0)".

For our computational experiment, assume that at a given time the feature $Y_{\text{inp }5}$ (the unit element of the input vector *A*) is guaranteed "important" (valuable/sensitive) feature of harmful information. This feature characterizes $\Delta\bar{x}_{\text{war}}$ – abnormal deviation of the average amount of direct calls for violence and cruelty, ethnic and religious hatred, or hostility in the content information. Other features of harmful information are undetermined. To get reasonable results of semantic content assessment and detect harmful information it is required to reconstruct the rest components of vector of "important" (valuable/sensitive) harmful information's features. In process of operation the two-layer artificial neural network reconstructs the rest components of vector *A*. Let us to consider this process with an example.

Suppose we are interested in the components of the vector characterizing "importance" of all harmful information's features considering that the fifth feature is obligatory for inclusion in the list of "dangerous" features, i.e. $Y_{\text{inp }5}$ value characterizing "importance" of this feature is equal to "1". Let us to pre-normalize the increments of all features relative to the scale of the activation function. Let the activation function have a stepwise form:

$$f(Y_{\text{inp}}) = \begin{cases} 1, & Y_{\text{inp}} \geq 1; \\ 0, & 0 \leq Y_{\text{inp}} < 1 \\ -1, & Y_{\text{inp}} < 0. \end{cases}.$$

Then $Y_{\text{inp }5}$ value characterizing "importance" of harmful information's feature $\Delta\bar{x}_{\text{war}}$ will correspond to the output value of fifth neuron that is equal to 1, while input vector will take the form $A = (0, 0, 0, 0, 1)$. In other words, the two-layer artificial neural network takes as input $Y_{\text{inp}} = (0, 0, 0, 0, 1)$.

Then, considering mathematical essence of the second neural network based branch of the common algorithm for uncertainty elimination (steps 10-17 of the algorithm described in Section 3), the output vector $B = (b_1, b_2, b_3, b_4, b_5)$ of the two-layer artificial neural network consequentially takes the following values:

$B(0) = f([0; 0; 0; 0; 1]) = [0, 0, 0, 0, 1]$;

$B(1) = f([0{,}667; -0{,}333; 1; 1; 0]) = [0, -1, 1, 1, 1]$;

$B(2) = f([3; -0{,}667; 4; 4; 7]) = [1, -1, 1, 1, 1]$;

$B(3) = f([3; -1{,}667; 4{,}667; 4{,}333; 7{,}667]) = [1, -1, 1, 1, 1]$;

$B(4) = f([3; -1{,}667; 4{,}667; 4{,}333; 7{,}667]) = [1, -1, 1, 1, 1]$;

$B(5) = f([3; -1{,}667; 4{,}667; 4{,}333; 7{,}667]) = [1, -1, 1, 1, 1]$.

The obtained results characterize intermediate and final dependencies of harmful information's features weight ("importance"/value/severity), i.e. characterize total preference (from the experts' point of view) of including of these features, that should be assessed in scope of detection and counteraction against harmful information, into the set of dangerous features. These results can be represented graphically as a diagram (Fig. 2).

As it can be seen from the diagram (Fig.2) the two-layer artificial neural network designed in the interests of evaluating the semantic content for the search and detection of harmful information, has stabilized after the third tact (step). Thus, using such artificial neural network containing two layers of neurons it is possible to implement assessment and short term normative weight ("importance"/value/severity) forecasting for harmful information's features in the conditions of incompleteness and inconsistency of input data.

The results of solving of the second computational experiment (example) allow constructing the vector of sensitive for the given conditions harmful information's features with a high degree of objectivity using accumulated in the neural network data. They allow selecting the volume and nomenclature of harmful information's features for including into the set of dangerous features mathematically correct and as objectively as possible. Moreover, the set of dangerous, obvious features constructed in the interests of detecting and counteracting harmful information, will be guaranteed to include such "important" (sensitive) features as $Y_{\text{inp }1}$ (its physical meaning is $\Delta\bar{x}_{\text{chil}}$), $Y_{\text{inp }3}$ (its physical meaning is $\Delta\bar{x}_{\text{porn}}(k)$), $Y_{\text{inp }4}$ (its physical meaning is $\Delta\bar{x}_{\text{drug}}$) and $Y_{\text{inp }5}$ (its physical meaning is $\Delta\bar{x}_{\text{war}}$) and won't include the feature $Y_{\text{inp }2}$ (its physical meaning is $\Delta\bar{x}_{\text{terr}}$).

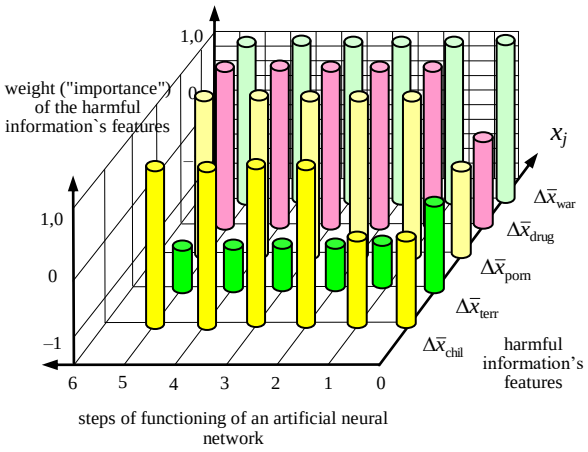

Fig. 2. Graph of the dependence of the weight ("importance") of the harmful information`s features on the cycle (step) of calculation new states of neurons of the output layer

Thus, the second computational experiment was considered. This experiment is based on the second neural network based branch of the common algorithm for uncertainty elimination (steps 10-17 of the algorithm described in Section 3). The experiment demonstrated that this algorithm allows eliminating incompleteness and inconsistency of source data. This distinguishes it from the first branch of the common algorithm for uncertainty elimination (steps 3-9 of the algorithm described in Section 3) that is considered in the first computational experiment and allows eliminating fuzziness.

The results of the computational experiments demonstrate that application of both branches of the common algorithm (described in Section 3) allows eliminating uncertainty of any type while constructing the set of dangerous explicit features for decision making in order to detect and counteract against harmful information.

6. Conclusion

The paper proposed a novel approach to developing the methodological foundations for harmful information's features assessment and decision making on counteracting against harmful information propagation considering uncertainty in data observations. These tasks were specified, and two variants of implementation of harmful information's countermeasures selection process were introduced. The stages of common algorithm for uncertainty elimination while assessing and categorizing information objects' semantic content using methods for fuzzy, incomplete and conflicting knowledge processing are specified for determination of input data for harmful information counteracting task. Thus, the common scheme of the process of eliminating uncertainty in semantic content of information objects and the selection of countermeasures against harmful information were described within the framework of

the common architecture of the system for intelligent analytical processing of network content.

We developed the models, algorithm and technique for harmful information's countermeasures selection on the basis of the proposed scheme for uncertainty elimination. The techniques include the information system model, harmful information counteracting model (including countermeasure model) and the harmful information counteracting algorithm. For the countermeasure selection we use traditional decision support theory methods and multicriteria optimization methods.

On the basis of the proposed scheme for eliminating uncertainty, models, an algorithm and a technique for selecting means of countering harmful information were developed. These solutions include an information system model, a harmful information counteracting model (including a countermeasure model), and a harmful information counteracting algorithm. To select countermeasures, traditional methods of decision support theory and methods of multicriteria optimization are used.

The future research will be devoted to enhancement of the developed algorithms and tools for harmful information's countermeasures selection. It is planned to make the proposed algorithms more universal, so that they would allow evaluating the characteristics of harmful information and choosing countermeasures taking into account both non-stochastic and probabilistic uncertainties.

Acknowledgment. The reported study was partially funded by RFBR project 18-29-22034 mk and by the budget project 0073-2019-0002.

References

1. Parashchuk, I., Doynikova, E., Saenko, I., Kotenko, I.: Selection of Countermeasures against Harmful Information based on the Assessment of Semantic Content of Information Objects in the Conditions of Uncertainty. In Proceeding of the 2020 International Conference on Innovations in Intelligent Systems and Applications (INISTA 2020), Novi Sad, Serbia, 1-7. (2020)
2. Vaismoradi, M., Turunen, H., Bondas, T.: Content Analysis and Thematic Analysis: Implications for Conducting a Qualitative Descriptive Study. Nursing & Health Sciences, Vol. 15, No. 3, 398-405. (2013)
3. Elo, S., Kyngas, H.: The Qualitative Content Analysis Process. Journal of Advanced Nursing, Vol. 62, No. 1, 107-115. (2008)
4. Krippendorff, K.: Content Analysis: An Introduction to Its Methodology. Sage Publications, California, USA. (2004)
5. Graneheim, U. H., Lundman, B.: Qualitative Content Analysis in Nursing Research: Concepts, Procedures and Measures to Achieve Trustworthiness. Nurse Education Today, Vol. 24, No. 2, 105-112. (2004)
6. Pashakhanlou, H.: Fully Integrated Content Analysis in International Relations. International Relations, Vol. 31, No. 4, 447-465. (2017)
7. Timmermans, S., Iddo, T.: Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. Sociological Theory, Vol. 30, No. 3, 167-186. (2012)
8. Marcus, S., Moy, M., Coffman, T.: Social Network Analysis. In: Cook, D. J., Holder, L. B. (eds.): Mining Graph Data. John Wisley & Sons, Hoboken. (2007)
9. UCINET documentation (2017). [Online]. Available: https://sites.google.com/site/ucinetsoftware/document (current February 2021)

10. Scott, J.: Social Network Analysis: Developments, Advances, and Prospects. Social Network Analysis and Mining, Vol. 1, No. 1, 21-26. (2011)
11. Qi, X., Davison, B. D.: Web Page Classification: Features and Algorithms. ACM Computing Surveys, Vol. 41, No. 2, 12:1–12:31. (2009)
12. Patil, A., Pawar, B.: Automated Classification of Web Sites using Naive Bayessian Algorithm. In Proceedings of the International Multi-Conference of Engineers and Computer Scientists, Vol. 1, Hong Kong, 466-467. (2012)
13. Kotenko, I., Chechulin, A., Shorov, A., Komashinsky, D.: Analysis and Evaluation of Web Pages Classification Techniques for Inappropriate Content Blocking. In: Perner, P. (ed.): Proceedings of the 14th Industrial Conference on Data Mining, Lecture Notes in Artificial Intelligence, Vol. 8557, 39-54. (2014)
14. Shibu, S., Vishwakarma, A., Bhargava, N.: A Combination Approach for Web Page Classification using Page Rank and Feature Selection Technique. International Journal of Computer Theory and Engineering, Vol. 2, No. 6, 897-900. (2010)
15. Kan, M.-Y., Thi, H. O. N.: Fast Web Page Classification using URL Features. In Proceedings of the 14th ACM International Conference on Information and Knowledge Management, Bremen, Germany, 325-326. (2005)
16. Baykan, E., Henzinger, M., Marian, L., Beber, I.: Purely URL-Based Topic Classification. In Proceedings of the 18th International Conference on World Wide Web, Madrid, Spain, 1109-1110. (2009)
17. Dumais, S., Chen, H.: Hierarchical Classification of Web Content. In Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Athens, Greece, 256-263. (2000)
18. Calado, P., Cristo, M., Moura, E., Ziviani, N., Ribeiro-Neto, B., Goncalves, M. A.: Combining Link-Based and Content-Based Methods for Web Document Classification. In Proceedings of the 12th International Conference on Information and Knowledge Management, New York, New Orleans, LA, USA, 394-401. (2003)
19. Belmouhcine, A., Idrissi, A., Benkhalifa, M.: Web Classification Approach using Reduced Vector Representation Model Based on HTML Tags. Journal of Theoretical and Applied Information Technology, Vol. 55, No. 1, 137-148. (2013)
20. Kotenko, I., Chechulin, A., Komashinsky, D.: Categorisation of Web Pages for Protection against Inappropriate Content in the Internet. International Journal of Internet Protocol Technology, Vol. 10, No. 1, 61-71. (2017)
21. Kriesel, D.: A Brief Introduction to Neural Networks. Cambridge University Press, Grate Britain. (2010)
22. Mehlig, B.: Artificial Neural Networks. University of Gothenburg, Sweden. (2019)
23. Rojas, R.: Neural Networks. Springer-Verlag, Germany. (1996)
24. Parashchuk, I.: System Formation Algorithm of Communication Network Quality Factors using Artificial Neural Networks. In Proceedings of the 1st IEEE International Conference on Circuits and System for Communications, Saint-Petersburg, Russia, 263-266. (2002)
25. Kosko, B.: Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence. Prentice-Hall, Englewood Cliffs, NJ, USA. (1992)
26. Kotenko, I., Saenko, I., Ageev, S., Kopchak, Y.: Abnormal Traffic Detection in Networks of the Internet of Things Based on Fuzzy Logical Inference. In Proceedings of the XVIII International Conference on Soft Computing and Measurements, Saint-Petersburg, Russia, 5-8. (2015)
27. Kotenko, I., Parashchuk, I., Omar, T.: Neuro-Fuzzy Models in Tasks of Intelligent Data Processing for Detection and Counteraction of Inappropriate, Dubious and Harmful Information. In Proceedings of the 2nd International Scientific-Practical Conference Fuzzy Technologies in the Industry, Ulyanovsk, Russia, 116-125. (2018)

Igor Kotenko obtained the Ph.D. degree in 1990 and the National degree of Doctor of Engineering Science in 1999. He is Professor of computer science and Head of the

Laboratory of Computer Security Problems of St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). He was a project leader in the research projects from the US Air Force research department, via its EOARD (European Office of Aerospace Research and Development) branch, EU FP7 and FP6 Projects, HP, Intel, F-Secure, etc. His research results were tested and implemented in more than fifty Russian research and development projects. His main research interests are innovative methods for network intrusion detection, simulation of network attacks, vulnerability assessment, verification and validation of security policy, etc. He has chaired several International conferences and workshops, and serves as editor on multiple editorial boards.

Igor Saenko obtained the Ph.D. degree in 1992 and the National degree of Doctor of Engineering Science in 2002. He is Professor of computer science and Leading Researcher of the Laboratory of Computer Security Problems of St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). His main research interests are security policy management, access control, management of virtual computer networks, knowledge modeling soft and evolutionary computation, information and telecommunication systems.

Elena Doynikova obtained her PhD in St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS) in 2017. In 2015 she was awarded the medal of the Russian Academy of Science in area of computer science, computer engineering and automation. Currently she is a senior researcher of computer security problems laboratory, SPC RAS. Research interests: information systems security, risk analysis and security decision support methods, security metrics, data mining. She participated in many projects devoted to information systems security research.

Igor Parashchuk, Doctor of Engineering Sciences, Professor; Leading Researcher of the Laboratory of Computer Security Problems in St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), areas of scientific interest: computer network security, automated information systems, data storage and processing.

Received: March 14, 2021; Accepted: August 31, 2021.