Procedia Computer Science
Volume 88, 2016, Pages 288–293
7th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2016

Information Security and Expert's Knowledge Autoformalization

Anatoly Malyuk[1,2] and Natalia Miloslavskaya[1]
[1]National Research Nuclear University MEPhI (Moscow Engineering Physics Institute)
[2]Financial University under the Government of the Russian Federation, Moscow, Russia
{AAMalyuk, NGMiloslavskaya}@mephi.ru

Abstract
To implement the proposed Information Security (IS) Maintenance Concept, the IS experts' knowledge autoformalization algorithm was created as the problems of IS assessment and protection level prediction are based mainly on the experts' informal professional knowledge.

Keywords: information security, unified information security maintenance concept, experts' professional knowledge autoformalization

1 Introduction

Intensive information and communication technologies (ICT) development has led to serious qualitative changes in all spheres of public life. Humankind is actually going through the formation of a new information society characterized by its great reliance on ICT. Information and ICT become the main strategic national resources. The ICT phenomenon sharply increasing the impact of 21st century's society was marked in the Okinawa Charter on Global Information Society adopted by the Group of Eight (G8) on July 22, 2000. At the same time the increasing role of ICT leads to an increase in the information security (IS) threats and brings IS issues to the forefront of the any system security that requires the development of science-based approaches to solving them. These IS threats relate to violation of the established modes of ICT systems usage, infringement of the constitutional rights and freedoms of citizens, malware spread, as well as the usage of modern ICT capabilities for the implementation of hostile, terrorist and other criminal acts. Managing IS correctly requires a comprehensive vision of the issues emerging and informed decision making. Hence, the IS maintenance (ISM) issues and, above all, reliable information protection (preventing its distortion, unauthorized modification, malicious collection, etc.) are now of special urgency.

We refers to *IS of a system (system's IS)* as its quality to be characterized, on the one hand, by its ability to resist the destabilizing effects of external and internal threats, and, on the other hand, by the level of threats posed by its operation to the elements of the system and its external environment. And *ISM* is a complicated process divided into many sub-processes of maintaining the secure (protected)

Selection and peer-review under responsibility of the Scientific Programme Committee of BICA 2016

doi:10.1016/j.procs.2016.07.437

state of information, characterized by its confidentiality, integrity, availability, etc. via information protection tools/systems (Malyuk, 2014), (Malyuk, 2015).

Thus the paper is organized as follows. The proposed unified ISM concept is described in section 2. The general ISM processes' model is presented in section 3. IS experts' knowledge autoformalization algorithm is introduced in section 4. The future research area concludes the paper.

2 Unified ISM Concept

The unified ISM concept's structure, worked out by selective integration of the best ISM practices (analyzed in detail in (Malyuk, 2014), (Malyuk, 2015)) and based on the general methodological approaches of the classical systems theory and modelling methods and our extension of these approaches and methods to the specific field of IS, is shown in Fig. 1.

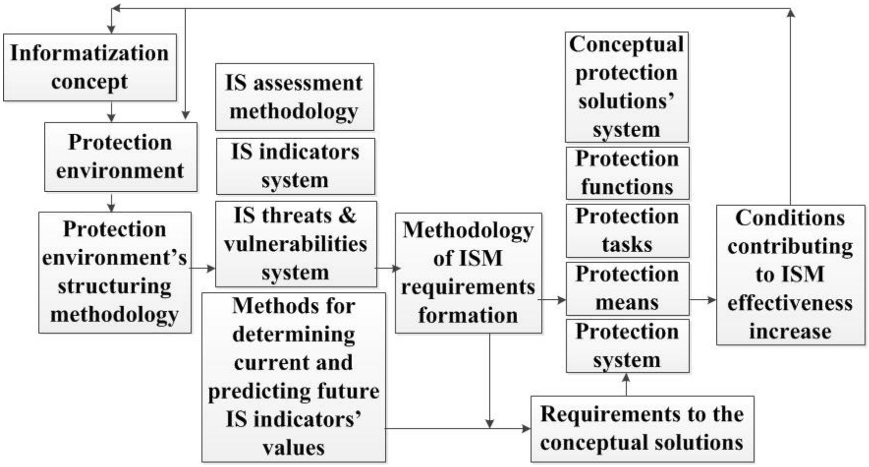

Informatization concept
Protection environment
Protection environment's structuring methodology
IS assessment methodology
IS indicators system
IS threats & vulnerabilities system
Methods for determining current and predicting future IS indicators' values
Methodology of ISM requirements formation
Conceptual protection solutions' system
Protection functions
Protection tasks
Protection means
Protection system
Requirements to the conceptual solutions
Conditions contributing to ISM effectiveness increase

Figure 1: Unified ISM concept's structure

First of all the concepts defining the protection content are developed. They are formed from the concepts of protected automated systems' construction and usage (an informatization concept) and its operation conditions (a protection environment). Further the protection content description is carried out, implying a strictly formal, or, if it is impossible, structured (in the form of a set of interacting elements) representation of the corresponding automated systems' architecture and operation. The IS assessment methodology comprises methods, models and tools to determine the current and predict future values of each of the system's IS indicators under the influence of each of the potential threats and vulnerabilities presence and any their aggregation. The ISM requirements formation methodology determines the approaches, tools and methods of practical organization of information protection. The conceptual protection solutions' system creates objective prerequisites for the formation of various tools and means necessary and sufficient to effectively address the set of the relevant ISM tasks on a regular basis and in accordance with the requirements to their solution which, in turn, are determined by the objectives of the operation of the respective system. The requirements to the conceptual solutions enable to justify such requirements to each of the conceptual solutions that achieve the goals of their adoption in the most rational way. The conditions contributing to ISM effectiveness are required for the formation and study of the list and content of those conditions (including the protection content), compliance with which will significantly increase the protection level in the expenditure allocated for this purpose funds or provide the required protection level while spending the smallest possible amount of funds.

The unified ISM concept is applicable to all protection levels called the protection objects: a separate computer/mobile device, an enterprise or its part, region or state. It creates all the necessary objective conditions for a transition to a new stage in solving the problems – a stage of ISM processes intensification. To describe the given object and to work out its ISM policy in accordance with the unified ISM concept means to answer the following twelve questions: What types of the IS threats are possible in that case? What is the nature of the IS threats origin? What type of the channels of unauthorized information gathering can be used? What are the IS threats sources? What are the reasons leading to the information confidentiality violation? What are the potential malicious acts? What are the requirements to ISM? How the protected object can be classified? What factors will affect the required information protection level? What are the applicable information protection systems (IPSs)? How IPS's architecture can be presented? What are the recommendations for improving the protection level?

3 General Model of ISM Processes

ISM problem analysis can be regarded as a formal system presented by the quartet

$$\mathrm{Z} =< \mathrm{R}_0, \mathrm{R}_\mathrm{p}, \mathrm{K}, \mathrm{U} >,$$

where R_0 is an initial state of the protected system defined by available data; R_p is a predictable state of the system corresponding to its potential capabilities to counter IS threats; K is knowledge about the system (elementary and composite models and their relationship, constraints on individual parameters, etc.); U is a system utility function, balancing the operation efficiency and its provision costs.

A principal element of the ISM methodological basis is to construct adequate models for the systems and processes under study. The proposed generalized ISM processes' model with the block diagram shown in Fig. 2 may become a basis for the ISM problem solution.

The model operates with the following parameters: $\{K\}$ is a set of IS indicators; $\{P^{(s)}\}$ is a set of environmental parameters affecting the automated system's functioning; $\{R^{(s)}\}$ is a set of system resources involved in the processing of protected information; $\{P^{(m)}\}$ is a set of internal system parameters that can be controlled directly in the processing of protected information; $\{P^{(i)}\}$ is a set of internal system parameters, which are beyond the direct control, but are affected (for example, in the process of the system components' reorganization or improvement); $\{S^{(m)}\}$ and $\{R^{(m)}\}$ are means and resources of the current management; $\{S^{(i)}\}$ and $\{R^{(i)}\}$ are means and resources of impact; $\{R^{(g)}\}$ is a set of general management resources.

To solve the problems of the analysis with this model (i.e. to determine IS indicators' values) the following generalized expression can be used:

$$\{K\} = F\left[\{P^{(m)}\}, \{\mathrm{P}^{(i)}\}, \{R^{(s)}\}, \{P^{(s)}\}\right].$$

The synthesis problem can be formulated in general form as to

1) find such $\{R^{(m)}\}$ and $\{R^{(i)}\}$ ($\{R^{(m)}\}$ + $\{R^{(i)}\} \le \{\bar{R}^{(g)}\}$, ($\{\bar{R}^{(g)}\}$ are the given resources) to satisfy the condition $\{K\}\rightarrow$max for given $\{R^{(s)}\}$ and $\{P^{(s)}\}$;

2) choose such $\{R^{(m)}\}$ and $\{R^{(i)}\}$ that the condition $\{K\} \ge \{\bar{K}\}$ ($\{\bar{K}\}$ is a given protection level) is satisfied for $\{R^{(g)}\}$ = $\{R^{(m)}\}$ + $\{R^{(i)}\}\rightarrow$min for given $\{R^{(s)}\}$ and $\{P^{(s)}\}$.

Thus the management problems can be reduced to optimize the distribution of $\{R^{(m)}\}$, $\{S^{(m)}\}$, $\{R^{(i)}\}$, $\{S^{(i)}\}$.

A modification of the generalized model can be easily expressed as: a model of the system's functioning in the absence of ISM (the model allows to determine the IS indicators' values, i.e. to solve the analysis tasks); a model of the current ISM based on optimizing the IPSs directly built-in the system; a model for managing the resources dedicated to ISM, which in addition to the previous tasks allows to optimize the formation of the current ISM tools and means; a model for managing the impacts on the parameters that do not allow current ISM, but can be affected; a model for managing the resources dedicated to the system's development; a full ISM model that allows to optimize the use of all resources allocated to ISM in addition to all the features discussed above.

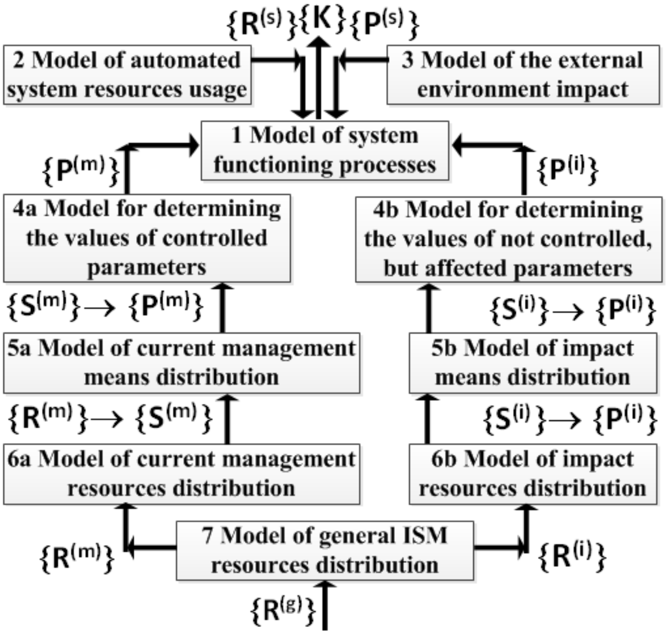

{R(s)}{K}{P(s)}
2 Model of automated system resources usage
3 Model of the external environment impact
1 Model of system functioning processes
{P(m)}
{P(i)}
4a Model for determining the values of controlled parameters
4b Model for determining the values of not controlled, but affected parameters
{S(m)}→ {P(m)}
{S(i)}→ {P(i)}
5a Model of current management means distribution
5b Model of impact means distribution
{R(m)}→ {S(m)}
{S(i)}→ {P(i)}
6a Model of current management resources distribution
6b Model of impact resources distribution
{R(m)}
7 Model of general ISM resources distribution
{R(i)}
{R(g)}

Figure 2: Generalized model of ISM processes

Our studies have shown that taking into account the required protection level and the degree of freedom in activities while organizing ISM three basic strategies of information protection may be distinguished: defensive, offensive and proactive. Each of these strategies can be effectively implemented within the unified ISM concept.

The practical usage effectiveness of the model described depends significantly on the representativeness and adequacy of statistical data, allowing to determine the functional dependencies establishing relationships between IS indicators, IPSs' parameters and volumes of resources invested in the implementation of ISM processes. In this regard, the model under consideration can be used only along with the informal methods of analysis and forecasting, in particular, using the IS expert's knowledge autoformalization algorithm.

4 IS Experts' Professional Knowledge Autoformalization

The whole ISM strategy of finding effective protection solution as well as most of the result interpretation stages should be based mainly on the informal IS experts' professional knowledge and intuitive methods. In these circumstances, IS expert's knowledge autoformalization can be the only way to create models of the investigated protection content on the basis of analytical activity algorithms. Thus the problem of IS experts' professional knowledge formalization arises. The autoformalization results in new information that the IS expert receives during the experiment with the models and the models themselves that reflect his deeper understanding of the structure of a particular ISM process under investigation and its inherent qualitative and quantitative dependencies.

Based on the experts' experience the following sequence and interrelation of the knowledge autoformalization algorithm stages is proposed in Fig. 3.

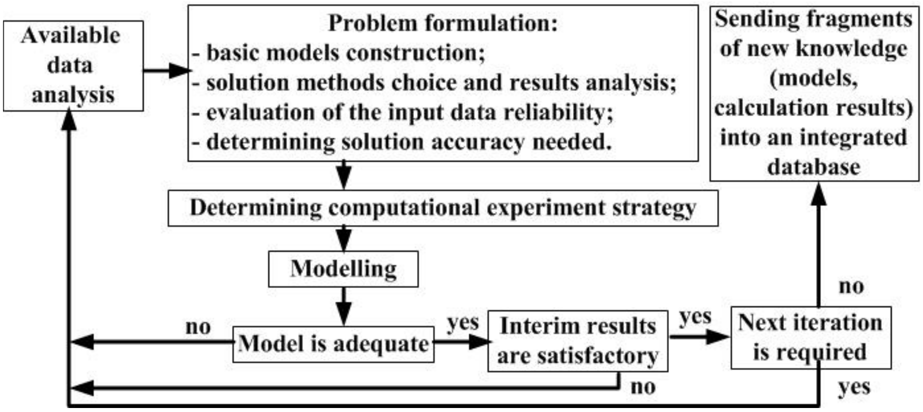

Available data analysis
Problem formulation:
- basic models construction;
- solution methods choice and results analysis;
- evaluation of the input data reliability;
- determining solution accuracy needed.
Sending fragments of new knowledge (models, calculation results) into an integrated database
Determining computational experiment strategy
Modelling
no
Model is adequate
yes
Interim results are satisfactory
yes
Next iteration is required
no
no
yes

Figure 3: Sequence of knowledge autoformalization stages

Analyzing all available data the IS expert formulates a problem to be solved. For that purpose he constructs the basic models, made a choice of methods applicable for solution, analyses the results obtained, evaluates the input data reliability and determines accuracy needed for the problem solution. All that helps him to determine a computational experiment strategy and after that to fulfil modelling. If the created models are adequate and the interim results satisfy the expert, the fragments of new knowledge can be sent to the integrated knowledge database (DB). If not, the expert starts from the initial data revision and repeats analysis. The described process is iterative till the satisfactory results will be achieved.

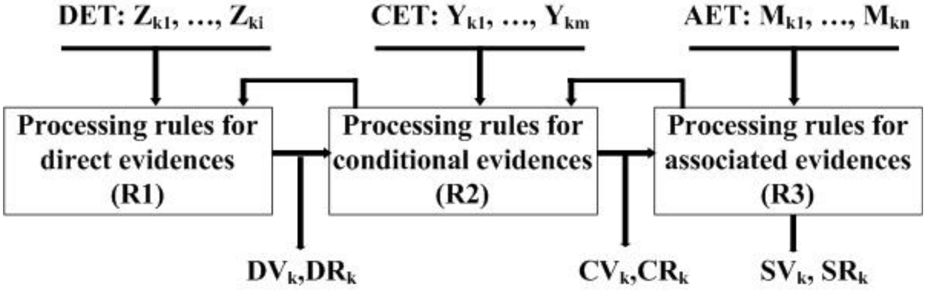
It is clear that in order to arrange the organizational support of the complex of information protection activities to implement a collection, accumulation and systematization of the input data is of fundamental importance in these circumstances. Let us define the accuracy as a level of reasonable assurance of the truth of a statement that satisfies some consistency rules and in accordance with those rules can be formally expressed by a number (Zelner, 1980). Using the idea of the Bayesian approach, it is possible to set the question about the reliability of expert's integrated DB fragments (DB, knowledge base and models base) considering any fragment as a hypothesis, and the fragments with which it is associated as an evidence for the fragmentation hypothesis. Then each new fragment entering the integrated database (NPK – a new piece of knowledge) can be represented as a pair <V, R>, where V is a fragment value and R is its reliability. NPK included in the integrated DB interacts with the fragments and hypotheses already contained therein, changing both their values and reliabilities. This reaction is quite complex and causes a modification of values and reliabilities of all the old fragments, anyway connected with the newly entered.

The notion of fragment's system value and system reliability defined taking into account all the evidences contained in the integrated DB should be introduced to describe the modifications. Modification of fragment's value and reliability when the evidence composition is changed can be accomplished using an algorithm with a flowchart shown in Fig. 4.

Different methods can be applied for evidence processing in the modification algorithm. An approach based on the use of filtering methods seems to be the most convenient. Wherein the general problem statement is a system of two equations, the first of which describes the structure and dynamics of the ISM process, and the second determines the generation mechanism for data available for the expert, then

$$x(k+1) = F[x(k), w(k), d(k)]\,,\; y(k) = H[x(k), v(k), d(k)]\,,$$

where *x (k)* is a state vector of the process under investigation; *w(k)* is a random vector of the investigated process's noises related to modelling methods' errors; *y(k)* is an observation vector; *v(k)* is a random vector of observation noises related to the errors of information reception channel (*w* and *v* are uncorrelated); *d(k)* is a variability vector that characterizes the current state and structure of the process and channel of information reception (in this case a failure is seen as a change in parameters or structure).


DET: $Z_{k1}, \ldots, Z_{ki}$
CET: $Y_{k1}, \ldots, Y_{km}$
AET: $M_{k1}, \ldots, M_{kn}$
Processing rules for direct evidences (R1)
Processing rules for conditional evidences (R2)
Processing rules for associated evidences (R3)
DV_k, DR_k
CV_k, CR_k
SV_k, SR_k

DET is a tuple of direct evidences for a given fragment; CET is a tuple of conditional evidences; AET is a tuple of associated evidences; DV and DR are fragment's value and reliability considering all direct evidences; CV and CR are fragment's value and reliability considering all conditional evidences; SV and SR are fragment's system values and reliabilities

Figure 2. Flowchart of an algorithm of calculating system value and system reliability for a fragment from the integrated DB

5 Conclusion

Main ideas of the proposed unified ISM concept in a concentrated form are the following: a structured description of the protection environment; a comprehensive quantitative analysis of an information protection degree for a particular protection object; a scientifically based definition of the required protection level at each specific object under specific conditions of its functioning; development of optimal IPSs on a single unified methodology. The further research is expected in the concept application and the development of IS expert's information environment model and behavioral intruders' models that improve the effectiveness of protection environment description.

6 Acknowledgement

This work was supported by the Competitiveness Growth Program of the Federal Autonomous Educational Institution of Higher Education National Research Nuclear University MEPhI (Moscow Engineering Physics Institute).

References

Malyuk, A., Miloslavskaya, N. (2014). *Information Security Theory Development*. Proceedings of the 7th Int. Conference on Security of Information and Networks SIN2014, Glasgow (UK). Pp. 52-55.

Malyuk, A., Miloslavskaya, N. (2015*). Information Security Theory for the Future Internet*. Proceedings of the 3rd Int. Conference on Future Internet of Things and Cloud FiCloud 2015. Rome (Italy). Pp. 150-157.

Zelner, A. (1980). *Bayesian Methods in Econometrics*. Moskva, Statistics, 1980 (in Russian).