A model for detecting vulnerabilities in UMV interfaces, based on probabilistic automata

A V Skatkov, A A Bryukhovetskiy and D V Moiseev*
Sevastopol State University», 299053, Sevastopol, Russia

*dmitriymoiseev@mail.ru

Abstract. The main features associated with the development and research of methods of adaptive intelligent technology for monitoring the state of objects of computer systems are considered. The proposed approach is focused on detecting moments of change in the state of controlled UMV resources, which are: communication channel, processor, memory. The aim of this work is to develop an adaptive model using a Bayesian classifier for estimating the state of UMV resources. The model is based on a probabilistic automaton with parametric self-tuning.

1. Introduction
This paper is devoted to the application of a model of probabilistic automata for decision support in assessing the information state of transport infrastructure objects in a smart city. Such objects, for example, include intelligent control systems for unmanned aerial and ground vehicles, systems that provide inter-machine interaction using Internet of things technology, and others. The heterogeneity of applications and wireless communications in the smart city infrastructure significantly complicates the security of objects [1].

The approach proposed in the article is focused on detecting moments of change in the state of controlled UMV objects, which are resources: communication channel, processor, memory. However the speed and reliability of the situation assessment can be crucial. In conditions of a lack of a priori information most of the problems of data analysis are related to research of stochastic systems [2-5]. One of the most effective tools for modeling complex stochastic systems is the methodology of probabilistic-automatic modeling [6,7]. The effectiveness of this methodology is determined by the following two features:
1) availability of tools that provide an adequate description of complex stochastic systems and their functioning processes;
2) ability to build unified models for a wide class of systems.

Various probabilistic automata models are currently known. The reason for the variety of automatic models is explained by the breadth of their application. Probabilistic automata are used in such areas as, for example: logical management, mathematical linguistics, theory of formal languages, modeling of human behavior, when describing models of enterprise information security, etc.

The criterion for the applicability of the automatic approach is best expressed in the concept of «complex behavior» [8]. We can say that an object has complex behavior if it can perform one of several output reactions as a response to some input action. It is significant that the reaction may depend not only on the input effect, but also on the background.

2. Problem statement
The aim of this work is to develop an adaptive model using a Bayesian classifier for estimating the state of UMV resources. The model is based on a probabilistic automaton with parametric self-tuning. The main task is to form transition probability matrices in the process of monitoring the state of UMV resources and evaluate the values of matrix elements formed for normal resource behavior in the time
interval \( T_N \) and the values obtained as a result of external influence for a certain time interval \( T_{2N} \). The use of this model will increase the level of reliability of classification of information situations that occur when monitoring the characteristics of UMV resources.

With a fairly General statement of the problem we are talking about the need to control the results of observations on the state of UMV resources. Denote \( R = \{ R_1, \ldots, R_j, \ldots, R_r \} \) – set of controlled UMV resources. Define the controlled characteristics of the resources, according to which we will assess their condition (the values of the characteristics are normalized, defined in the range \([0;1]\)):

- \( D_j \) – loading of \( j \)-th resource,
- \( V_j \) – rate of change \( D_j \), where \( V = (D(t_i) - D(t_{i-1}))/\Delta t \).

The specified characteristics are vector quantities with components of set elements \( R_j \).

We assume that the state of \( R_j \) at a given time \( t \) depends on the values of the characteristics \( D_j, V_j \). Let these states at time \( t \) be denoted by \( S_j = \{ S'_0, S'_1, S'_2 \} \) – the set of possible states of the object \( R_j \).

State values are defined and normalized in the range \([0;1]\).

Let the states be defined at intervals \([I_k; I_{k+1}]\), where \( I_k \) – threshold for setting the resource state detection area, \( k = 0, 1 \). Without loss of generality for the example (Fig. 1), we will assume that we have three intervals on which possible states \( S'_j \) are determined:

- \( S'_0 \in [0;I_0] \), \( S'_1 \in (I_0;I_1] \), \( S'_2 \in (I_1;1] \),

Assume that the area with the number "0" indicates a normal state, and the area with the number "2" – a critical state, the area "1" - correspond to a precritical state.

3. Methods and results.

We will record changes in the \( S'_j \) state of the resource in the time interval \( T_N \) at moments of time \( \{ t_0, t_1, \ldots, t_j, \ldots, t_n \} \).

To evaluate the dynamic probabilistic states of UMV resources it is proposed to use an automaton model. For this purpose we define a probabilistic automaton as a system

\[
\sum = (S, X, Y, \Phi, \Psi, S_i) ,
\]

where

- \( S_i \in S \) – initial state,
- \( S \) – set of values of the state vector,
- \( X \) – set of values of the input vector,
- \( Y \) – set of values of the output vector,
- \( \Phi \) – transition function,
- \( \Psi \) – output function,

The automaton defined by the scheme (see figure 1) functions in a discrete automaton time, the moments of which are the clock cycles \( t_0, t_1, t_2 \ldots \), each of which corresponds to the values of the input signal – \( X \), output signal – \( Y \) and internal states – \( S \).

\[
S'_0 \quad S'_1 \quad S'_2
\]

\[
0 \quad I_0 \quad I_1 \quad 1
\]

**Figure 1.** Areas of difference in resource states \( S'_j \).

To evaluate the dynamic states of resources it is proposed to use the automaton probabilistic Moore model. A special feature of the probabilistic dynamic model is that the elements of the transition and output matrices represent the values of the corresponding transition probabilities between states and outputs. Denote the specified probabilities – \( P \) as:

- \( P(s_i(t+1)) = \Phi(s_i(t),x(t)) \),
- \( P(y_j(t+1)) = \Psi(s_i(t+1)) \),

where \( j=1,m \), \( m \) – the number of the automaton states, \( t = 0,1,2 \ldots \).

At the initial moment of time \( t_0 \) the automaton is in the \( S_i \) state. At every moment of time \( t_i \) the automaton is in one of the internal states, is able to receive an input signal, form a new internal state,
and issue the corresponding output signal. The transition table sets the transition probabilities to the $s_j(t+1)$ state at time $t+1$ depending on the $s_j(t)$ state at time $t$ if the signal $x_i(t)$ is received. Each row of the matrix contains the probabilities of transition from state $i$ to any possible state $j$, which form a complete group, and therefore the sum of the probabilities of these events is equal to one. For a probabilistic automaton, we need as many transition tables as there are input signals $x_i(t)$. Below is one transition table (table 1) for the input signal $x_i(t)$. Denote $P_{ijk}$ – probability of transition of the automaton to the state $S_k(t+1)$ provided that the automaton was in the state $S_j(t)$ and the input signal $x_i(t)$ was received.

**Table 1. Table of the automaton transitions.**

| $x_i(t)$ | $S_0(t)$ | $S_1(t)$ | $S_j(t)$ | $S_m(t)$ |
|----------|----------|----------|----------|----------|
|          | $P_{00}$ | $P_{01}$ | $P_{0j}$ | $P_{0m}$ |
|          | $P_{10}$ | $P_{11}$ | $P_{1j}$ | $P_{1m}$ |
|          | $...$    | $...$    | $...$    | $...$    |
| $x_i(t)$ | $S_j(t)$ | $P_{j0}$ | $P_{ji}$ | $P_{jm}$ |
|          | $...$    | $...$    | $...$    | $...$    |
| $x_i(t)$ | $S_m(t)$ | $P_{m0}$ | $P_{mi}$ | $P_{mj}$ |
|          | $...$    | $...$    | $...$    | $...$    |

The output table of the Moore automaton (table 2) is simplified, since the output state $Y(t+1)$ depends only on the internal state $S(t+1)$ and does not depend on the input signal $x(t)$. Denote $P_{ij}$ – probability of an automaton output $Y_j(t+1)$ provided that the automaton was in the state $S_i(t+1)$.

**Table 2. Table outputs of a Moore automaton**

| $S_0(t+1)$ | $S_1(t+1)$ | $...$ | $S_j(t+1)$ | $...$ | $S_m(t+1)$ |
|------------|------------|-------|------------|-------|------------|
| $Y_0(t+1)$ | $P_{00}$   | $P_{01}$ | $...$ | $P_{0j}$ | $...$ | $P_{0m}$   |
| $Y_1(t+1)$ | $P_{10}$   | $P_{11}$ | $...$ | $P_{1j}$ | $...$ | $P_{1m}$   |
| $...$      | $...$      | $...$  | $...$     | $...$ | $...$      |
| $Y_j(t+1)$ | $P_{j0}$   | $P_{ji}$ | $...$ | $P_{jj}$ | $...$ | $P_{jm}$   |
| $...$      | $...$      | $...$  | $...$     | $...$ | $...$      |
| $Y_m(t+1)$ | $P_{m0}$   | $P_{mi}$ | $...$ | $P_{mj}$ | $...$ | $P_{mm}$   |

At the initial moment of time $t_1$ automaton is in the state $S_i$. The output signal is not generated in the initial state. Then in the first clock cycle when the $x_i$ signal is received the automaton with the probability $P_{ij}$ goes to the $S_j$ state and the output signal $Y_j$ is generated. The initial distribution of states is shown in the table 3.

**Table 3. Initial distribution of the automaton states**

| $\Psi(s_j(t+1))$ | $Y_0(t+1)$ | $Y_0(t+1)$ | $...$ | $Y_j(t+1)$ | $...$ | $Y_0(t+1)$ |
|-------------------|------------|------------|-------|------------|-------|------------|
| $\Phi(s_j(t),x(t))$ | $S_0(t+1)$ | $S_0(t+1)$ | $...$ | $S_j(t+1)$ | $...$ | $S_m(t+1)$ |
| $S_j(t)$          | $P_{j0}$   | $P_{j0}$   | $...$ | $P_{jj}$   | $...$ | $P_{jm}$   |

The initial probability distribution of states is formed on the basis of a priori information that can be obtained during the normal functioning of the UMV in the time interval $T_N$, when there are no external influences. Accordingly, for this mode of operation, the probability distributions $P_{ijk}$ of transitions between states $S_j \rightarrow S_k$ at the input signal $x_i$ can be obtained and the transition, output, and initial probability distribution matrices are formed.

In the operating mode, the UMV is subject to external influences that lead to changes in the values of a priori probabilities. In order to compensate for the influence of external factors based on a posteriori information, it is proposed to use a probabilistic adaptive classifier. The main task is to get the values of
conditional distributions that determine the probability that an observation belongs to each of the possible classes. One of the known methods is based on the application of Bayes’ theorem. In the context of solving the problem of detecting changes in the state of resources, the formula gets the following interpretation:

\[
P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{i=1}^{n}P(H_i)P(A|H_i)}.
\]

If \( H_i \rightarrow S_i, A \rightarrow (S_i \rightarrow S_j) \):

- \( P(S_i) \rightarrow \) probability of occurrence of the condition \( S_i \);
- \( P(S_i | (S_i \rightarrow S_j)) \rightarrow \) conditional probability of observing the \( S_i \) state when transitions appear \( S_i \rightarrow S_j \).

The algorithm for constructing transition probability matrices and the initial probability distribution of states contains the following sequence of actions:

- the state vector is played \( S_i \) and the sequence is built \( S_0, S_1, \ldots, S_j, \ldots, S_l \), where \( i, j, l \in \{0,1,2\} \);
- the number of \( N_{ij} \) obtained pairs of the form \( (S_i \rightarrow S_j) \) is calculated;
- the number of \( N_i \) obtained pairs of the form \( (S_i \rightarrow S_j) \), where \( j=0,1,2 \), i.e. transitions from \( S_i \) to any state of \( S_j \) are counted;
- probabilities of state transitions are found \( P_{ij} = N_{ij}/N_i \);
- transition/exit matrices are constructed;
- the initial probability distribution of States is determined – \( P_{ii} = N_i/N \), where \( N \rightarrow \) the total number of states.

In the working mode of operation, the adaptive self—tuning of the automaton—the matrix of transient probabilities is used. To do this, the forecast of the state of the automaton made at step \( t_i \) is evaluated in each clock cycle \( t_i \). If the forecast is correct, the corresponding element \( P_{ijk} \) of the transition matrix accumulates the number of \( NP_{ijk} \) such "rewards". Otherwise, if the forecast is incorrect, the number of \( NP_{ijk} \) penalties is accumulated. Whenever the value \( NP_{ijk} > \eta \) (the value of \( \eta \) is set by the expert), the transition matrix elements are corrected by the value \( NP_{ijk} \Delta P_{ijk} \), where \( \Delta P_{ijk} \) – part of change the value of a matrix element in a single clock cycle.

4. Conclusion

The proposed approach is based on the use of a probabilistic automaton and a Bayesian classifier. The method is based on the use of adaptive parametric self-tuning of elements of transition probability matrices. The development of adaptive intelligent technology based on the assessment of changes in the state of UMV resources, based on an adaptive automatic probabilistic model, will lead to an increase in the reliability and efficiency of decision support processes when solving security problems of critical information infrastructure objects "Smart city".

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (projects № 19-29-06015/19, 19-29-06023/19).

References

[1] Strite S and Morkoc H 1992 J. Vac. Sci. Technol. B 10 1237
[2] Nakamura S, Senoh M, Nagahama S, Iwase N, Yamada T, Matsushita T, Kiyoku H and Sugimoto Y 1996 Japan. J. Appl. Phys. 35 L74
[3] Zegzhda P.D. Systematization of cyberphysical systems and assessment of their security / P.D. Zegzhda, M.A. Poltavtseva, D.S. Lavrova // Problems of information security. Computer system. 2017. № 2. Pp. 127–138.
[4] Shiryaev A.N. Probabilistic and statistical methods in the theory of decision-making. 2-ed., new. M.: MCCME, 2014. 144 p.
[5] Information technologies for critical infrastructures: monograph / edited by A.V. Skatkov. Sevastopol: SevNTU, 2012. 306 p.
[4] *Skatkov A.Y., Brykhovetskiy A.A., Moiseev D.V.* Intelligent monitoring system for solving large-scale scientific problems in cloud computing environments // Information and control systems. SPb.: «GUAP>, 2017. № 2 (87). Pp. 19–25.

[5] Skatkov A., Brykhovetskiy A., Moiseev D. Detecting changes simulation of the technological objects’ information states // MATEC Web of Conferences. Vol. 224, 2018. 02072 (ICMTMTE 2018). https://doi.org/10.1051/matecconf/201822402072

[6] *Varshavskiy V.I., Pospelov D.A.* The orchestra plays without a conductor: Reflections on the evolution and management of certain technical systems. – M.: Science. Main editorial office of physical and mathematical literature, 1984.

[7] *Pospelov D.A., Pushkin V.N.* Thinking and automata. M.: Sov. Radio, 1972. 244 p.

[8] *Shalyto A. A.* Automata-based programming paradigm // Scientific and technical bulletin of Saint Petersburg state University ITMO. Automata-based programming. 2008. 53, Pp 3–23.