Implementation of a combined algorithm designed for increasing the reliability of information systems

S S Bychkov\textsuperscript{1}, A M Popov\textsuperscript{2} and A K Gorbacheva\textsuperscript{2}

\textsuperscript{1}Krasnoyarsk State Government Agency “Information Technology Center of Krasnoyarsk Krai”, 2 Robespierre St., Krasnoyarsk, 660021, Russia
\textsuperscript{2}Reshetnev Siberian State University of Science and Technology, 31 Krasnoyarskii rabochii prospekt, Krasnoyarsk, 660037, Russia

E-mail: ampopov1959@gmail.com

Abstract. This paper considers the results of experimental studies of the previously presented algorithm designed for increasing the reliability level of information systems. The data on the organization and conduct of the studies is provided to clarify their meaning. Within the framework of the study conducted, the obtained experimental data of simulation modeling and the functioning data of existing information systems are compared. The hypothesis of homogeneity of the structure of information systems has been formed, thereby reconfiguring the presented algorithm into a model for the analysis of information systems. The results presented can be used in further research activities. The obtained data on the opportunity to predict the functioning of information systems can be used for economic planning.

1. Introduction
Disruption in the operation of the company in modern economic conditions leads to significant economic losses and reputational issues. In this regard, increasing attention is being paid to improving the reliability of production elements that have a direct impact on the organization’s business process, while the level of process automation is steadily growing. Thus, increasingly greater attention is being paid to ensuring the efficiency of automation systems (AS), as well as control information systems (IS).

As a result, the efficiency (capacity to carry out a part of the production process correctly) of IS and AS is a parameter that affects the functioning of the company directly and must be given due particular consideration and analysis (no costs to restore IS/AS and correction of internal errors, fulfillment of obligations by hiring third parties to enforce contractual obligations) [6].

The authors have previously presented the paper describing the notion of the application options for the combined algorithm designed for increasing the level reliability of information systems [1,8] (Algorithm). While designing the Algorithm, the following mathematical methods have been selected:

− methods of calculating the reliability of the information system index, in particular, the dynamic method (dynamic methods work on the condition that an information system is being operated, thus indicating effectively the level of reliability of the IS);
− mathematical modeling of IS/AS (Mathematical modeling is used to predict the level of reliability of the information system at a certain point) [2,4];
− methods of decreasing mean time between failures (MTBF) and decreasing the failure rate of the information system (these methods are applied to tackle the identified problems in IS/AS).
Aiming to test the operability and correctness of the Algorithm, an experiment has been set up on the model of IS. The results of the experiment have been compared with the results of the existing IS. Based on the analysis of the data obtained, the conclusion has been formulated: within the framework of the assumptions and limitations coming from the use of the model, these ISs can be represented as a multiphase multichannel open-loop queuing system that is defined by the following features:

- input data stream;
- processing and combining data modules;
- data verification queues;
- multichannel service devices (control modules, processing module, data transmission modules);
- output data streams.

Exactly the same conclusion can be made as regards the majority of automatic control systems, where the data goes through the stages of modules. In this case, the data is being processed and changed, achieving the stage of output data, which can be used in managerial decision-making. In addition, it is also possible to use this approach for systems with dynamics of change, including that which is based on change management algorithms [9].

One of the main tasks of the experiment was to obtain data on the probable location (software/hardware module) and time (from the start of the computing process) of the occurrence of an error in the mathematical model of the IS. For the experiment, a simulator has been used. The decision of constructing a simulation model was made due to the availability of data on the functioning of the existing IS (the period for collecting statistical data is 5 years), as shown in table 1. During this period of time statistical duty has recorded all errors that occur in the IS, and the measures to eliminate them have been taken as well. For simulation, functional simulation software, AnyLogic produced by The AnyLogic Company, was used.

The necessity to apply the simulation in the presented experiment is based on the following factors:

- no need to use large amounts of computing power to run IS;
- possibility of altering the key parameters of the IS without exposing the IS to crash risk;
- possibility to select and combine the parameters of the simulation model that will allow to obtain accurate data based on the collected statistics (congruence of simulation results and actual statistical surveys) due to the proper selection of control coefficients;
- possibility of making predictions about the occurrence of failures in the IS based on the current simulation model;
- assessing the correlation of parameters of the simulation model, as well as reaching the adequate conclusions and recommendations concerning the real IS, including the use of time-to-failure methods.

Therefore, the use of the simulation model provides us with ample opportunities for carrying out a detailed analysis of the IS by setting up many experiments, without the risk of economic loss. There is an exceptional opportunity to make changes to the key parameters of the IS and its components [5,7].

Description of the studied IS.

| No | IS | Parameter | Description | Quantity |
|----|----|-----------|-------------|----------|
| 1  | IS 1 | IS module (unique type of module) | - receiving information - initial verification of received information - recording of received information | 6 |

1 To preserve the confidentiality of the information, the IS names have been changed, no unique names of the modules as well as no unique features are included in the description, and the IS structure is simplified as well, which allows preserving the accuracy of the data during the experiment, but not allowing to discredit the specific IS.
Data processing module (unique type of module)
- data alteration (the following types of modules are known: processing (conversion of information that involves minor changes), changes (getting totally different information comparing to initially received), storage and generation (getting new data without preliminary processing of received data, where the former is related to the latter))

Data type (unique data category)
The number of types of incoming and outgoing data streams that cannot be merged and that are processed in the separate IS modules

Time of functioning of IS (days)
The period of time during which the observation of the real IS has been conducted

IS 2

| IS module (unique type of module) | Data processing module (unique type of module) | Data type (unique data category) | Time of functioning of IS (days) |
|----------------------------------|-----------------------------------------------|---------------------------------|---------------------------------|
| - receiving information          | - recording of received information            | The number of types of incoming and outgoing data streams that cannot be merged and that are processed in the separate IS modules | 25 |
| - initial verification of received information | The IS function module, which performs certain operations on the data alteration (the following types of modules are known: changes (getting altered (moderated) information comparing to initially received), storage) | | 27 |
| | | The period of time during which the observation of the real IS has been conducted | 2000 |

IS 3

| IS module (unique type of module) | Data processing module (unique type of module) | Data type (unique data category) | Time of functioning of IS (days) |
|----------------------------------|-----------------------------------------------|---------------------------------|---------------------------------|
| - receiving information          | - data alteration (the following types of modules are known: processing (getting totally different information comparing to initially received), storage and generation (getting new data without preliminary processing of received data, where the former is related to the latter)) | The number of types of incoming and outgoing data streams that cannot be merged and that are processed in the separate IS modules | 5 |
| - initial verification of received information | | | 2000 |
| - recording of received information | | | |

Formal description.
According to the data presented, the IS models should consist of the following elements:
- simulation of data (of various types) entry;
- simulation of IS modules functioning;
- simulation of data processing modules;

These modules receive the same type of information from different sources and send it to the same type of modules for processing. The result of processing will be identical regardless of the module. This IS is a striking example that indicates the option for reducing the simulation to a linear structure.
− simulation of output data verification modules functioning;
− simulation of data output.
Logic diagrams for the IS models 1, 2 and 3 are represented in figure 1, 2 and 3.

Figure 1. Logic scheme of the simulation model for the IS 1.

Figure 2. Logic scheme of the simulation model for the IS 2.
In order to analyze the operation of the IS the following IS-model experiment has been set up:

Several types of data are fed through the input of the IS model. The intervals between occurrence of data of the same type are random. Each type of data enters the corresponding IS module, after which it is exposed to the processing, the result of which is arbitrary (data can be received, modified and fed through the output of the IS, as well as can cause the generation of new data in the IS), and it cannot take more than a predetermined period time that have been laid down for processing the received data. After processing, the data are verified in one of the emission modules. The data verification time is arbitrary. After verification, part of the data can be returned for revision, the rest of the data comes out of the IS model.

As a result of the IS operation, errors can occur in the modules. The number of errors in the modules depends on the workload (the amount of processed data at the same time) of the modules, while an error in an arbitrary module leads to downtime of the whole model of the IS. The IS downtime is contingent on the restoration process of the IS and it is clearly fixed.

As can be noted from the represented diagram of the simulation IS model, possible errors in the IS operation will occur in modules that are aligned with the reception, processing, storage, generation of new data and data alteration, which is confirmed by the statistical data on the operation of the existing IS.

In this regard, errors that caused failures in the IS operation were installed (artificially programmed) in the IS models. Programmed failures are of probabilistic (conditionally random) nature, which is related to the frequency of operation of a particular IS module. Moreover, each failure is recorded leading to the decreased probability of the next one. Thus, the simulation of failures caused by disruption of the software and hardware parts of the IS is carried out, and therefore the human factor impact is eliminated completely.

The refusal to consider errors caused by the human factor is motivated by the lack of objective ways how to preclude these errors. Moreover, data of related to this type of errors were excluded from the statistical data compilation.

2. Results from the analysis
The results of this experiment are:

- The operation stage of the IS was simulated, which can be compared with real data, the weights are adjusted, the simulation results are close to real ones.
- The predictions about the timing and location of failure occurrence in the IS is made.

The diagram represented in figure 4 shows the deviation of the operating results of the simulation model from the data gained from the statistical compilation based on the functioning of the existing IS. The unit of time is 1 calendar day. The diagrams are presented for three simulated ISs. As can be seen from the diagrams represented, the deviation in the estimated dates of the information system is not more than 10 calendar days, which is not significant within the long-term and medium-term planning.
(for a period of 3 or more years). The variability of deviation for IS1 is 28.51, for IS2 it is 34.27, and 39.84 for IS3.

Figure 4. Deviations in the values obtained by IS models.

3. Conclusion
As a result of this research the following conclusions can be drawn:

- The simulation model allows us to carry out an analysis with high accuracy, pending the formalization of the model and the application of assumptions.
- A hypothesis has been formed: all ISs (including those that include hardware, such as ACS, APCS, etc.) can be represented as a queuing system. The confirmation of the hypothesis will simplify the construction of a unified model of the IS, the use of which will make it possible to simulate an arbitrary IS.
- The results obtained, subject to adjustments, can be used in planning operational indicators.
- The Algorithm can be applied in ensuring information security of the IS as related to identifying potential location for errors and zero-day vulnerabilities [3].
- The model can be used to design complex information security systems for automated and information parts of strategic and critical information infrastructure objects.

References
[1] Popov A, Zolotarev V and Bychkov S 2014 Improving the reliability of information systems Siberian Journal of Science and Technology 3 55
[2] Azarnova T, Asnina N, Proskurin D and Polukhin P 2017 The formation of the Bayesian network structure of the process of testing the reliability of information systems Bulletin of VSTU 6
[3] Mitrokhin V and Ringenblum P 2014 Assessing the impact of information security threats on the accessibility of the telecommunications network TUSUR reports 2 32
[4] Litvyak R, Vorobyov S and Katsupeev A 2016 The model for studying the reliability of a high-availability information system based on architecture with redundant servers University news. North Caucasus region. Series: Engineering 2 190

[5] Lvov A, Svetlov M and Martynov P 2014 Improving the information reliability of digital systems with qam/cofdm modulation Izv. Saratov Univ. (N.S.), Ser. Math. Mech. Inform. 4-1

[6] Dorofeev A and Markov A 2015 Planning for business continuity and recovery Issues of cybersecurity 3 11

[7] Melnik E and Klimenko A 2017 A comprehensive method for organizing the management of a fault-tolerant information-management system based on multi-agent interaction Izvestiya TulGU. Technical Science 9

[8] Popov A, Zolotarev V and Bychkov S Implementation of a combined algorithm designed to increase the reliability of information systems: simulation modeling IOP Conf. Series: Materials Science and Engineering 155 (1) 012010 (doi:10.1088/1757-899X/155/1/012010)

[9] Styugin M, Zolotarev V and Parotkin N 2019 Open information systems protection by its continuous change Proc. 2019 RusAutoCon 8867600 (doi: 10.1109/RUSAUTOCON.2019.8867600)