IOP Conf. Series: Materials Science and Engineering **155** (2016) 012010 doi:10.1088/1757-899X/155/1/012010

Implementation of a combined algorithm designed to increase the reliability of information systems: simulation modeling[1]

A Popov, V Zolotarev, S Bychkov
Reshetnev Siberian State Aerospace University
31 "KrasnoyarskiyRabochiy" prospect, Krasnoyarsk, 660037, Russia.

E-mail: bychkovss_va@mail.ru

Abstract. This paper examines the results of experimental studies of a previously submitted combined algorithm designed to increase the reliability of information systems. The data that illustrates the organization and conduct of the studies is provided. Within the framework of a comparison of As a part of the study conducted, the comparison of the experimental data of simulation modeling and the data of the functioning of the real information system was made. The hypothesis of the homogeneity of the logical structure of the information systems was formulated, thus enabling to reconfigure the algorithm presented, – more specifically, to transform it into the model for the analysis and prediction of arbitrary information systems. The results presented can be used for further research in this direction. The data of the opportunity to predict the functioning of the information systems can be used for strategic and economic planning. The algorithm can be used as a means for providing information security.

Introduction

Strategic and economic planning are some of the most important directions of the company's activities, because they determine the development trends of the company.

The more parameters/indicators are taken into account, the higher accuracy of the planning is achieved. From an economic and strategic perspective, uninterrupted functioning of the information system (IS) an important variable that indicates the functioning of the company (no spendings on IS restoration and fixing of IS errors, smooth performance of business-related tasks)[1].

The authors of the current study have previously presented the paper in which they deal with the application options for the combined algorithm designed to increase the level reliability of information systems. During the stage that involved designing the combined methodology to analyze the reliability of information systems[2], 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 information system);

1 This work supports by the president's grant for young Russian scientists. Contract No. 14.Z56.15.6012.

– methods of mathematical modeling of information systems (Mathematical modeling is used to predict the level of reliability of an information system at a particular/specific moment/period of time);
– methods of decreasing MTBF and decreasing the failure rate of information system (aimed at tackling the problems that have already been identified)[3];

As a part of the experimental mathematical modeling of IS, a detailed analysis of the existing IS (their functioning and structure) has been done. Based on its results, the conclusion has been formulated: in terms of mathematical modeling, these IS can be presented as a polyphase multichannel open-loop queuing system that is defined by the following characteristics[4]:

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

The primary task of the experiment was to obtain data on the location and time estimated for the occurrence of error in the IS model that has been put to the test. For the organization of the experiment the aim of which was to determine the reliability of the IS the decision to use the simulation[5] was made. The decision to prepare a simulation model was made in view of the availability of the data on the functioning of this IS system over the period of 5 years. During this period of time, all the errors that have occurred in the IS have been recorded, and the measures to eliminate them have been taken as well. FTo construct the simulation, functioning software, AnyLogic produced by The AnyLogic Company, was used.

The necessity to apply the simulation in the presented algorithm is determined by the following factors:

– the possibility of altering the key parameters of the IS without exposing the real IS to crash risk;
– the 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);
– the possibility of making predictions about the occurrence of failures in the IS based on the current simulation model;
– the possibility to analyze the dependence of some parameters of the simulation model on the others, and to come to the adequate conclusions and recommendations concerning the real IS.

Therefore, the construction of a simulation model provides us with great opportunities for carrying out a detailed analysis of the IS performance, including an exceptional one, that is, to alter key parameters of the IS without causing damage to the system and for completing subsequent analysis of the obtained data in order to make conclusions and suggestions on how to fix the bugs in the IS[6].

The description of the studied IS: (In order to maintain confidentiality, the names of IS are changed; no unique names of the modules as well as no unique features are included in the description; the IS structure is simplified to the level where it still allows to maintain the accuracy of data during the experiment, but doesn’t allow to discredit the specific IS. The IS function modules, which performs the task of receiving information, its primary inspection and making of its registry record, receive the same type of information from different sources. This information is directed at the processing modules of the same type. This IS is a prime example that indicates the possibility of reducing the simulation to a linear structure.

Table 1. Input data for further research

No	IS	Parameter	Description	Unit	Quantity
1	IS 1	IS module	The IS function module, which performs the task of receiving information, its primary inspection and making of its registry record	PC.	6
		Data processing module	The IS function module, which performs certain operations on the data alteration (these are the types of modules 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))	PC.	6
		Data type	The number of types of incoming and outcoming data streams that cannot be merged and that are processed in the separate IS modules	PC.	8
		Time of functioning	The time during which the observation of the real IS has been conducted	days	2000
2	IS 2	IS module	The IS function module that performs the task of the receiving information, its primary inspection and making of its registry record	PC.	25
		Data processing module	The IS function module, which performs certain operations on the data alteration (these are the types of modules known: changes (getting totally different information comparing to initially received), storage)	PC.	27
		Data type	The number of types of incoming and outcoming data streams that cannot be merged and that are processed in the separate IS modules	PC.	25

		Time of functioning	The time during which the observation of the real IS has been conducted	days	2000
3	IS 3	IS module	The IS function module that performs the task of the receiving information, its primary inspection and making of its registry record	PC.	4
		Data processing module	The IS function module, which performs certain operations on the data alteration (these are the types of modules known: changes (getting totally different information comparing to initially received), generation (getting new data without preliminary processing of received data, where the former is related to the latter))	PC.	6
		Data type	The number of types of incoming and outcoming data streams that cannot be merged and that are processed in the separate IS modules	PC.	5
		Time of functioning	The time during which the observation of the real IS has been conducted	days	2000

Formal description

According to the data presented, the IS models should consist of the following elements (simulations of):

– data (of various types) entry;
– operation of IS modules;
– data processing modules;
– output data verification modules functioning;
– data output.

Logic diagrams for the IS models 1, 2 and 3 are presented in Fig.1., Fig. 2, and Fig. 3.

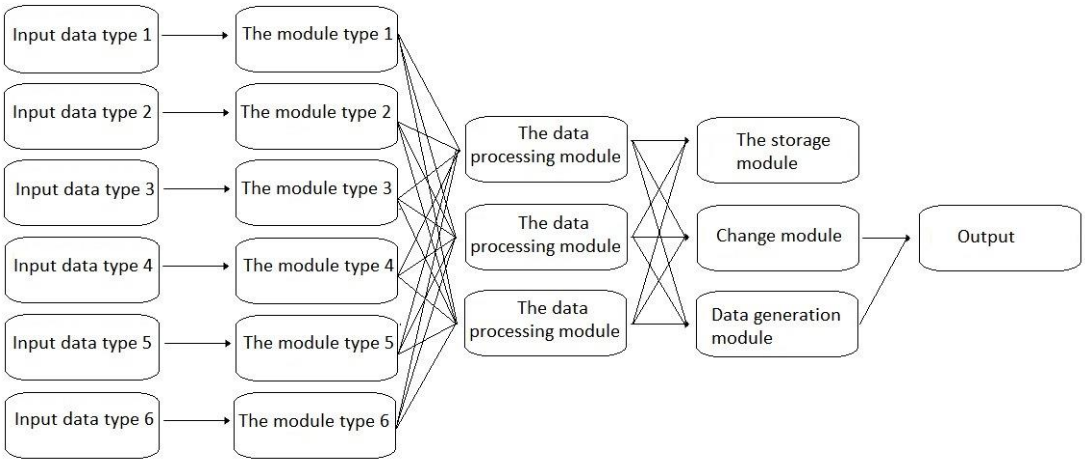

Figure 1. Logic scheme of the simulation model for the IS 1.

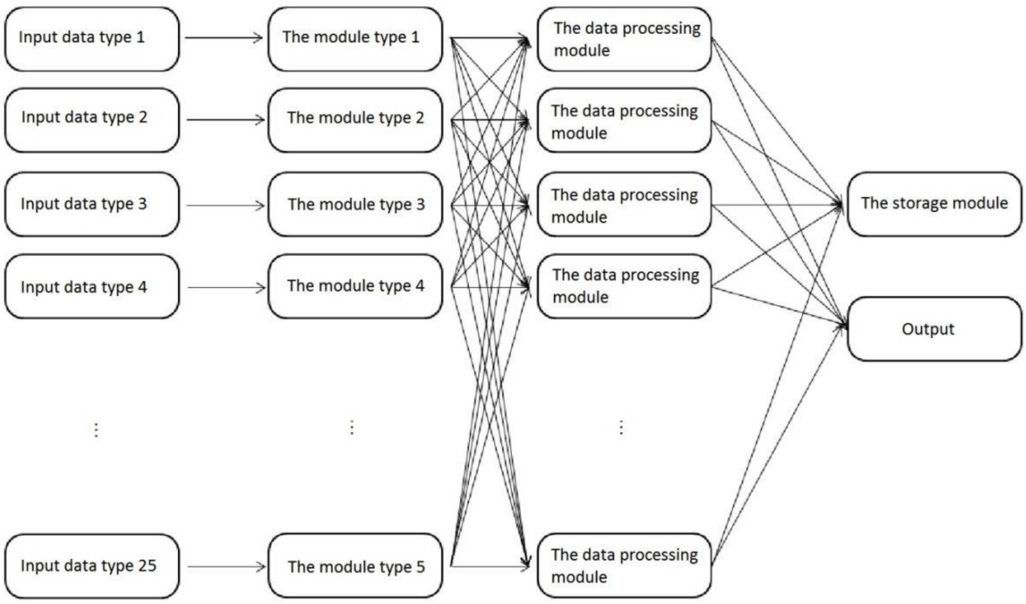

Figure 2. Logic scheme of the simulation model for the IS 2.

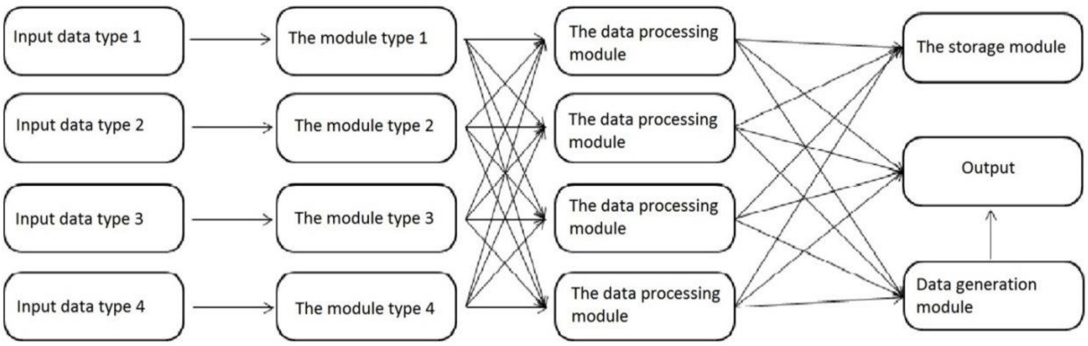

Figure 3. Logic scheme of the simulation model for the IS 3.

The following experiment was conducted on the IS models in order to analyze the operation of the IS.

There is a number of different data types fed through the IS input. The intervals between the feedthroughs of the data of the same type are random. Each type of the data enters the corresponding IS module and then it is exposed to the processing, the result of which is arbitrary (the data can be accepted, modified, and fed through the output IS, or it can lead to a generation of new data in the IS), and it cannot take more than a predetermined period time. After processing, the data is exposed to the procedure of verification in one of the emission modules, where the time taken to perform this procedure is arbitrary. After verification, a part of the data may be returned for revision, while the rest of the data comes out of the IS.

As a result of the IS operation, there can occur errors in modules. The number of errors depends on the level of workload of the modules, and the error in an arbitrary module leads to an IS downtime. The IS downtime is dependent on the IS restoration process and it is fixed and unchangeable.

As it can be noted from the presented circuit of the simulation IS model, possible errors in the IS operation will occur in the modules that perform the tasks of reception, processing, storage, generation of new data and data alteration, which is confirmed by statistical data gathered from the observations of the IS operation.

In this regard, artificially programmed errors that caused haltings (failures) in the IS operation were installed in the IS models. Programmed failures are probabilistic per se, which is related to the frequency of functioning of the particular IS module. Each halting (failure) is registered, thus leading to the decreased probability of the next one. Therefore, the simulation of failures caused by haltings in hardware and software parts of the IS is performed, which completely eliminates human factor effect. The decision to not consider the errors caused by human factor was based, first of all, on the absence of objective ways to avoid these errors. The statistical compilation of the data gathered from the observations of the current IS operation did not include the data related to this type of errors either.

Analysis of the results

As a result of the experiment, the following results have been obtained:

- both the modules of the IS that are prone to fail and the possible timing of its occurrence have been defined;
- a stage of the IS operation that can be compared to the real data has been simulated.
- The following table (2) presents selected results of the simulation IS model operation as well as the data gathered from the observations of the IS operation. It is important to emphasize that the dates presented are almost or completely identical to each other.

Table 2. Research results.

№	Information system	Fact	Experimental data (date)	Observational data (date)
1	IS 1	Module error 1	07/03/2012	15/03/2012
2		Module error 1	21/07/2013	17/07/2013
3		Module error 2	28/04/2013	24/04/2013
4		Module error 2	27/11/2013	02/12/2013

IOP Conf. Series: Materials Science and Engineering **155** (2016) 012010 doi:10.1088/1757-899X/155/1/012010

5		Module error 2	11/06/2014	06/06/2014
6		Module error 3	04/08/2014	08/08/2014
7		Module error 3	19/11/2015	17/11/2015
8		Module error 4	14/06/2015	14/06/2015
9		Module error 5	25/02/2015	28/02/2015
10		Module error 6	20/06/2014	18/06/2014
11		Module error 6	15/04/2015	14/04/2015
12	IS 2	Module error 1	11/02/2012	08/02/2012
13		Module error 1	10/08/2014	14/08/2014
14		Module error 1	23/07/2015	23/07/2015
15		Module error 2	9/03/2013	12/03/2013
16		Module error 2	21/09/2013	23/09/2013
17		Module error 2	12/04/2014	14/04/2014
18		Module error 3	16/07/2012	16/07/2012
19		Module error 3	27/11/2015	24/11/2015
20		Module error 4	23/04/2012	26/04/2012
21		Module error 4	19/02/2013	18/02/2013
22		Module error 5	27/01/2013	31/01/2013
23		Module error 5	24/08/2015	24/08/2015
24		Module error 6	05/11/2013	01/11/2013
25		Module error 6	25/03/2015	30/03/2015
26	IS 3	Module error 1	11/04/2012	11/04/2012
27		Module error 1	10/03/2014	14/03/2014
28		Module error 2	13/04/2012	09/04/2012
29		Module error 2	11/04/2013	11/04/2013
30		Module error 2	24/04/2015	27/04/2015
31		Module error 3	12/06/2012	08/06/2012

32		Module error 3	09/01/2014	09/01/2014
33		Module error 4	27/05/2012	30/05/2012
34		Module error 5	20/05/2014	15/05/2014
35		Module error 6	28/08/2013	29/08/2013
36		Module error 6	26/05/2015	25/05/2015

Also, a series of tests aimed at identifying interdependence between changes in the parameters of failures and errors occurrence in some modules and the occurrence of failures and errors in other modules was completed. The sample of the results of these tests is presented in the table below (3).

An important parameter in this case is a factor increasing the frequency of errors. Example: the given level of error in the test module = 0.1, depending on the module that logged n errors, the higher the level of errors is in the test module and in the identification of errors – the less regular flow of information will be transmitted to the next stage (as a result of the decrease in the input stream related to the dependent module, where the module detects the reduction of errors completed by 6%).

Table 3.Dependence of error occurrence on changes in the parameters (Impact of change in the coefficient on the grand total is chosen according to the initial standard levels of coefficient values).

№	Module name	Parameter	Dependent Modules (name)	Trend (correctly amount of processed data)	Rate of increase in the incidence of errors	Impact of change in the coefficient on the grand total
1	The module type 1	0,3	The data processing module	decline	10%	5%
2	The module type 2	0,5	The data processing module	decline	23%	12%
3	The module type 3	0,15	The data processing module	raise	1%	1%
4	The module type 4	0,01	The data processing module	raise	2%	1%
5	The data processing module	0,2	The storage module	raise	5%	4%
6	The data processing module	0,18	The storage module	decline	4%	4%

7	The data processing module	0,7	The storage module	decline	21%	18%
8	The data processing module	0,35	The storage module	decline	4%	4%
9	Data generation module	0	Output	raise	3%	3%
10	Data generation module	0,17	Output	decline	6%	6%

With reference to these data and according to the IS values of error coefficients, it can be concluded that the optimum values for the construction of three IS that correlate with the real statistical data gathered from observations are as follows:

Table 4. Optimal coefficients optimal for the analysis of the IS 3.

№	Module name	Parameter
	The module type 1	0,12
	The module type 2	0,08
	The module type 3	0,17
	The module type 4	0,04
	The data processing module	0,15
	The data processing module	0,03
	The data processing module	0,09
	The data processing module	0,21
	Data generation module	0,05

Conclusion

As for the results of the data analysis, some discrepancies in the results obtained and the actual IS model can be indicated. However, these results are based on the operation of such a IS model that does not take into account all the peculiarities of the IS operation and its architecture. The future studies will be concerned with the fixation of this error.

In spite of few flaws in the results, the authors of this study can list the following advantages and the positive results of the experiment:

– Authors propose the hypothesis that all the IS (the IS that complete within themselves a real business process that involves human participation) can be represented as queuing systems. In case this hypothesis is proved, the expenses on the construction of the improved IS model to obtain reliable results can be reduced. Thus, the opportunity to create a generalized model for the analysis of any IS becomes available.

– Adjusted for the error in calculations, the results can be used for economic and strategic planning. As seen from the data, the degree of readiness of the analyzed IS is high.

– Algorithm can also be applicable for dealing with the issues of the IS information security. The analysis that demonstrates the potential problems of the IS [7], as well as their actual occurrence, will prove its usefulness in the design of information protection systems. It is also possible to model the information security system itself in order to reveal its vulnerabilities and subsequently adjust them.

References

[1] Anderson R 2013 Measuring the Cost of *CybercrimeThe Economics of Information Security and Privacy*, Part IV, Springer Berlin Heidelberg Pp. 265-300.
[2] Baskerville R, Trevor Wood-Harper A A 2013 Critical Perspective on Action Research as a Method for Information Systems Research. *Enacting Research Methods in Information Systems* Vol. **2** Pp. 169-190.
[3] Allen S 2014 Tam Performance Monitoring with Application of Reliability Growth Analysis. *Engineering Asset Management – Systems, Professional Practices and Certification, Part of the Series Lecture Notes in Mechanical Engineering* Pp 1281-1289.
[4] Kim H, Choi B 2016 The Effects of Three Customer-to-customer Interaction Quality Types on Customer Experience Quality and Citizenship Behavior in Mass Service Settings *Journal of Services Marketing* Vol. **30.** Issue. 4.
[5] Macal C, North M 2014 Introductory Tutorial: Agent-based Modeling and Simulation. *Proceedings of the 2014 Winter Simulation Conference, IEEE Press* Pp. 6-20.
[6] Sargent R 2013 Verification and Validation of Simulation Models *Journal of Simulation* Vol. **7**, Iss. 1,. Pp 12-24.
[7] Shudrak M, Zolotarev V 2016 Improving Fuzzing Using Software Complexity Metrics *Lecture Notes in Computer Science* Vol. **9558** Pp. 246-261.