Preface

Special Issue on "Computer Science and Information Technologies" (CSIT)

Background

The 19th International Workshop on Computer Science and Information Technologies (CSIT 2017) was held in Germany, Baden-Baden, October 8-10, 2017. The Workshop was a forum for presentation of new results in research, development and applications in computer science and information technologies.

The previous Workshops on Computer Science and Information Technologies (CSIT) were conducted in Moscow (1999), Ufa (2000, 2001, 2003, 2005, and 2007), Patras, Greece (2002), Budapest, Hungary (2004), Karlsruhe, Germany (2006), Antalya, Turkey (2008), Crete, Greece (2009), Moscow-Saint-Petersburg, Russia (2010), Garmisch-Partenkirchen, Germany (2011), Ufa - Hamburg - Norwegian Fjords (2012), Vienna-Budapest-Bratislava (2013), Sheffield, England (2014), Rome, Italy (2015), Prague- Kunovice, Czech Republic (2016).

There were two more CSIT meetings since 2017 in Bulgaria (Varna, 2018) and in Austria (Vienna, 2019), their papers and results may be used for further special issues in Acta Polytechnica Hungarica or in other journals, depending on the demand and on the success of the recent work

As the title of the meetings (Workshop according to its name; conference, according to its value and size) is quite general, there are several topics, which may be discussed and were discussed, from the theoretical side and from the application side as well to give forum to several people. Just to give a short list of paradigms involved:

> *Artificial Intelligence, Decision Support Systems, Modeling, Simulation, Networking, Integration, Date- and Knowledge-Base Management, Data Mining, Big Data, Robotics, Industrial Automation, Green Production, etc.*
>
> *Applications in industry, management, finances and agriculture, etc,*

The workshops aim at bringing together researchers from different areas, including Distributed Systems, Management of Data and Knowledge, Computer Control. These areas are heavily related nowadays. Agent technology, Peer-to-Peer Information management and ubiquitous computing are just examples of the strong interrelationship between these disciplines. The Workshop organizers in the last years regularly provided a platform for cross disciplinary discussions in the framework of the announced topics.

Goals and Approach

During the past 20 years history of the workshops, taking into consideration the scientific level of the papers presented and published in the proceedings, having an increasing international participation we believe that the meeting became adult and strong enough to attract a broader international interest. The managers of CSIT and of the Acta Polytechnica Hungarica (the international journal of the Óbuda University) decided to jointly prepare and publish a special issue devoted to CSIT. Prof. Nafisa Yusupova (Ufa) and Prof. George Kovács (Budapest) were invited as guest editors.

Most of the papers were selected from the contributions of the 2017 Baden Baden meeting, however some earlier papers were added to increase the scientific level.

The initiator and main organizer of all CSIT workshops is the Ufa State Aviation Technical University (USATU) of the city Ufa.

USATU (UGATU in Russian) has a rather special role in the Russian University system, as fundamental university with old traditions. Founded in 1932, Ufa State Aviation Technical University is one of the largest universities of higher education in the Russian Federation. Serving 1500 faculty members and 20 000 students from 27 countries around the world. The university is situated on a 500-acre campus in Ufa, Republic of Bashkortostan.

The Faculty of Informatics and Robotics (Dean – Prof. Nafisa Yusupova) offers more than 40 academic options in several subject areas. The Faculty is proud of strong position of leadership in computing in the region, representing numerous industrial links and projects. A central part of our Informatics and Robotics Faculty's goals is to perform modern research in Informatics and Robotics by focusing our efforts in various specialization areas represented but not limited by: Algorithms and Artificial Intelligence, Programming Languages and Software Engineering, Computer Engineering and Systems.

Ufa is the capital of Bashkortostan, Russia. The city is situated in the Ural mountains closed to the Europe-Asia border, inhabited mostly by Russians, Tatars and Bashkirs. The city is cut to two parts by the Bjelaja river similarly to the river Danube, which cuts Vienna, Bratislava, Budapest and Belgrade to two-two parts.

According to most researchers recent Finland and Hungary both were populated by those people leaving Bashkir territories moving to the west in the early medieval years.

Organizing/Editing the Special Issue on CSIT

The papers of this special issue were selected from the 2017 workshop material. All authors were kindly requested to rewrite their contributions to match with journal requirements, and then the improved papers went through the regular journal review procedures. Finally 18 papers were selected for publication in this special issue.

The basic difference between a normal and a special issue is, that something keeps the papers of the special issue together, there is something common.

A special issue is generally published based on a given topic (open call or invitations) or selected from a meeting of 80-120 presentations. It generally consists of 15 to 30 papers collected into groups of 3-6 contributions, which are similar in some points of view, and there is a reasonable sequence of papers within each group. All are defined by the logic or structure of the papers. In our recent case the basis was one meeting (CSIT 2017) having about 60 contributions representing a huge selection of topics.

The editors wanted to organize the papers according to some guidelines, as for example the place of research activity on the "From Theory to Application" Line (TAL). The steps of technical innovation from basic research to industrial applications through different levels of applied research are always different and changing from time to time. Innovation is a very broad area, the selected papers may cover only some little segments of it.

> *We warn the reader from believing that the long forgotten slogan: "Science Goes To Work" is warmed up, just most aspects of innovation involve science working directly for us, for innovation, for our common future*

THEORY	*BASIC RESEARCH*			*APPLIED RESEARCH*			*PRACTICE*
TTTT	*G1*	*G2*	*G3*	*G4*	*G5*	*G6*	*RRRR*

Groups Sequences, Theory, Dreams, Reality, etc.

Based on the above given approach we planned to **start with pure theory** (basic research, TTTT), then the so called applied results is coming up, and finally applications, industrial realizations (RRRR) are to be discussed.

> *From a more scientific point of view one should mention PLM (Product Life-cycle Management), what details the complete life-span of every product (all innovation results and more) from birth of an idea through several design and implementation steps, not forgetting about tests, service, maintenance and reuse and recycling at the end.*

The normally simple question: which contributions belong together, which ones should be in the same group ?turned out to be too complex to answer, as all papers deal with 2 to 6 topics, and use 1-10 tools and/or methodologies; "the most relevant" ones of them is hard to choose. Finally the clusters of papers were formed by means of taking into account of several keywords of type topics ad of type tools & means, which were not necessarily the same, which were given by the authors.

We assigned simple keywords (as identifiers) to every paper, to decide their best places using several aspects, to help in using the "theory to industry" line (TAL).

All papers will be outlined soon, now it is just to see which papers fit into which basic categories (selected more or less arbitrarily), as, for example below

Information management (Im)	2 papers
Application oriented information management (Ima)	x papers
Application oriented information management (Ima)	11-x papers
CAD-CAM-industry oriented (ind)	2 papers
Maps, geography, government (gov)	2 papers
Control, industry (cont)	7 papers

Naturally several items could be defined differently, some issues may belong to more groups at the same time. All papers have their (proper) place in the theory to industrial application line (TAL), and the number of members in the groups is the most important.

In the 6 lines above the numbers identify the papers along the TA line. It is clear that their distribution is rather uneven, as the group Ima (No. 2. applications of information management) is overloaded, almost all papers belong to it. Pure theory (TTTT) and pure industry (RRRR) are very small subgroups, which cannot yet be seen.

Two remarks:

a) It will be seen that most papers have 3 or more authors, what suggests that these topics are complex enough to request that groups should work on them.

b) Another important information is that one or more universities and research organizations are involved in the special issue from the following cities: Ufa, Ekaterinburg, Dresden, Sankt-Petersburg, Moscow, Orenburg, Irkutsk, Izhevsk, Chelyabinsk, Baden-Baden, Budapest.

Groups/Clusters and Sequences

A commonly used clustering could be just to put the papers in a sequence according to the place of the first authors in the alphabet.

G1 - G6 six groups (or six clusters) for 12 papers, introducing all papers

The goal of the guest editors is to assist the reader in choosing papers to read or to study them, or to find relationships between old and newer contributions, or just to find given authors or topics, etc. The o IDs in the contents and in the coming lists have the form G.n.m, (G is constant, n is the group number (1-6) and m (1-18) is the sequence number of the paper in our list. The groups and sequences are based on the TAL (if possible and reasonable) in general and used for given tasks.

- Group No. 1 and No. 2. and partially No. 3 are dealing with basic software solutions, TTTT (plus risk based approach and data mining)

- Group No. 4 gives papers solving some 3D mapping, picture and territory management, which are not for the general public, but for any local authorities.
- Group No. 5 contains intelligent, multilevel control solutions, approaching to the RRRR level of the TAL.
- The last group (G6) presents industrially useful, practical applications (RRRR).

G1. Software design and development/first part *(general applications in different application fields, including applications of ontologies)*

G1.1 In their paper *"Analytical Toolset for Model-based Stochastic Error Propagation Analysis: Extension and Optimization Towards Industrial Requirements"* the authors (T. Fabarisov, N. Yusupova, K. Ding, A. Morozov and K. Janschek) discuss and solve some problems to help industry the application of MBSI by means of extending and optimizing the available analytical toolset to give industry a useful tool.

The paper with the hard, complex problems, which are properly defined and solved, could be called the flagship of CSIT, and it demonstrates the research strength and devotion of all participants of the workshops, and it reminds us that only the very best papers were selected for this issue.

G1.2 In the paper "*JSON Documents-Processing Using Situation-oriented Databases"* by V. Mironov, A. Gusarenko, N. Yusupova, Y. Smetanin the description of two sound approaches to an important problem of processing heterogeneous information are given. The authors use several tools to minimize home-made, unique software. For example, see the paper for details: JSON, XML, Document Object Model, web service, associative array, Query, Socrata Quary Language, Data extraction, Hypertext pre-processor, Situation Oriented DB. And these are working together to produce the two solutions. Heterogeneous data are managed in the previous paper as well.

G2. Software design and development/second part *(general to specific applications)*

G2.3 G. Kulikov, V. Antonov, A. Fakhrullina and L. Rodionova *in "Formal representation of the model for implementing system engineering functions because of the necessary diversity of structural relationships with the polycubic data organization"* define a formal representation of subject domain based on system model in the form of category of processes. The possibility of constructing a number of identical formal algorithms for controlling the integrity of the structure of initial processes is shown, and the necessary "external" control in case of violation is defined.

G3 Intelligent solutions – using data mining in medicine/toxicology *(data mining and risk based approach)*

G3.4 The paper of A. Massel ad D. Gaskova.3.09 is entitled "*Application of a Risk-based Approach to Identify Critical Ties"* describes the identification of critical facilities (infrastructure, equipment), which appear in the energy sector. The authors formulated a risk-based approach to decision-making in identification of critical facilities (infrastructure, equipment) in terms of cyber threats.

G3.5 In the paper *"Complex Analysis of Medical Data with Data Mining Usage"* the authors (N. Yusupova, G. Shakhmametova and R. Zulkarneev) present a new approach to the medical, in particular, the toxicological data-analysis, data-processing and a multilevel system realization. The three-stage technique for data analysis with data mining usage is offered. The results of the research, as relationship of gender and age with poisoning with drugs and/or narcotics are discussed in details.

G4 Working for government – tasks of territories &. map generation, university education

G4.6 In their paper "*The new combined method of the generation of a three-dimensional dense map of environment based on history of camera positions and the robot's movements"* the authors (A. Vokhmintcev and M. Timchenko) solve some problems of the development of adaptive methods to generate a three-dimensional combined dense map of the accessible environment. The map provides a required accuracy of reconstruction and it determines a pos4tion of a robot in a relative coordinate system. A new dynamic matching algorithm combining visual features and depth information are able to build global consistent maps is proposed. The performance of the proposed system in real environment is presented and discussed. Moving robots and cameras assist in realization.

G4.7 In the paper "*Analysis of the Territorial Systems State Based on a Complex of Indicators*" V. E. Gvozdev, O. I. Khristodulo and D. V. Blinova discuss and solve some reliability problems of the basic topological structures of GRID systems containing a minimal enough set of constructive elements. The construction of structural reliability models based on series- parallel reliability schemes are described. The application of these approaches create the basis for a multidimensional analysis of the GRID-systems' reliability characteristics in the design and development stages. This, in turn, creates the prerequisites for assessing the functional safety of such systems.

G4.8 In the paper entitled *"Energy-efficient technologies in the educational program of the architectural higher school"* the authors (Zakharova, Krivonogov, Kruglikov and Petunin) discuss the world trends in the modern construction in relation to ecological rational design according to "green" standards and applications of the building information modeling – BIM. There are ideas and

solutions how to implement "green" building in Russia. The key issue is chaging eduational systems in higher level schools. Intreresting examples prove the success of the efforts of the teachers ad students of the University of Architecture and Art and of the Ural Federal University meeting requirements of "green" standards.

The next two clusters contain 2-2 papers with useful, RRRR solutions. G5 details complex control problems, while G6 is devoted to CAD/CAM, CNC and scheduling.

G5 Complex, multilevel intelligent control – for industry

G5.9 In their paper "*Design of Multi-Level Intelligent Control Systems of Complex Technical Objects on The Basis of Theoretical-Information Approach*" the authors (B. Ilyasov, V. Vasilyev and S. Valeev) discuss an approach of designing advanced multi-level intelligent control systems for complex technical objects on the basis of theoretical-information. Their system allows to make project decisions during the design of control algorithms. The project demonstrates the usability of their ideas is the design of intelligent control systems for aviation gas-turbine engines.

G5.10 Another paper devoted to aircraft problems is "*Architecture of the security access system for information on the state of automatic control systems"* by I. Frid, A. M. Vulfin, V. V. Berkholts, D. Ju. Zakharov and K. V. Mironov. Security of the aircraft control and control system is a natural, basic demand in aircrafts. This paper discusses a problem of secure access provision via Web-service to a data base. This data base contains critical information about the parameters of the architecture of the security access system life cycle of complex technical devices (CTD). This article presents a logic model of access to a data base as well.

G6 Scheduling and production solutions - working in industry

The last group of papers in this issue deal with the solution of scheduling industrial problems, namely CNC control data-diode development and cloud scheduling.

G6.11 The paper "*Routing in CNC Cutting Machines: Engineering Constraints*" by A. Petunin and P Chentsov is devoted to some problems of sheet cutting optimization with engineering specifics. There are two hard restrictions often forgotten by CNC designers ad users: thermal expansion and sheet rigidity restrictions. The solution uses the dependence of the cost functions on the route, where the machine works (the cutting tool moves) and on the direction of tool motion on the contour.

G6.12 In the paper "*Effective Scheduling Method of The Cloud System of Collective Access to Virtual Working Environments"* the authors I. Bolodurina, L. Legashev describe a cloud system of collective access (CSCA) to virtual working environments as a means of providing an economically profitable remote access to paid and free software for educational institutions of secondary education. The problem of efficient CSCA scheduling to optimize the usage of cloud virtual machines and software licenses has been studied in details. The mathematical model of CSCA is presented. The UML class diagram of CSCA simulator is described. The statistical analysis of fitness function value distribution is performed.

Conclusion of the introduction

The guest editors made their best to make this special issue as user friendly as possible with commenting and introducing the serious scientific contributions – often using easy to understand expressions.

The reader can see and experience the broad spectrum and high scientific level of the CSIT meetings, the diversity of contributions from theory to practice in several different domains with different aspects of information management.

The guest editors are grateful to the editors of Acta Polytechnica Hungarica for the hard work they assisted to manage and publish this special issue and they are thankful to all authors, who contributed and last, but not least to all reviewers and their assistance powers who worked a lot to make the papers acceptable by increasing the quality, explaining the authors what and how should they improve.

Prof. Nafisa Yusupova
Ufa, Russia

Prof. György Kovács
Budapest, Hungary

Guest Editors

Model-based Stochastic Error Propagation Analysis for Cyber-Physical Systems

Tagir Fabarisov[1], Nafisa Yusupova[2], Kai Ding[3], Andrey Morozov[4], Klaus Janschek[1]

[1]Institute of Automation, Technische Universität Dresden, Georg-Schumann-Str. 11, 01187 Dresden, Germany, {tagir.fabarisov, klaus.janschek}@tu-dresden.de

[2]Faculty of Computer Science and Robotics, Ufa State Aviation Technical University, ul. K. Marx 12, 450000 Ufa, Russia, yussupova@ugatu.ac.ru

[3]Bosch (China) Investment Ltd., Corporate Research, Shanghai, China, kai.ding@cn.bosch.com

[4]Institute of Industrial Automation and Software Engineering, University of Stuttgart, Germany, andrey.morozov@ias.uni-stuttgart.de

Abstract: Industry 4.0 is the current trend of automation and data exchange in manufacturing technologies that is focusing on the creation of smart factories with the modular structured Cyber-Physical Systems (CPS), in tight cooperation with humans. This trend also implies that the systems become more complex, heterogeneous, and distributed especially their network and software parts. This makes the CPS highly critical subject to failures at different levels, including software, hardware, and human operators. Consequently, ensuring reliable and safe operation under the presence of non-avoidable threats also becomes a more complicated task. The proper analysis of the CPS requires thorough comprehension of both the dependability properties of system components and their interactions as well as structural and behavioral aspects of the complete system. Such an analysis of complex and mutually interlinked system properties puts considerable challenges on appropriate methods for modeling and analysis, as well as, on the related applied software tools. The Dual-graph Error Propagation Model (DEPM), developed in our lab, is a mathematical abstraction of the main future system's properties, which are vital for the determination of the error propagation processes. It is a useful analytical instrument for the evaluation of the influence of particular faults and errors to the overall system behavior. OpenErrorPro is our analytical software tool for stochastic error propagation analysis that supports the DEPM framework. Using OpenErrorPro, a DTMC model could be automatically generated from a DEPM, and the reliability metrics, in addition to, error propagation path, can be computed. This could be implemented for the analysis of the heterogeneous CPS components. The necessary steps for the DEPM framework extension, required for such an implementation, are discussed in this paper.

Keywords: Cyber-Physical System; Industry 4.0; Markov chain model; Probabilistic Model Checking; Error propagation model; reliability; safety; dependability; model-based system; model-based analysis; control flow; data flow; optimization

1 Introduction

Industry 4.0 is the current trend of automation and data exchange in manufacturing technologies that is focused on the creation of smart factories with the modular structured Cyber-Physical Systems (CPS) in tight cooperation with humans. This trend also implies that the systems become more and more complex, heterogeneous, and distributed, especially their network and software parts. This makes the CPS highly critical subject to failures at different levels including software, hardware, and human operators. Consequently, ensuring reliable and safe operation under the presence of non-avoidable threats also becomes more and more complicated task. The proper analysis of the CPS requires thorough comprehension of both dependability properties of system components and their interactions as well as structural and behavioral aspects of the complete system. Such an analysis of complex and mutually interlinked system properties puts considerable challenges on appropriate methods for modeling and analysis and on applied software tools.

Model-based System Engineering (MBSE) [1] approaches help both to simplify and speed up system development and provide semi-formal information for earlier system analysis. However, reliability and safety evaluation methods like Fault Tree Analysis and Failure Mode and Effect Analysis that are recommended in nowadays industrial standards fail to describe system behavioral aspects in a sufficiently deep manner. Additional sophisticated and highly specialized methods for the analysis of the effects of faults are required, such as the method for error propagation analysis which is discussed in this paper. Model-based analysis is MBSE-oriented analysis for the earlier design phases. One of such is the error propagation analysis (EPA). Employing EPA allows a design engineer to find and fix, or at least mitigate, the errors and flaws of the system design on its earlier design phase.

EPA can be based either on formal methods, e.g. probabilistic model checking techniques [2], or fault injection approaches. The formal analysis can be performed in early system design phases and strongly depends on the system model quality. Fault injections require either an already implemented system or its executable model. Analytical methods suffer from the state space explosion. The time required for the analysis can grow exponentially with the system model complexity. Usually, there is a kind of system model complexity threshold. After this threshold is exceeded, the required time grows so fast that the analytical approach becomes practically inapplicable. Conducting simulation-based analysis

with fault injections allows extending the number of experiments for reaching a wanted confidence level. However, fault injection methods have problems with the simulation of rare events, especially when we want to model faults that have extremely low activation probabilities.

2 State of the Art

2.1 Reliability and Safety Analysis

Classical methods, such as Fault Tree Analysis (FTA, IEC 61025), Markov Analysis (IEC 61165) [3], Failure Mode and Effect Analysis (FMEA, IEC 60812), and Hazard and Operability Analysis (HAZOP, IEC 61882) [4] are recommended in current safety standards. However, due to the increasing system complexity, these methods fail to describe aspects of modern mechatronic systems and CPS in sufficiently deep manner. FMEA, FMECA, and HAZOP are top-level qualitative methods that require numerical input from quantitative methods like FTA or Markov Analysis. For instance, the methods like FTA are not suitable for the reliability analysis of the systems with complex interactions between components and enhanced software parts. Markov methods can cover these aspects, but in most cases, the described solutions are either simple and high-level or prone to the state space explosion.

Model checking proposes advanced methods that allow a system designer, given a formal model of a system, exhaustively and automatically, check whether this model meets a given specification. Available model checking tools are based on state-based models such as state machines, discrete and continuous Markov chains, and Markov Decision Process models (e.g. PRISM, MRMC, STORM, SHAREP, SPIN, NuSMV) [5], Petri Net models (e.g. TAPAAL, ROMEO, ORIS, TimedNet) [6], or more high-level formal descriptions like AADL-based SLIM language (COMPASS) [7] or AltaRica language (OpenAltarica, ARC Studio) [8]. These powerful methods require specific and rather deep knowledge of the model checking such as discussed formal models and the formalization of the required system properties using temporal logic. Most of these methods are oriented to the manual and top-level analysis.

2.2 System Modeling

Several examples of widely used system modeling paradigms can be employed for the Industry 4.0 CPS systems. Structured Analysis [9] is rather old, relatively simple and straightforward design method that fits well for high-level functional design, small projects, and fast prototypes. In comparison with the modern

modeling methods, such as UML/SysML, the structured analysis doesn't provide enough design capabilities for complex system aspects. UML [10] with its extension SysML (Unified and Systems Modeling Languages) [11] is a popular and universal design approach. UML models can cover all phases of the system development life cycle. A great many UML/SysML design tools are available and integrated into industrial processes, including, MagicDraw, IBM Rational family, and UML Enterprise Architect. Another tools support automated UML-based MBSE including the "executable UML" and auto code.

Simulink and Stateflow [12] is the dominating modeling paradigm for dynamic systems, embedded control systems, and digital signal processing. Simulink provides a "native" interface for control engineers in the form of combined block diagrams (Simulink) and state charts (Stateflow) [13]. This Mathworks toolset includes built-in code generation and deployment mechanisms. This is also a perfect method for fast prototyping and software and hardware in the loop. The Mathworks tools do not support well early phases of the development such as functional design, or top-level composition. Therefore, Simulink/Stateflow is often used in together with other composition models like general UML or AADL.

AADL (Architecture Analysis & Design Language) [14] is a united framework for the model-based engineering of embedded real-time systems. AADL is strongly oriented to software and hardware co-design of real-time systems. One of the key features of this method is the inherited analytical capabilities. AADL has interfaces with analytical tools including COMPASS [7] for reach dependability and performance analysis, PRISM [5] for stochastic model checking, and OpenFTA [15] for fault tree analysis. AADL, like the Simulink/Stateflow, is also very specific and has relatively poor tool support (OSATE [16]) in comparison with Simulink/Stateflow or UML.

Human-in-the-Loop, Human-in-the-Loop Cyber-Physical Systems, Cyber-Physical-Human Systems, and "Internet of All" concepts share the common idea to consider the human as a part of a larger Cyber-Physical System (CPS). This idea has drawn considerable interest in recent years, as it is comprehensively surveyed in [17].

2.3 Dual-Graph Error Propagation Model

The analytical methods for model-based system reliability and safety analysis is required in order to ensure the reliable and safe operation under the presence of non-avoidable in the numerous applications in Industry 4.0. The proper analysis of the CPS requires thorough comprehension of both dependability properties of system components and their interactions as well as structural and behavioral aspects of the complete system. Such a method will find its usage in safety critical industrial domains including aerospace, automotive, transportation, medical and robotics applications, where a failure or malfunction may result in environmental

harm, severe equipment damage or even serious injuries of the personal. Industry 4.0, which is the current trend of automation and data exchange in manufacturing technologies, creates what is called a Smart Factory that consists of the modular structured Cyber-Physical Systems in tight cooperation with humans. With a robust increasing of these trends, aforementioned will find its implementation in the context of smart factory's reliability and safety and, therefore, will benefit to the environmentally friendly manufacturing technologies.

Therefore, the analysis of fault activation, error propagation, and timing properties of a given CPS is viable in order to ensure the safe operation. The error propagation analysis can be based either on formal methods, e.g., probabilistic model checking techniques, or fault injection approaches. The formal analysis can be performed in early system design phases and strongly depends on the system model quality, and subsequently be used when using the fault injection approaches.

The Dual-graph Error Propagation Model (DEPM) [18], developed in our lab, is a mathematical abstraction of the main future system's properties, which are vital for the determination of the error propagation processes. It is a useful analytical instrument for the evaluation of the influence of particular faults and errors to the overall system behavior.

DEPM framework describes the process-oriented model that allows the reliability modeling of heterogeneous CPS components, their interaction, nontrivial failure scenarios, multiple failure modes, hierarchical compositions, data errors propagation, timing aspects, and sophisticated control and data flow structures with branching, loops, and guarded stochastic transitions. The DEPM allows the computation of the reliability metrics using underlying Discrete-time Markov chain (DTMC) models. DEPMs can be automatically generated from common CPS models including the Simulink/Stateflow models [19], UML models [20], SysML models [21], AADL [22] models, as well as software source code using LLVM [23]. This allows not only the automated application of our method, but also the analysis of systems developed with a combination of modeling paradigms. For example, Simulink/Stateflow is often used together with other composition models developed using UML or AADL. The DEPM is a mathematical model that captures system control and data flow structures and reliability properties of system components.

An example of a Dual-graph Error Propagation Model, that describe a reference Cyber-Physical System is presented in Figure 1. This system consists of two autonomous and connected cars, and a navigation system connected via a network. Two cars are equipped with controllers and a set of sensors. Controllers collect the data from sensors regarding the obstacles on the course, as well as the latitude and longitude coordinates of a car. These data are being transmitted between two cars in order to ensure the safety. The critical failure of the system is defined as *"sensor does not provide necessary data."* Another failure of the system could be

described as *"network is down."* Assume that the sensor system is prone both to transient, bounded in time, and permanent, continuous in time, faults. Our system is tolerant to transient faults since the coordinates from the navigation system allows the system to survive during short sensor downtime. The permanent sensor fault leads to the system failure in case that the network is down unless an operator detects the failure and fixes it.

The DEPM combines two directed graph models: A Control Flow Graph (CFG) and a Data Flow Graph (DFG). The nodes of the graphs represent executable system elements and (generic) data storages. The rounded rectangles with black borders represent the elements *aController*, *aCar*, *aSensors*, *Network*, *NavSystem*, *bCar*, *bSensors*, *bController*, and one special element *RepairMode* that models the corrective actions. The CFG arcs model the control flow transitions (black lines) between the elements. The DFG arcs model the data flow transitions (blue lines) between the elements and data storages (rectangles with blue borders). The data storages are specified with finite sets of string or integer values. For example, data *Ping* takes values in a range *[0, 300]* that represent the network latency and *aNetState* takes values between *{OK, FAIL},* representing the operational and fail states of the network. The DEPM also supports hierarchical systems. Elements *aSensors* and *bSensors* of a top-level DEPM contain an internal DEPMs with a set of *Sensor* elements. These elements are sub-systems, that contains a DEPM models itself. While the reliability metrics are being compute, these models should be calculated first in order to compute the high-level model.

The cyclic process starts with the execution of the initial element *Network* and in a nominal case continues with the execution of either *aCar*, *bCar* or *NavSystem*. The control flow transition from *Network* to *RepairMode* represents the network failure detection. This is a guarded stochastic transition specified by control flow commands of the element *Network* shown in the gray rectangle expressed in PRISM input language.

The command *"0.01:(cf'=RepairMode)&(NetState'=FAIL)"* specifies that the process jumps from *Network* to *RepairMode* with the probability 0.01 only if *Ping* less than 200. Similar to the control flow commands, the elements can have data flow commands that specify the fault activations during the element execution and the propagation of the errors from data inputs to outputs. For example, the data flow commands of the element *aController* manage the data *aObstacles*. The commands of the element *aController* model transient and permanent generator faults. The transient fault *"(aNetState'=FAIL)&(aSenState'=OK)"* occurs with the probability 0.04 and the permanent fault *"(aNetState'=FAIL)& (aSenState'=FAIL)"* with the probability 0.02. As we can see, in case of the permanent fault, the value of *aSenState* is changed to *FAIL* and stays *FAIL* unless we jump to the element *RepairMode*. The red arrow-shaped nodes such as *SensorFailure* and *PackageDrop* specify system failures using expressions that contain the values of data storages and the special variable *cf* that models the current control flow state. A fixed execution time is defined for each element in

order to compute time-related reliability metrics such as MTTF or mean downtime. In this example, the execution of the element *RepairMode* takes much longer (300s) than the execution of the *Network* elements (1s).

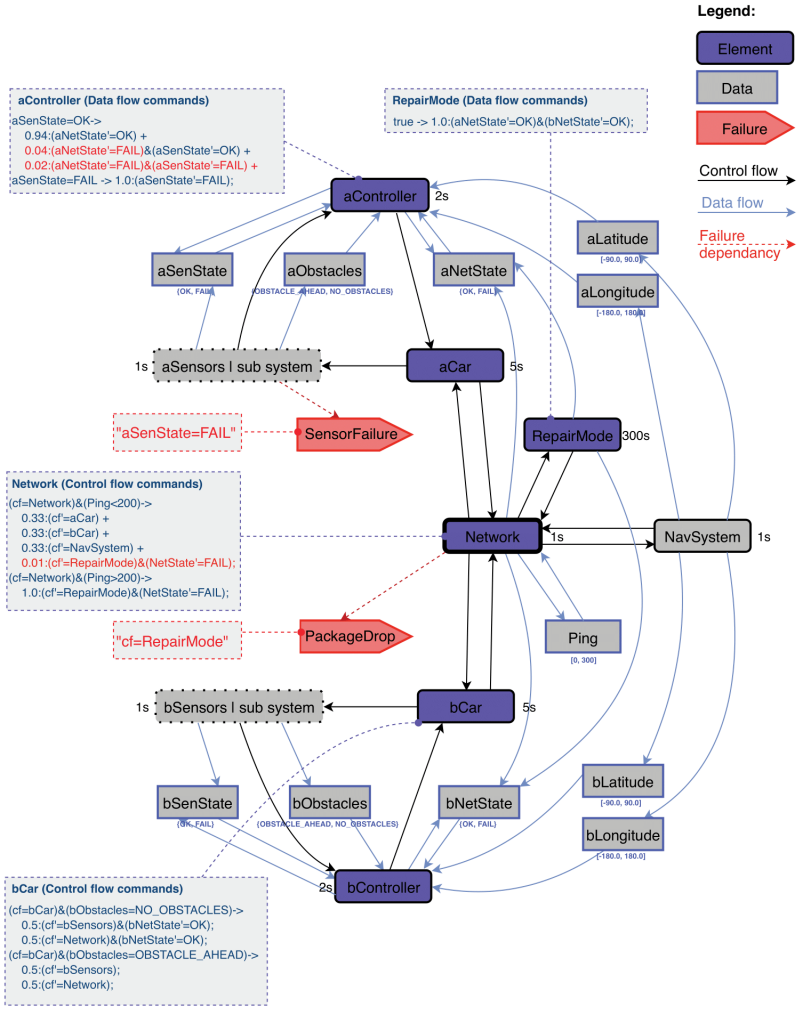


Figure 1
An example of a Dual-graph Error Propagation Model for Cyber-Physical System

3 Analytical Toolset for Error Propagation Analysis

3.1 Analytical Toolset for Model-based Stochastic EPA

Using the DEPM framework described in the previous chapter allow a designer to calculate the dependability metrics, e.g. a mean number of errors, mean time to failure etc., using discrete-time Markov chain (DTMC) models. A DTMC model describes system dynamics as a stochastic process in terms of errors occurrence and propagation. OpenErrorPro [24] is our analytical software tool for stochastic error propagation analysis, that support DEPM framework. Using OpenErrorPro a DTMC model could be automatically generated from a DEPM and the reliability matrices cold be calculated, for instance, the MTTF metric for *"SensorFailure"* failure (Figure 1).

OpenErrorPro is implemented in Python and consists of three main parts: error propagation library, graphic user interface (GUI), and model transformation algorithms. For efficient computation of the generated DTMC models, that can be formally analyzed, the OpenErrorPro uses a built-in interface with PRISM software [5]. The tool also integrates optimization methods against state space explosion of DTMC models such as the automatic nesting algorithm and data flow slicing [25]. The model transformation algorithms allow automatic generation of DEPMs from the baseline system models UML/SysML [21], Simulink/Stateflow [19], AADL [22]. After that OpenErrorPro automatically creates discrete time Markov chain models using the DEPM representation and computes required numerical reliability properties using PRISM software. The most relevant and close methods and tools are high-level model checkers COMPASS, OpenAltaRica, and Figaro, which are discussed and compared with OpenErrorPro in [24]. OpenErrorPro enables the reliability-related features analysis of CPS that cannot be modelled by other methods for quantitative system-level reliability analysis. In Figure 2 a GUI of the toolset with an open DEPM model of the CPS presented in previous chapter is shown.

3.2 Model-based Stochastic EPA for Cyber-Physical Systems

The DEPM framework as well as OpenErrorPro tool have been applied for several case studies during the last years [26]. Different types of mechatronic systems have been analyzed, including the moving and flying robots, and specific model-based software systems from the space and automotive industrial domains. The DEPM is a powerful analytical instrument for the modeling of a broad set of dependability-related features that influence error propagation processes. The feasibility of the DEPM approach has been witnessed by a number of case studies including navigation system of a mobile robot [27], automated medical patient

table motion control [22], robotic software for space-to ground haptic feedback control [18], and embedded flight control software of a UAV [21]. DEPM framework could be applied for CPS in the analysis of the propagation of typical CPS errors from potentially faulty components to technically accessible system variables, e.g. network, sensor outputs, the software of the controllers, embedded boards, and other computing units. As it was presented in [24], the DEPM framework could be applied for the complex model design. OpenErrorPro helped to perform the sensitivity analysis of system reliability to identify the places where it makes sense to apply system redundancy. In the work, authors have encountered state space explosion issues of the underlying Markov chain models. Nevertheless, they resolved them with the optimization trick using the fact, that on the DEPM-level OpenErrorPro provides a fully automated access to the control and data flow structures.

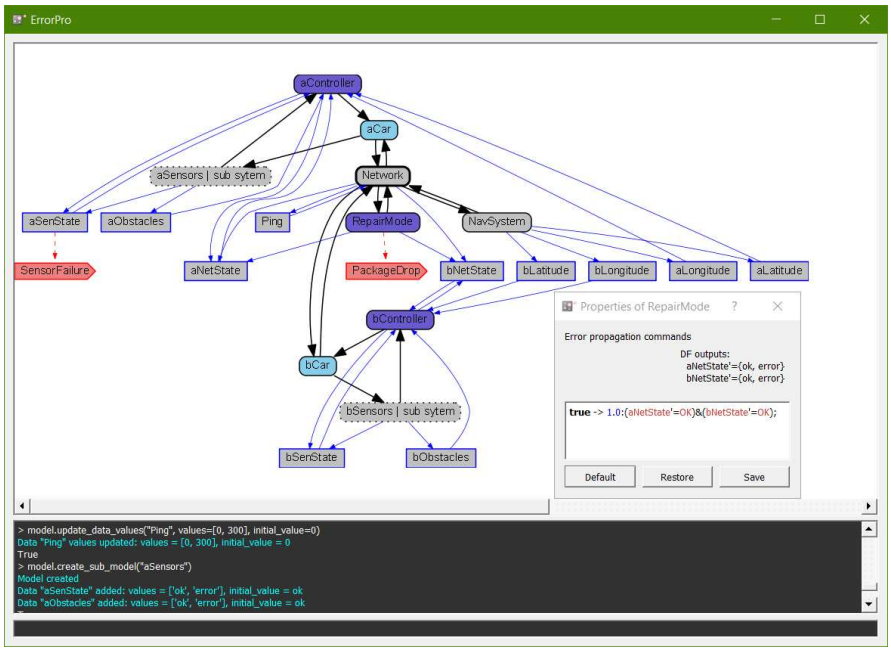


Figure 2

A GUI of the OpenErrorPro toolset with DEPM of a Cyber-Physical System. Note how in *"Properties of RepairMode window"* the data flow commands represents a repairing after the failure.

For the implementation of DEPM in CPS reliability analysis, there is required a generic approach for the automated creation of a formal hierarchical DEPM from semi-formal heterogeneous CPS models. It can be divided in the following steps. First, the given CPS is being decomposed into individual components and mapped into the DEPM elements. While performing this step, there must be a persistent conceptual separation between hardware and software parts, as it's done

previously in [28] for hardware and in [19] for software. The hardware-level DEPMs will model sensors, actuators, network, and computing units. The software of the computing units will be modelled as sub-DEPMs of the corresponding hardware components. For the model-based software, e.g. implemented in Simulink, the block functions will play the role of the DEPM elements [18]. For the manually implemented software LLVM-based method [29] will be used that decomposes the code into single instructions. Second step is to map the CPS data flow structure into DEPM data storages and data flows. This information is commonly available in some specific format, e.g. SysML Internal Block Diagrams, UML Component Diagrams, AADL models for hardware or Simulink block diagrams for model-based software. The key challenge is the integration of the DEPMs generated from several types of baseline models in such a way that the final hierarchical DEPM will capture comprehensively the propagation of the common CPS errors and their combinations. For that the design of a generic approach to extend the DEPM with the CPS error types and define corresponding error propagation rules is needed. After that, the next step is to mark the DEPM data storages that model technically accessible variables and critical variables, e.g. actuator commands.

The propagation of an error to a critical DEPM part will cause a system failure. The DEPM, generated following the discussed steps, will contain no probabilistic parameters. The further extension of the DEPM with the probabilistic parameters will enable the quantitative estimation of the probability that errors will be detected before they propagate to the critical system parts. First, the information about the system operational profile will be added, such as sensor rates, network transmission times, sample times and the control flow structure of the software components. This information will be mapped into the DEPM control flows either from available behavioral models, e.g. UML/SysML Activity Diagrams, State Machines, or AADL models with timing properties, or via the extension of the system with logging, execution, and evaluation of the gathered statistical information about the execution patterns of the CPS components. One of the challenges is to distinguish between stochastic and deterministic control flow events taking into account the imperfect synchronization of the CPS components. Finally, the DEPM must be extended with the probabilities of the error propagation through the components and fault activation. The error propagation from inputs to outputs of hardware components can be described either deterministically from the known component specification or using local fault injections. The error propagation through the software parts can be computed with the corresponding sub-DEPMs, as it shown in [24]. Fault activation probabilities can be evaluated using the available sources of the failure data such as FIDES [30], the Electronic Parts Reliability Data (ERPD) [31], or the Non-Electronic Parts Reliability Data (NRPD) [32]. However, these sources address the “complete” component failures. The probabilities for different error types are extremely hard to find. Thus, all the available probabilistic information will be added into the DEPM and cover unknown probabilities with local

nondeterminism. OpenErrorPro, discussed in previous chapter, supports automatic generation of DTMC (Discrete-time Markov Chain) models from a DEPM and analyzing them using the interface with modern probabilistic model checkers such as PRISM [5] and Storm [33]. These model checkers support local nondeterminism, however, in the future we will need to switch from DTMCs to (Markov Decision Process) models. The more information we feed to DEPM the more precise the estimation will be. The nondeterminism will result in several equally suitable sets, that have to be considered, in the subsequent processing steps of DEPM approach.

Conclusion

The Dual-graph Error Propagation Model has been discussed in this paper. DEPM is a process-oriented model that allows the reliability modeling of heterogeneous components, their interaction, nontrivial failure scenarios, multiple failure modes, hierarchical compositions, data errors propagation, timing aspects, and sophisticated control and data flow structures with branching, loops and guarded stochastic transitions. The DEPM allows the computation of the reliability metrics using underlying Discrete-time Markov chain models. DEPMs can be automatically generated from common CPS models, including the Simulink/Stateflow, UML/SysML, AADL models, as well as, software source code using LLVM software for stochastic error propagation analysis. OpenErrorPro error propagation tool, based on the described DEPM, has also been discussed in this article. This tool plays the role of an intermediate analytical model between the baseline system model and the formal mathematical model that can be formally analyzed. It can automatically create discrete-time Markov chain models using the DEPM representation and calculate the required numerical reliability properties. In addition to traditional reliability metrics, OpenErrorPro enables the evaluation of customizable reliability metrics and the application of effective optimizations, against the state space explosion of underlying Markov chain models already on the DEPM level. This toolset has been proven to have functionality and capability, for implementation in the analysis of the heterogeneous components of a Cyber-Physical System. As it was discussed herein, OpenErrorPro could be applied for the analysis of CPS. In order to be able to generate a formal DEPM from the CPS representation, future efforts will expand the DEPM approach, including such capabilities as discussed in this paper.

Acknowledgment

This work has been supported by DAAD “Michail Lomonosov” program jointly with Russian Ministry of Science and Higher Education funds.

References

[1] Gianni, Daniele; D'Ambrogio, Andrea; Tolk, Andreas, eds. Modeling and Simulation-Based Systems Engineering Handbook (1 ed.) USA: CRC Press, 2014, ISBN 9781466571457

[2] C. Baier and J.-P. Katoen. Principles of model checking. MIT press, 2008

[3] N. B. Fuqua, "The applicability of markov analysis methods to reliability, maintainability, and safety," Selected Topic in Assurance Related Technologies (START), Vol. 2, No. 10, pp. 1-8, 2003

[4] E. Ruijters and M. Stoelinga, "Fault tree analysis: A survey of the stateof-the-art in modeling, analysis and tools," Computer science review, Vol. 15, pp. 29-62, 2015

[5] Marta Kwiatkowska, Gethin Norman and David Parker. PRISM 4.0: Verification of Probabilistic Real-time Systems. In Proc. 23rd International Conference on Computer Aided Verification (CAV'11), volume 6806 of LNCS, pp. 585-591, Springer, 2011

[6] Katoen J P, Zapreev I S, Hahn E M, et al. The ins and outs of the probabilistic model checker MRMC [J]. Performance evaluation, 2011, 68(2): 90-104

[7] Bozzano M, Cimatti A, Katoen J P, et al. The COMPASS Approach: Correctness, Modeling and Performability of Aerospace Systems[C]//SAFECOMP. 2009, 5775: 173-186

[8] Arnold A, Point G, Griffault A, et al. The AltaRica formalism for describing concurrent systems [J]. Fundamenta Informaticae, 1999, 40(2, 3): 109-124

[9] E. Yourdon, Modern Structured Analysis. Yourdon Press, 1989

[10] OMG. OMG Unified Modeling Language (OMG UML), 2019

[11] Delligatti, Lenny. SysML Distilled: A Brief Guide to the Systems Modeling Language, Addison-Wesley Professional, 2013, ISBN 978-0-321-92786-6

[12] Mathworks, Matlab & Simulink: Simulink User's Guide R2019b. Retrieved 2020

[13] Mathworks, Matlab & Simulink: Stateflow User's Guide R2019b, Retrieved 2020

[14] P. Feiler, D. Gluch, "Model-based Engineering with AADL: An Introduction to the SAE Architecture Analysis & Design Language," Addison-Wesley Professional, 2012

[15] Formal Software Construction Limited, OpenFTA, Version 1.0, User Manual

[16] Carnegie Mellon University, "Welcome to OSATE," Online: http://www.osate.org

[17] D. S. Nunes, P. Zhang, J. Sá Silva, "A survey on Human-in-the-Loop applications towards an Internet of All", IEEE Communications Surveys & Tutorials, Vol. 17, No. 2, 2015

[18] Morozov, Andrey & Janschek, Klaus & Krüger, Thomas & Schiele, André. (2016) Stochastic Error Propagation Analysis of Model-driven Space Robotic Software Implemented in Simulink. MORSE '16, Proceedings of the 3rd Workshop on Model-Driven Robot Software Engineering: 24-31

[19] K. Ding, A. Morozov, and K. Janschek. Reliability evaluation of functionally equivalent simulink implementations of a pid controller under silent data corruption. In 2018 IEEE 29th International Symposium on Software Reliability Engineering (ISSRE), pp. 47-57, IEEE, 2018

[20] K. Ding, T. Mutzke, A. Morozov, and K. Janschek. Automatic transformation of uml system models for model-based error propagation analysis of mechatronic systems. IFAC-PapersOnLine, 49(21):439-446, 2016, 7th IFAC Symposium on Mechatronic Systems MECHATRONICS 2016

[21] M. Steurer, A. Morozov, K. Janschek, and K.-P. Neitzke. Sysml-based profile for dependable uav design. In 10th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS), 2018

[22] A. Morozov, T. Mutzke, B. Ren, and K. Janschek. Aadl-based stochastic error propagation analysis for reliable system design of a medical patient table. In IEEE Annual Reliability & Maintainability Symposium (RAMS), 2018

[23] A. Morozov, Y. Zhou, and K. Janschek. Llvm-based stochastic error propagation analysis of manually developed software components. In Proceedings of the European Safety and Reliability Conference (ESREL18), 2018

[24] Morozov, A., Ding, K., Steurer, M., & Janschek, K. (2019, October) OpenErrorPro: A New Tool for Stochastic Model-Based Reliability and Resilience Analysis. In 2019 IEEE 30th International Symposium on Software Reliability Engineering (ISSRE) (pp. 303-312) IEEE

[25] Zhao F, Morozov A, Yusupova N I, et al. Nesting algorithm for dual-graph error propagation models[C]//CSIT'2016. 2016: 106-110

[26] Fabarisov, T., et al. "The efficiency comparison of the prism and storm probabilistic model checkers for error propagation analysis tasks." Industry 4.0 3.5 (2018): 229-231

[27] A. Morozov and K. Janschek. Probabilistic error propagation model for mechatronic systems. Mechatronics, 24(8):1189-1202, 2014

[28] A. Morozov, K. Janschek. Flight Control Software Failure Mitigation: Design Optimization for Software-implemented Fault Detectors. In Proceedings of 20th IFAC Symposium on Automatic Control in Aerospace ACA 2016, Sherbrooke, Quebec, Canada, 21-25 August 2016

[29] Vidineev, Viacheslav, et al. "LLVM-based C to DEPM transformation tool: New functionality and performance improvements." Информационные технологии интеллектуальной поддержки принятия решений. 2019

[30] F. Guide et al. Reliability methodology for electronic systems. FIDES group, 2009

[31] W. Denson, P. Jaworski, W. Crowell, and D. Mahar. Electronic parts reliability data 1997. 1996

[32] W. Denson, G. Chandler, W. Crowell, A. Clark, and P. Jaworski. Nonelectronic parts reliability data 1995, Technical report, RELIABILITY ANALYSIS CENTER GRIFFISS AFB NY, 1994

[33] C. Dehnert, S. Junges, J.-P. Katoen, and M. Volk. A storm is coming: A modern probabilistic model checker. In International Conference on Computer Aided Verification, pp. 592-600, Springer, 2017

JSON Documents Processing Using Situation-Oriented Databases

Valeriy Mironov[1], Artem Gusarenko[1], Nafisa Yusupova[1], Yuriy Smetanin[2]

[1]Ufa State Aviation Technical University, Computer Science & Robotics Dept., Karl Marks st., 12, 450008 Ufa, Russia;
{mironov, gusarenko}@ugatu.su, yussupova@ugatu.ac.ru

[2]Russian Foundation for Basic Research, Head of Information Technologies & Computing Systems Dept., Leninsky Prospekt st., 32a, 119334 Moscow, Russia;
smetanin@rfbr.ru

Abstract: Situation-oriented databases provide processing of documents from heterogeneous data sources under the control of a hierarchical situational model. This article discusses the problem of processing database documents in JSON format, along with XML. Two implementation approaches are discussed: (1) on the fly JSON to XML document conversion and using Document Object Model for processing XML, and (2) loading the JSON document into an associative/indexed array followed by applying the template engine. The database interpreter works with external heterogeneous data extracted from files, databases, archives, web services, data is processed using virtual documents. Examples of processing JSON documents received from a web service are analyzed. Data from the San Francisco Open Data web server is used as the JSON test source. Query in Socrata Query Language used for JSON data extraction is presented. The implementation of approaches in the research situation-oriented database prototype based on Hypertext Preprocessor is considered.

Keywords: hierarchical model; situation-oriented databases; NoSQL; databases; model-driven development; open data

1 Introduction

Web applications that are widely used now, in the course of performing their functions, should support the interaction, on the one hand, with users and other applications, on the other hand, with a variety of local and remote data sources. Traditionally, relational databases were used as data sources, which caused the problem of mapping [1] relational data to data [2] in other formats, such as XML [3]. However, Now the range of data formats used has significantly expanded, in

particular, thanks to the success of the NoSQL [4] movement (such as Key-Values databases, JSON-oriented document databases, Graph databases) and Polyglot Persistence [5]. In this regard, new approaches to the flexible processing of heterogeneous data in web applications are needed. The development of modern web applications is strongly influenced by new approaches such as Service-Oriented that estimated in work [6] in aspect of characteristics of Model-Driven Architectures which use domain-specific languages [7]. Recently, the experience of creating situation-oriented web applications has become relevant [8], applicable for NoSQL in model-driven [9] database development such as SODB together with Big Data technologies [10].

The ideas put forward in the framework of these approaches have had multiple impacts, including on Situation-Oriented Databases (SODB) [11]. SODB is a new approach to developing web applications, driven by the embedded Hierarchical Situation Model (HSM) [12]. It is based on the uniform processing of heterogeneous data [1]. The SODB approach novelty in the field of heterogeneous data integration is the concept of virtual documents that are mapped to real data in various formats (Fig. 1). When a virtual document is declared in the HSM model, it's mapping to real data is specified. After that, the processing of this virtual document can be specified with the reference to its declaration. The advanced invariance principle assumes that changing the mapping [2] specifications should not affect the specifications for processing virtual documents.

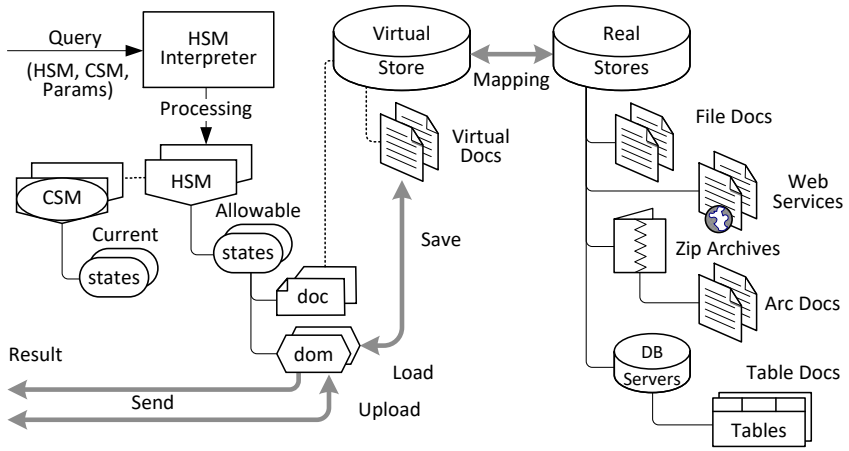


Figure 1
SODB architecture

Initially, SODB focused on processing virtual documents in XML format and on the using of appropriate XML technologies such as XSLT [11]. Recently, for XML has appeared a competitor such as JSON. Due to its simplicity and laconism, JSON has become popular and is actively used along with XML when exchanging data between web servers [3] and between the browser and the server.

In this regard, there was an implementation problem of the JSON data processing functions in the SODB.

2 Problem Discussion

The SODB architecture, presented in Fig. 1, explains the principles of processing heterogeneous documents [12]. The basis of the SODB is the hierarchical situation model [6] HSM. HSM specifies a hierarchical set of allowable business process states, conditions for state transitions, actions provided in the current states.

The HSM Interpreter executes the HSM model on external queries, identifying the current states and performing the actions associated with the current states. The sets of current states (CSM – Current State Model) are saved between the interpretation cycles.

There are two types of elements for processing data in the HSM: (1) the doc element specifies a virtual XML document, and (2) the dom element specifies a DOM object (XML Document Object Model) for loading and processing a virtual document. A virtual XML document can be mapped on a real document in the form of a local XML file, a remote XML web service, a ZIP archive containing compressed XML files, a relational database that stores XML documents. Similarly, real JSON document can exist in the form of a local JSON file, remote JSON web service, ZIP-archive with JSON-packed files, relational database, storing JSON-documents.

The conceptual model [7] for processing a virtual XML document is shown in Fig. 2 *a*. Here is a fragment of the HSM model, including the state sta, which contains three child elements: (1) doc is a definition of a virtual document, (2) dom is a virtual document processing object, and (3) wdg is a widget that forms the HTML-code for sending to the user's browser. A real XML document that maps a virtual document is loaded into the DOM object. Here, it is processed, for example, in the form of XSL transformation in accordance with the specified XSLT stylesheet. The result of the processing is used to generate a view in the client's browser using a widget.

Two approaches can be proposed to implement this scheme in the case where a virtual document is mapped to a JSON document: (1) JSON to XML transformation, and (2) processing JSON in special objects.

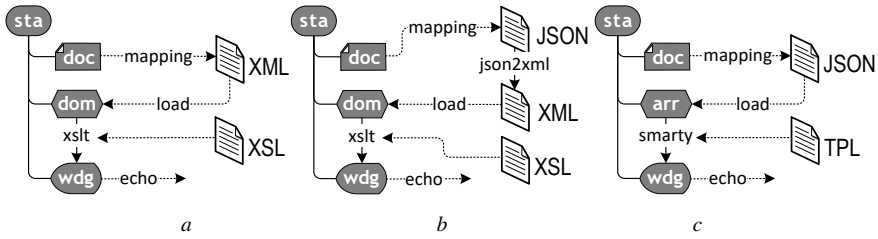


Figure 2
Models of XML / JSON processing in SODB

2.1 JSON to XML Transformation

The model for processing a JSON document with transformation to XML format is shown in Fig. 2 *b*. The conversion is done on the fly before loading the document into a DOM object, further processing is performed just like processing an XML document. The advantage of this approach is due to the large capabilities of XSLT with respect to JSON. Some difficulties originate if the result of processing should be in JSON format. An additional XSL transformation is required to reverse the conversion of XML to JSON in this case.

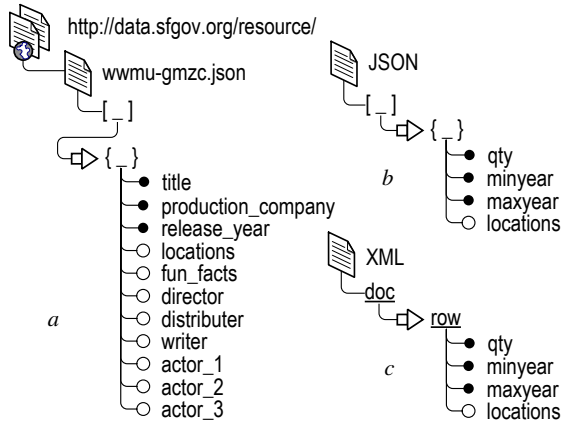
2.2 JSON Processing Objects

The JSON document processing model using JSON processing objects is shown in Fig. 2 *c*. In this case, the JSON document is loaded into a special object for processing. Object arr is an associative array (an abstract data type composed of a collection of 'key, value' pairs), which provides access to particular elements of the JSON document. A suitable template engine, such as Smarty, can be used to handle an associative array and form the widget's code. The capabilities of the template engine are more modest than XSLT, but they are usually sufficient to solve a wide range of transformation tasks.

Both approaches discussed above were implemented in the HSM Interpreter research prototype. Interpreter is written in PHP [3] and is used as an engine for the server part of the web application. The performance of the approaches has been tested on several test cases, two of which are detailed below.

3 PHP-based Examples

Consider two examples of workable HSM models that illustrate the technique of applying both approaches.



d

Qty	Range	Locations
54	1935--2010	
27	1941--2014	Golden Gate Bridge
22	1936--2017	City Hall
19	1938--2015	Fairmont Hotel (950 Mason Street, Nob Hill)
14	1995--2017	Treasure Island
10	1961--2008	Chinatown
10	1936--2015	Coit Tower
10	1956--2000	St. Peter & Paul's Church

Figure 3
Web service JSON data source example

3.1 Used JSON Data Source

Data from the SF OpenData web server, which provides a variety of information about San Francisco, including in JSON format, was used as the JSON test source. Namely, Film Locations service was used (a filming locations listing of movies shot in San Francisco). The structure and composition of the facts about the films and locations provided by the data source are shown in Fig. 3 *a*.

Let us need summary data, not just facts. For example, suppose for each `location` you need to calculate the number of facts (`qty`) and their time range (`minyear, maxyear`) (Fig. 3 *b*). Equivalent representation in XML-format is given on Fig. 3 *c*. Thus, we want to obtain the result in the form presented in Fig. 3 *d*. This problem can be solved in two ways: (1) to download fact-document in the original format Fig. 3 *a* and convert it to the desired format by XSL transformation, and (2) to receive the document in the desired format Fig. 3 *b* from the web service by sending a request to the server. In the tests we use the second way.

The SF OpenData service allows you to formulate GET queries using the SoQL – SQL-like query language. To solve our problem, a query was constructed, which provides grouping by attribute locations, calculation of statistical indicators for groups by means of aggregate functions count, min, max, restriction and sequencing of output results:

```
$select = locations, count (*) as qty,
        min (release_year)   as minyear,
        max (release_year)   as maxyear
$group = locations
$having = qty > 9
$order = qty desc
```

Before sending to the server, the query must be coded in accordance with the rules for setting parameters in HTTP GET requests.

3.2 JSON to XML Transformation Example

In Fig. 4 is an example of the HSM model, which involves processing a JSON document based on JSON to XML transformation. In the state `sta:JSON-Processing` two virtual multi-documents are specified.

Multi-document `doc:SFGov` is provided for downloading JSON-data. It is mapped to web services SF OpenData [8]. The type attribute indicates the JSON-format documents. `Entry`-element `ent:FilmLocs` defines a separate virtual document in the multi-document, mapped to the Film Locations in San Francisco Web service [8]. The `path` attribute specifies a resource, and the get attribute specifies an SoQL query that is coded in accordance with the rules for creating GET parameters.

Multi-document `doc:TmpDoc` is provided to demonstrate the possibility of local saving of downloaded data. It serves to store data in a relational database on a MySQL server. The `action` attribute instructs to establish a connection with MySQL using the default parameters. Element `ent:putDoc` specifies a virtual document that is stored in the cell of table `xmldocs` in the line with the given value of the identifier id in column doc. Element `ent:getDoc` is specified a virtual document extracted from the same cell of the table `xmldocs`.

In the root state a submodel [9] `sub:proc` is envisaged, providing processing of virtual documents. The internal state `sta:proc` is intended to handle under normal conditions and states `sta:loadErr` and `sta:saveErr` – when an error occurs.

In the state `sta:proc` first element `dom:Films` creates and loads a DOM-object. The `srcDoc` attribute refers to the virtual document `ent:FilmLocs` in the multi-document `doc:SFGov`. Therefore, a corresponding JSON document is downloaded from the corresponding web service (see Fig. 3 *b*). This document is converted on the fly into XML format before loading into the DOM object. During the conversion, the JSON array element (enclosed in square brackets) wraps itself in tags `doc`, and the JSON object element (bounded by curly brackets) turns into tags `row` (see Fig. 3 *c*). The `onLoadErr` attribute instructs to jump to the `sta:loadErr` state if errors are detected during the load process. Presence of the attribute `saveDoc` means that after processing the contents of the DOM object must be stored in the corresponding virtual document. Attribute refers to a virtual document `ent:putDoc` of the multi-document `doc:TmpDoc` therefore XML-based representation JSON-document will be saved in the corresponding MySQL table cell. The `onSaveErr` attribute instructs to jump to the `sta:saveErr` state if error is detected during data saving.

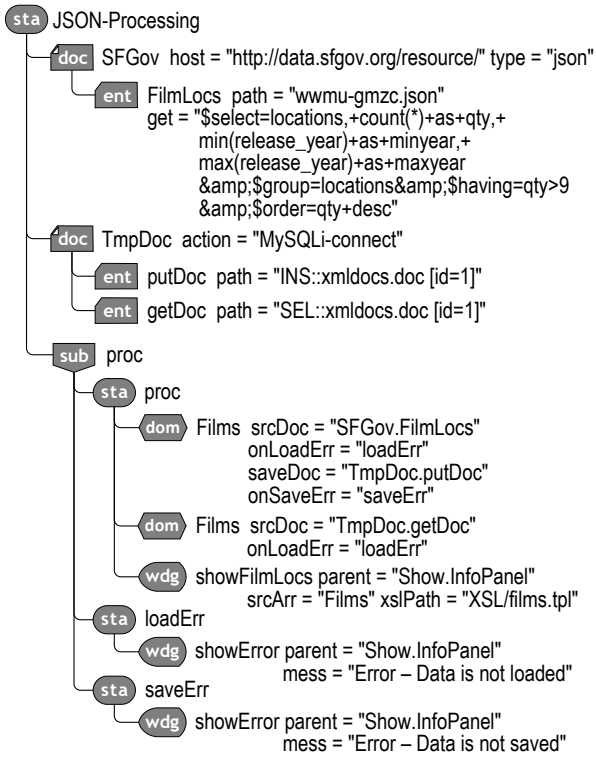


Figure 4
HSM model examples based on JSON to XML transformation

The second `dom:Films` element reloads the DOM object. The `srcDoc` attribute refers to the `ent:getDoc` virtual document in the multi-document `doc:TmpDoc`. Therefore, the XML-document will be loaded from the cell MySQL table that has been saved earlier. (Note that the same content is loaded into the DOM object. This is done only to demonstrate the technique of loading an XML document from the MySQL database.)

The `onLoadErr` attribute prescribes go into a state `sta:loadErr` if errors are found in the load process.

Widget-element `wdg:ShowFilmLocs` generates HTML-code snippet to display the result in the client browser (see Fig. 3 *d*). The attribute `srcDom` refers to `dom:Films`, i.e. the HTML code of the widget is formed based on the content of this DOM object. The attribute `xslPath` prescribes to generate HTML-code by XSL Transformation of XML-content in accordance with the style sheet file `XSL/films.xsl`. The attribute parent points to the parent widget, the HTML code of which should be embedded with the resulting HTML code of this widget, i.e. it is assumed: (1) That in some state that is a parent of the state `sta:proc`, widget-element `wdg:Show` is defined, and (2) that the HTML code generated by the widget `wdg:Show`, contains an identifier attribute `id = "InfoPanel"`.

In the states `sta:loadErr` and `sta:saveErr` widgets `wdg:ShowError` generate an error message when loading and saving data [10]. These messages are inserted into the same point in the HTML of the parent widget, into which the result is inserted in the absence of errors.

3.3 JSON Processing Objects Example

In Fig. 5 is an example of the HSM model, which involves processing a JSON document based on JSON processing objects. Syntactically, this model is very like the previous one (see Fig. 4), it solves the same problem, but in a different way. Definitions of virtual documents `doc:SFGov` and `doc:TmpDoc` have not changed, except that the document `doc:TmpDoc` maps to the table `jsondocs`. The main difference is that `arr`-elements are used instead of the `dom`-elements.

`Arr`-elements are provided in the HSM-model specifically for JSON documents. When interpreted, they generate `Arr` objects – multidimensional arrays designed to load and process JSON documents during interpretation of the HSM model. When loaded, each JSON object is represented as an associative array, and the JSON array is represented as an indexed array.

The first `arr:Films` element creates and loads an `Arr`-object. The attribute `srcDoc` refers to a virtual document `ent:FilmLocs` in multi-document `doc:SFGov`. Therefore, a JSON document is downloaded from the web service and loaded into the `Arr`-object. As before, the `onLoadErr` attribute prescribes

to go to the `sta:loadErr` if errors are found in the boot process. The presence of the attribute `saveDoc` means that after the processing, the contents of the `Arr`-object must be saved in the `ent:putDoc` of the multi-document `doc:TmpDoc`. Therefore, `arr:Films` content is converted back into a JSON document that is stored in the MySQL table cell. The attribute `onSaveErr` prescribes go into a state `sta:saveErr` if there are errors in the save process. The second `arr:Films` element reloads the `Arr` object. The attribute `srcDoc` refers to virtual document `ent:getDoc` in the multi-document `doc:TmpDoc`. Therefore, the JSON document will be extracted from the MySQL table cell in which it was previously saved. The attribute `onLoadErr` prescribes go into a state `sta:loadErr` if errors are found in the load process.

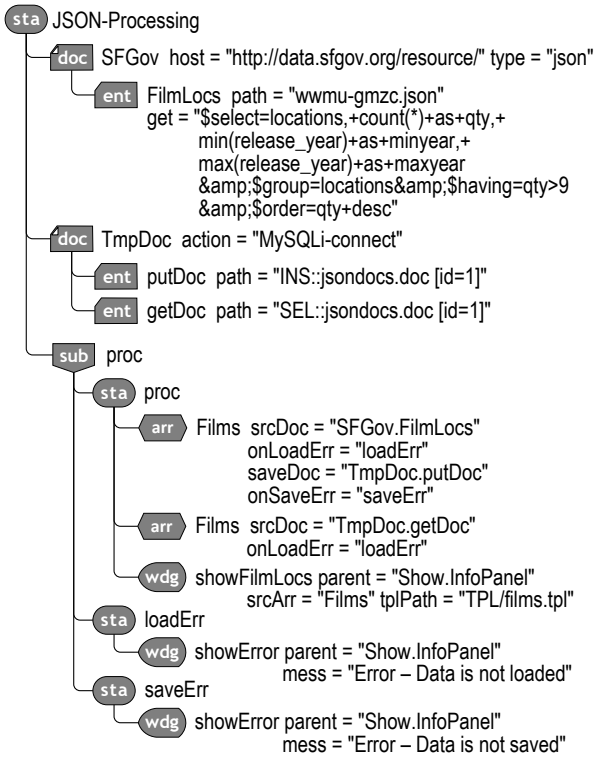


Figure 5
HSM model examples based on JSON processing objects

Widget element `wdg:ShowFilmLocs` generates HTML-code snippet to display the result in the client browser (see Fig. 3 *d*). The attribute `srcArr` refers to `arr:Films`, i.e. HTML-code of the widget is formed on the basis of the content of this `Arr`-object. The `tplPath` attribute instructs to generate the HTML code from the `Arr`-object content using the Smarty template engine per the template

file `TPL/films.tpl`. The attribute `parent` indicates that the resulting HTML code of this widget should be embedded in the parent HTML code of the widget `wdg:Show`.

In states `sta:loadErr` and `sta:saveErr` the `wdg:ShowError` widgets generate an error message when loading and saving data. These messages are inserted into the same point in the HTML of the parent widget, into which the result is inserted in the absence of errors.

4 The Results Discussion

`Two approaches`. Now DBMS can work with JSON-documents in the sense that it can maps virtual documents for real data in JSON-format. When using the first approach, internal processing of JSON documents is performed, as before, in XML format. When using the second approach, internal processing is performed in the format of multidimensional associative / index arrays. If the first approach allows using the power of XML technologies, then the second approach is more familiar for JSON documents.

`High abstraction`. As in the case of XML, JSON specifications are set at a high level of abstraction. Essential aspects of JSON documents are specified in the HSM model declarations, and technical details such as using an algorithm for converting JSON to XML format, creating objects and loading JSON documents into them, etc. are hidden from the developer and executed by the interpreter. For example, the HSM model developer does not need to explicitly specify which of the two approaches to use for processing the JSON document. Based on the context, the interpreter chooses either transformation into an XML-format if the document needs to be loaded into the DOM-object, or conversion to a multidimensional array if the document needs to be loaded into the Arr-object.

`Uniformity`. With both approaches to providing JSON functionality, the structure of the HSM models remains unchanged. The structure is the same as in the case of using XML-documents, the differences appear only in the attributes. Thus, a successful linguistic notation has been found both for specifying virtual documents and for processing them: using a doc-element, a virtual multi-document is defined and its mapping to real data is specified, and using dom/arr-elements, a uniform processing of the virtual document is specified.

`Invariance`. The above invariance principle is observed within the same approach, i.e., changing the address of the JSON document and the physical store will require changes to the specifications of the doc-element and will not require changes to dom/arr elements that reference this doc-element. At the same time, the principle of invariance is violated when moving from the JSON to XML method to the JSON processing objects method, since it will require you to replace the

dom-elements with arr-elements and change the attributes that is, you need to make changes to the document processing specifications. In the future, it is supposed to investigate the possibility of unification so that in the HSM-model one could use some generalized Document Processing Object element, which, depending on the context, can be interpreted as a dom-element or as an arr-element.

`Implementation`. The above functionality of virtual multi-documents is implemented in the research prototype of the HSM interpreter as part of the engine for managing web applications based on DBMS. The prototype of the HSM interpreter, written in PHP language, as well as test HSM-models, other programs, and data of Web applications are located on the webserver http://hsm.ugatu.su. When a client accesses the root script of a web application, the HSM interpreter runs, which executes a cycle of interpretation of the corresponding HSM model. In the process of interpretation, the corresponding CSM-model is corrected and the resulting HTML-code sent to the client. The built-in features and add-ons of the PHP-platform are actively used for access to local files, archives, relational databases [4], remote web servers.

`Effect`. Using the proposed approach reduces the complexity of programming web applications based on DBMS, compared with traditional "manual" programming in the languages of server-side scripts such as PHP. This is achieved due to a higher level of abstraction of the HSM model when the set of routine operations specified in the declarative form when determining virtual multiloquent is assigned to the HSM-model interpreter.

Conclusions

Being an information processor for processing documents from heterogeneous data sources, situationally-oriented databases need the functionality of processing JSON documents. The approaches discussed above make it possible to solve this problem by transforming the JSON document either into an XML format on the fly or into an associative-index array. In the second case, it was required to provide in the situational model special arr-elements, which generate associative-index arrays for loading and processing JSON documents during interpretation. A comparative consideration of both approaches in a practical example of processing the document retrieved from the web service showed a high uniformity of the situational models used for this. JSON-functionality is implemented in the prototype of the HSM interpreter on the PHP platform as part of the control mechanism of web applications based on SODB.

Acknowledgment

This work is supported by RFBR grant 19-07-00682.

References

[1] Janga, P., Davis, K. C.: Mapping heterogeneous XML document collections to relational databases. Lecture Notes In Computer Science, Vol. 8824, 2014, pp. 86-99

[2] Cobo, M. J. et al.: A relational database model for science mapping analysis. Acta Polytechnica Hungarica, 12 (6), 2015, pp. 43-62

[3] Amanatidis, T., Chatzigeorgiou, A.: Studying the evolution of PHP web applications. Information and Software Technology, 72, 2016, pp. 48-67

[4] Varga, V. et al.: Conceptual design of document NoSQL database with formal concept analysis. Acta Polytechnica Hungarica, 13 (2), 2016, pp. 229-248

[5] Fowler M., Sadalage P. J., NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence. Addison-Wesley, 2013

[6] Ameller, D., Burgués, X., Collell, O., et al.: Development of service-oriented architectures using model-driven development: A mapping study. Information and Software Technology, Vol. 62, 2015, pp. 42-66

[7] Kosar, T. et al.: Domain-Specific Languages: A Systematic Mapping Study. Information and Software Technology, 71, 2016, pp. 77-91

[8] Xuanzhe Liu et al.: Data-driven composition for service-oriented situational web applications. IEEE Transactions on Services Computing, Vol. 8, Iss. 1, 2015, pp. 2-16

[9] Agh, H., Ramsin, R.: A pattern-based model-driven approach for situational method engineering. Information and Software Technology, 78, 2016, pp. 95-120

[10] Osvaldo, S. S., Jr. et al.: Developing software systems to Big Data platform based on MapReduce model: An approach based on Model Driven Engineering. Information and Software Technology, 92, 2017, pp. 30-48

[11] Mironov, V. V., Gusarenko, A. S., Yusupova, N. I.: Situation-oriented databases: document management on the base of embedded dynamic model. CEUR Workshop Proceedings (CEUR-WS.org): Selected Papers of the XI International Scientific-Practical Conference Modern Information Technologies and IT-Education (SITITO 2016) Vol. 1761, P. 238-247, Moscow, Russia, 2016, pp. 238-247

[12] Mironov, V. V., Gusarenko, A. S., Yusupova, N. I.: The invariance of the virtual data in the situationally oriented database when displayed on heterogeneous data storages. Vestnik Komp'iuternykh i Informatsionnykh Tekhnologii [Herald of Computer and Information Technologies], No. 1 (151) 2017, pp. 29-36 (In Russian)

Formal Method of Structural-Logical Identification of Functional Model of Subject Area Polycubic Data Matrix

Gennady Kulikov, Vyacheslav Antonov, Almira Fahrullina, Ludmila Rodionova, Ansaf Abdulnagimov, Maria Shilina

Ufa State Aviation Technical University
Faculty of Computer Science and Robotics
K. Marx str, 12, 450008 Ufa, Russia
e-mail: kulikov.gg@net.ugatu.su, antonov.vv@net.ugatu.su,
fahrullina.almira@net.ugatu.su, rodionova.lyudmila@net.ugatu.su,
abdulnagimov.ai@net.ugatu.su, shilina.ma@net.ugatu.su

Abstract: This work proposes a technique of presenting a system model of a subject area in the form of a category of processes. The rule of synthesis of self-organizing structure of process models is proposed. It is shown that the sequence of process decomposition compositions has the property of fractal transformations (self-like). The core of these transformations are commutative triangles that define the structure of interacting processes in composition-decomposition. Such structures have been shown to retain the properties of self-organization in accordance with the laws of the required diversity of the EShBI. The possibility of building a number of identical formal algorithms of control of integrity of the structure of initial processes, in case of violation of which "external" control is required, is shown. The proposed transformation is demonstrated by the example of synthesis of the self-organizing process of the agreement, the architecture of whichб is defined in ISO/IEC 15288 standard "System Engineering." Polycubic data organization for a software system is proposed herein.

Keywords: software system; category theory; subject area; structural analysis and synthesis; formal model; system engineering standards; data processing and transmission; fractal

1 Introduction

At present, automated processing, transfer and use of personnel reserve data using software allows the enterprise not to find itself in a situation of shortage of personnel, upon dismissal or transfer of employees, which significantly increases the level of competitiveness of the organization. In these cases, the tool for

improving the processing, transmission and use of personnel reserve data is intelligent content, which allows to formalize and structure data of enterprise specialists forming this reserve for further multidimensional analysis of data.

Intellectual content in this article refers to a set of information objects, data of human capital of an enterprise, their connections, characteristics, parameters preserving knowledge about the subject area under study.

Human capital is the amount of accumulated knowledge and skills, motivation and energy that can be used, for a certain period of time, to realize the objectives of the production of goods and services [1].

The filling of intelligent content forms data and knowledge of the personnel reserve of the enterprise. In turn, the personnel reserve is a group of data on specialists potentially seeking a leadership position, passed competitive selection and individual training [2].

At the same time, in the process of forming intelligent content there is a problem related to the use of two interrelated business processes of personnel reserve and personnel agency (for example, educational institutions). Considering the information objects of two interrelated business processes (on the example of a personnel reserve and a personnel agency), it is possible to use the standard ISO/IEC 15288 "System Engineering" most common in system engineering, which allows to formalize and structure intelligent content and relations between the objects in question on the example of agreement processes [3].

It is possible to solve the current theoretical and practical problem related to the creation of a fully decentralized platform on which business processes are carried out. Namely, the application of Blockchain technology, which allows to improve business processes, as well as to create a single language of communication. And, as a result, the complex problem of big data analytics, categories of information objects - software modules, ensuring the functioning of real and virtual objects with information processes taking into account their mutual relations. In solving these problems, this study used the results of the works of prominent authors, including D. Gordon [4], J. Zahman [5], K. Lockira [6], in the field of application of topos models and category theory - R. Goldblatt [7], Z. Zhang [8], W. J. Xie [8], Y. I. Yang [8], G. Sun [8], Y.-Q. Gao [8] et al. [9-11].

2 Interaction of Considered Objects of Agreement Processes using Category Theory

One of the fundamental concepts of the methodology of designing information systems is the concept of the life-cycle of its software, which is regulated by a number of normative documents. The main ones are international standards

ISO/IEC 12207 and ISO/IEC 15288. The software lifecycle structure is divided into three process groups: core, support, and organizational processes, each with specific tasks, inputs, solution methods, and results. In the very life-cycle of the software, the results of the subsequent stage may necessitate a change in the solutions of earlier stages, i.e. a recursion component is present. The following categorical approach makes it possible to describe all processes under administrative regulations for absolutely all stages using the Chomsky hierarchy. As a result, we come to structuring the relations of objects, in the form of functional relations. With further allocation of objects inside large objects, we get reduction of unstructured part in them and increase of the general structured part of the subject area according to certain rules, which allows to maximize traceability and identification in automated processes [12] up to the whole chain of relations and a specific performer. The properties of ambivalence of relationships - transactions between virtual objects in the information environment are determined both by the specific objects and by the properties of the environment itself.

Using the formalization proposed in [12, 13] ISO/IEC 15288 "System Engineering" each stage can be represented as a process category. When considering the agreement process, causes and consequences can be used symmetrically.

The state of the object as a whole in the investigated category can be defined in terms of its "mixed" real and virtual states (Figure 1) [14].

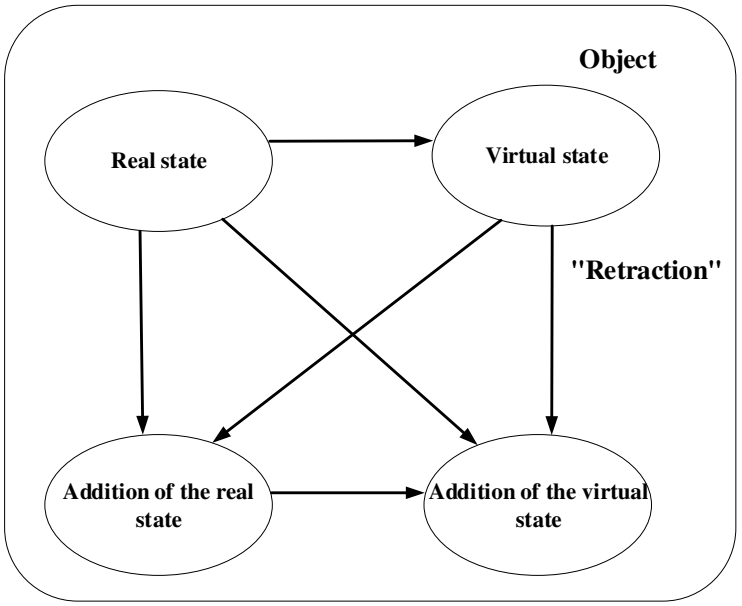


Figure 1
"Mixed" state of the object (similar to cubit)

Applying the rules of the international standard of system engineering ISO/IEC/IEEE 15288, we can formalize the main components, and accounting objects can be represented by different categories of information objects, related relations, with the relations between virtual objects and real objects expressed by the symbolic model of the data warehouse and define the structure [14].

The system goes through certain stages during its life (the life-cycle (LC) model should consist of stages) [15]. Phase - a period within the LC system that refers to the status of the system description or to the system itself. The stages are associated with significant changes in the life of the system, consistent with the passage of milestones in its development throughout the LC. The LC model may include one or more stage models and is assembled as a sequence of stages that may overlap or repeat depending on the application of the system in question, its size, complexity, varying needs and capabilities. Life-cycle stages form the structural basis for detailed modeling of life-cycles of the system using typical processes of its life-cycle. Each stage represents meaningful progress and achievement of planned stages of system development throughout the entire life-cycle and gives rise to critical decisions regarding inputs and outputs. These solutions are used by organizations to address uncertainties and risks directly associated with costs, timelines, and functionality in creating or implementing the system. In this way, the stages provide the organization with a structure of works, in which the management of the enterprise has a high ability to review and monitor the project and technical processes. The relationship between life-cycle processes can be represented as a diagram in Figure 2 [15].

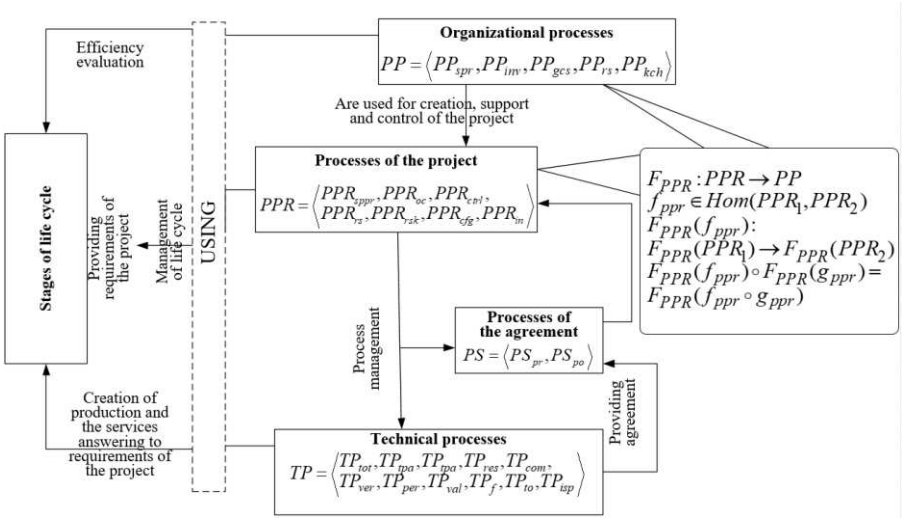


Figure 2

The scheme of communications between processes of life-cycle of system

The formula is presented as an example of the relationship between technical processes:

$$TP =< TP_{tot}, TP_{tpa}, TP_{tpa}, TP_{res}, TP_{com}, TP_{ver}, TP_{per}, TP_{val}, TP_{f}, TP_{to}, TP_{isp} >, \quad (1)$$

$$TP_{tot} = \{TP_{tot}^{1}, TP_{tot}^{2}, TP_{tot}^{3}\}, \ldots, TP_{isp} = \{TP_{isp}^{1}, TP_{isp}^{2}, TP_{isp}^{3}\}$$

Note that formula 1 defines the relationship between technical processes as an ordered set.

Thus, technical processes form, a class of objects, for each pair of objects of which and TP_2 a plurality of morphisms $Hom(TP_1, TP_2)$ are defined, for each pair of which (morphisms), for example $g_{TP} \in Hom(TP_1, TP_2)$ and $f_{TP} \in Hom(TP_2, TP_3)$ their composition $g_{TP} \circ f_{TP} \in Hom(TP_1, TP_3)$ is defined. Technical processes form a category of sets [16].

The relationship between the life-cycle processes of the system can be represented as a diagram in Figure 1, where:

– organizational processes – $PP = \langle PP_{spr}, PP_{inv}, PP_{gcs}, PP_{rs}, PP_{kch} \rangle$; (2)

– project processes – $PPR = \langle PPR_{sppr}, PPR_{oc}, PPR_{ctrl}, PPR_{rs}, PPR_{rsk}, PPR_{cfg}, PPR_{in} \rangle$; (3)

– technical processes $TP =< TP_{tot}, TP_{tpa}, TP_{tpa}, TP_{res}, TP_{com}, TP_{ver}, TP_{per}, TP_{val}, TP_{f}, TP_{to}, TP_{isp} >$; (4)

– agreement processes – $PS = \langle PS_{pr}, PS_{po} \rangle$. (5)

From the studies carried out [15, 17, 18] we can conclude that there is a display of categories of sets, preserving the structure of these categories-sets. I.e. the relationship between life-cycle processes, in light of the provisions of the system engineering methodology, is described by functionaries that match an object of one category with an object of another category.

Let some object be highlighted and designated by the term. As a rule, a new object is compared to already known objects, and its information model is formed as a set of comparison of information models of previously known objects. At the same time the model of the new subject area for this object will be built on the basis of the subject area of the object that became known first. As a result, knowledge of the subject area including the object will be structured as a set of properties of the first selected object and a sequence of changes of subject areas of subsequent objects. It turns out that previously selected objects are more privileged with respect to subsequent objects of the subject area, because models of these objects are constructed by changing the model of already known objects. At the same time, the sequence of selecting objects is the will of the case. When you try to select the object that is best suited as the first, you find it most

convenient to use as its image some idealized, average object of that subject area, the model of which is replaced by a set of variables describing the objects of the subject area. Selecting a set of variables to describe subject area objects and selecting valid values for these variables is largely arbitrary. However, it is this choice that will further determine the limits of the applicability of its model [15, 19].

Thus, the agreement process consists of the acquisition process and the delivery process. At the same time the process is symmetrical, the acquiring side can act both sides at the same time. Using the informational approach, to formalize the subject area, are the categories of objects and the relationships between them. That is, the category can be represented by a plurality of subject area objects and a plurality of relationships there between [12].

Consider the basic positions of category theory [20, 21] to simulate this process. We will use the following symbols: agreement processes PS, acquisition processes $PSpr$, delivery processes $PSpo$ (Figure 3).

Let's highlight PS the acquisition $PSpr$ and delivery agreement processes $PSpo$ into separate categories and subcategories, which consist of process objectives, process activities, and result [12]. Consider the interaction of subcategories within which information objects result in a communicative triangle by category and individual subcategories.

According to ISO/IEC 15288 "System Engineering," the agreement process consists of acquisition processes and delivery processes [22, 23].

This category PS forms a class of objects with a given relation for any pairs $(PS, PSpr_{pr})$ and $(PS, PSpo_{pr})$ whose relationships are defined by a set of morphisms: f_1, f_2, p_1, p_2.

By treating agreement processes PS as interrelated acquisition $PSpr$ and supply processes $PSpo$, the interaction process can be represented by a Cartesian work $g{:}PS \to PS_{pr} \otimes PS_{po}$, where g is the only morphism expressed as a result of the agreement process. Note that these operations comply with the rules of formal logic.

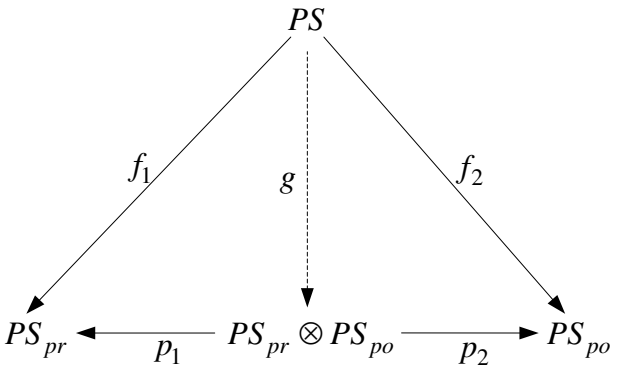


Figure 3
Simplified diagram of the commutative triangle of the agreement process

Thus, the agreement process is formalized as morphisms, and the result of interaction of information objects can be represented by a Cartesian $PS_{pr} \otimes PS_{po}$ product with morphisms $p_1{:}PS_{pr} \otimes PS_{po} \to PS'_{pr}$ and $p_2{:}PS_{pr} \otimes PS_{po} \to \in PS'_{po}$ such that for any object with PS morphisms $f_1 : PS \to PS_{pr}$ and $f_2 : PS \to PS_{po}$ there exists a single morphism $g{:}PS \to PS_{pr} \otimes PS_{po}$ in which a diagram in the form of a commutative triangle is highlighted, where:

$\otimes$ – Cartesian product

g – is the morphism resulting from the process,

Applying the logical convolution of the semantic model of the subject area under study [22], which also contains contradictions relations, we come to their formalization in the form of logical Descartes squares.

Using the research of N. A. Vasilyev [14] of the transition from a logical square to a logical triangle, which reflects the Associative Law of Composition between three vertices "all pairs of judgements-opposites cannot be simultaneously true, but can simultaneously be false," [8], we come to the expediency of applying the Law of Necessary Ashby Variety for two vertices. For example, two of any vertices must have a relationship on the principle of their superposition (nesting, while maintaining and as holistic objects). But simultaneously, relationships between objects can also be "private" (separable), when one object is part of another, then integrity retains one object and the other object is a complement to the first or vice versa, with the superposition property stored for only one object. Thus, to combine associative relations and "private" relations it is possible to use the principle of decomposition of logic triangles, selection of "virtual" objects - internal for the considered "large" triangle. It is obvious that additions will appear as new associative (relative to the sides of the triangle) and private as additions.

These are Rombs (essentially the logical squares of Descartes). This separation logic is well seen on the semantic triangle. To eliminate such potential contradictions, the logical structure of the subject area model should be complemented by cybernetic squares for self-organization. Thus, it is possible to obtain a structure of an open, consistent, self-organized controlled meta-module of an AIC subject area. In the application aspect, this can be realized by a two-step procedure: 1) finding conflicting squares; 2) elimination of detected contradictions by building metamodel with additional squares to cybernetic control system with feedback.

Consider the following separate subcategory: the acquisition process PS_{pr}, which consists of three components (Figure 4):

1) Acquisition process purposes $PS^1_{pr} = \{pr^1_1\}$.

2) Activities in the acquisition process $PS^2_{pr} = \left\{pr^2_1,...,pr^2_8\right\}$.

3) Result of process of acquisition $PS^3_{pr} = \left\{pr^3_1,...,pr^3_7\right\}$.

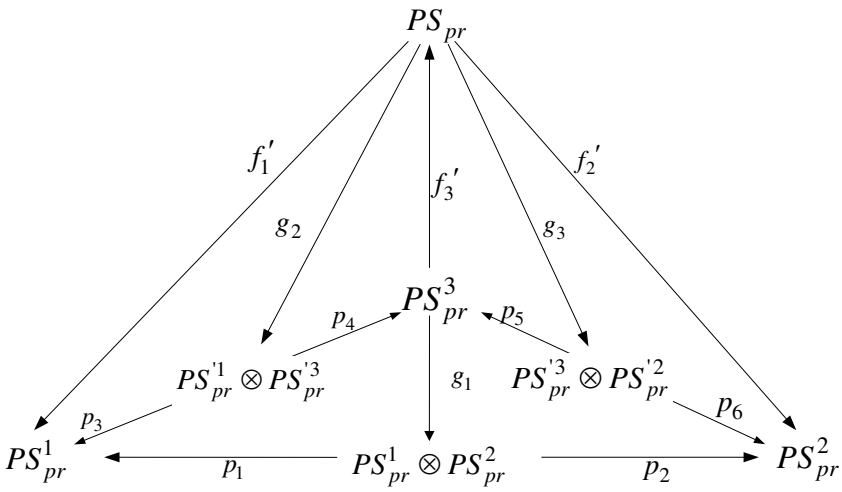


Figure 4
Purchase process commutative triangle

The interaction of information objects PS^1_{pr} and PS^2_{pr}, can be represented by a Cartesian work $PS^1_{pr} \otimes PS^2_{pr}$ – with morphisms $p_1{:}PS^1_{pr} \otimes PS^2_{pr} \to PS'^1_{pr}$ and $p_2{:}PS^1_{pr} \otimes PS^2_{pr} \to PS'^2_{pr}$ such that for any information object PS_{pr} with

morphisms $f_1^{'}:PS_{pr} \to PS_{pr}^1$ and $f_2^{'}:PS_{pr} \to PS_{pr}^2$, there is a single morphism $g_1:PS_{pr} \to PS_{pr}^1 \otimes PS_{pr}^2$ which is defined by a commutative diagram.

Similarly, the interaction of objects with morphisms is considered, $p_3:PS_{pr}^{'1} \otimes PS_{pr}^{'3} \to PS_{pr}^{''1}$ and $p_4:PS_{pr}^{'1} \otimes PS_{pr}^{'3} \to PS_{pr}^{''3}$ in which for any object PS_{pr} with morphisms $f_1^{'}:PS_{pr} \to PS_{pr}^1$ and $f_3^{'}:PS_{pr} \to PS_{pr}^3$, there is a single morphism $g_2:PS_{pr} \to PS_{pr}^{'1} \otimes PS_{pr}^{'3}$ that is defined by a commutative diagram.

In the same way, the interaction of objects $PS_{pr}^{'3} \otimes PS_{pr}^{'2}$ with morphisms $p_5:PS_{pr}^{'3} \otimes PS_{pr}^{'2} \to PS_{pr}^{''3}$ and $p_6:PS_{pr}^{'3} \otimes PS_{pr}^{'2} \to PS_{pr}^{''2}$ such that for any object PS_{pr} with morphisms $f_2^{'}:PS_{pr} \to PS_{pr}^2$ and $f_3^{'}:PS_{pr} \to PS_{pr}^3$, there is a single morphism $g_3:PS_{pr} \to PS_{pr}^{'3} \otimes PS_{pr}^{'2}$ that is defined by a commutative diagram.

As a result, the acquisition process can be represented as a Cartesian product of the following projections:

$$PS_{pr} = (PS_{pr}^1)^{'} \otimes (PS_{pr}^2)^{'} \otimes (PS_{pr}^{'1})^{''} \otimes (PS_{pr}^{'3})^{'} \otimes (PS_{pr}^{'3})^{''} \otimes (PS_{pr}^{'2})^{''} \tag{6}$$

Similarly, the interaction of delivery process objects is discussed. The result can be represented as a Cartesian product, and if the result is an independent quantity, it can be intermediate. This allows you to identify the dependence of a recursive point on one another. In other words, if the goal changes, you can predict the result, and information objects are represented as categories, and the ratio of the transition from one stage to another is represented as morphisms.

The pattern of forming a multiple Descartes square for the subject area under study and their decomposition in terms of formal logic and category theory is shown in Figure 5.

When considering agreement processes where the acquisition process PS_{pr}, the delivery process PS_{po}, each presented as categories, object interactions can be represented by a set of Cartesian works:

$$\begin{aligned} PS_{pr} \otimes PS_{po} = \langle & PS_{pr}^1 \otimes PS_{pr}^2, PS_{pr}^{'1} \otimes PS_{pr}^{'3}, PS_{pr}^{'3} \otimes PS_{pr}^{'2}, PS_{po}^1 \otimes PS_{po}^2, \\ & PS_{po}^{'1} \otimes PS_{po}^{'3}, PS_{po}^{'3} \otimes PS_{po}^{'2} \rangle. \end{aligned} \tag{7}$$

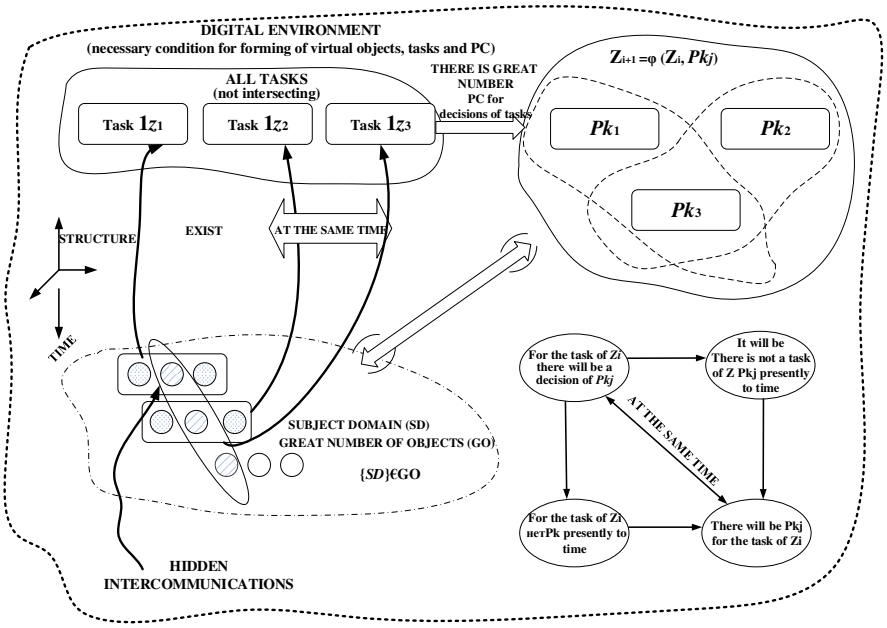


Figure 5
A scenario diagram for forming a multiple Descartes square and decomposing them

Obviously, the agreement process in this case can be represented by the formula:

$$PS = \langle (PS_{pr} \otimes PS_{po}), (PS_{po} \otimes PS_{pr}) \rangle . \tag{8}$$

Similarly, the acquisition process PS_{pr} can be represented by the formula:

$$PS_{pr} = \{\langle (PS^1_{pr} \otimes PS^2_{pr}), (PS^2_{pr} \otimes PS^1_{pr}) \rangle, \langle (PS^1_{pr} \otimes PS^3_{pr}), (PS^3_{pr} \otimes PS^1_{pr}) \rangle, \\ \langle (PS^3_{pr} \otimes PS^2_{pr}), (PS^2_{pr} \otimes PS^3_{pr}) \rangle\}, \tag{9}$$

And the delivery process PS_{po} with the formula:

$$PS_{po} = \{\langle (PS^1_{po} \otimes PS^2_{po}), (PS^2_{po} \otimes PS^1_{po}) \rangle, \langle (PS^1_{po} \otimes PS^3_{po}), (PS^3_{po} \otimes PS^1_{po}) \rangle, \\ \langle (PS^3_{po} \otimes PS^2_{po}), (PS^2_{po} \otimes PS^3_{po}) \rangle\}. \tag{10}$$

Thus, the interaction between different categories can be described as a functionary based on the Cartesian works presented above.

In a category PS_{pr}, any object will appear in the category objects PS_{po} that follow: $f \in \mathrm{Hom}(PS_{pr}, PS_{po})$ $\{f \in \mathrm{Hom}(PS^1_{pr}, PS^1_{po}), f \in \mathrm{Hom}(PS^2_{pr}, PS^2_{po}).$ $f \in \mathrm{Hom}(PS^3_{pr}, PS^3_{po})\}.$

It also describes how to display objects in a category PS_{po}:

$g \in \mathrm{Hom}(PS_{pr}, PS_{po})$ $\{g \in \mathrm{Hom}(PS^1_{pr}, PS^1_{po}),$

$g \in \mathrm{Hom}(PS^2_{pr}, PS^2_{po}),$

$g \in \mathrm{Hom}(PS^3_{pr}, PS^3_{po})\}.$

Similarly, the "Cartesian product" is described, because we will consider it as a new object:

$f \in \mathrm{Hom}(PS_{pr} \otimes PS_{po}, PS_{po}),$

$g \in \mathrm{Hom}(PS_{po}, PS_{po} \otimes PS_{pro}),$

$f \circ g \in \mathrm{Hom}(PS_{pr} \otimes PS_{po}, PS_{po} \otimes PS_{pro}).$

Thus, we get that objects in categories are determined by interaction of functionators. Here we can use the Descartes square, where the composition of the above two processes g, indicated, is a direct mapping of the given agreement process to the result. The result of the process can be considered as a composition of the product of goals and activities. It turns out to be another triangle. At the same time, the goal cannot be achieved in one way only. Applying the principle of the necessary diversity of the Ashby [7], according to which, there cannot be the same relations between objects, there must be additional ways to achieve results, we get additional squares in the diagram. In other words, the Descartes square will be determined by the Ashby principle of necessary diversity [2]. These charts satisfy the provisions of the topos, which allows to infer the dependence of, for example, this point on another, by changing the result, for this purpose we determine what needs to be changed for the purposes that in the activity, or we do not change the goals, but only the activity.

It is obtained from the parameters described above, the process can be considered in the form of a combined process and for it the obtained conclusions are fair. That is, the agreement process can be represented by a hierarchical set of nested processes. the dependence of $y=f$ is fair $(x, y) => y = f(x, f(x, y))$, the self-similarity of large-scale invariancy turns out (fractals). In other words, the software system has many variants of development, and the state of the system is determined by the place in which it is currently located.

The presented rule of synthesis of self-organizing structure of agreement process is based on decomposition - composition of objects-processes by logical addition of fractals in the form of commutative triangles into squares of Descartes and further into Cartesian polyhedron.

When designing a software complex, we will divide it into categories of objects from which it will consist. Given that fractal relations can be established between them. We can understand that once established relation between categories in the form of a functionary, will determine functionary relations. We get many functionaries, which we highlight as a separate category. For instance, relations between fractal elements are built at the level of categorical relationship [3, 18].

Thus, sub-objects constituting a fractal can be described as functionary relationships, i.e. elements of different categories belong to each other via a functionator. As a result, the categories are related to each other by a self-similarity relation for a particular subject area.

Fractal in this case acts no geometric place of points, but as a categorical fractal. We will understand the certain categorical object connected by some similarity as a fractal, the similarity is connected by versatility, multi-hierarchy, self-similarity. In terms of category theory, fractal properties will satisfy certain conditions. In this case, each object is a software module, the lower the chain, the greater the number of relationships (Figure 6).

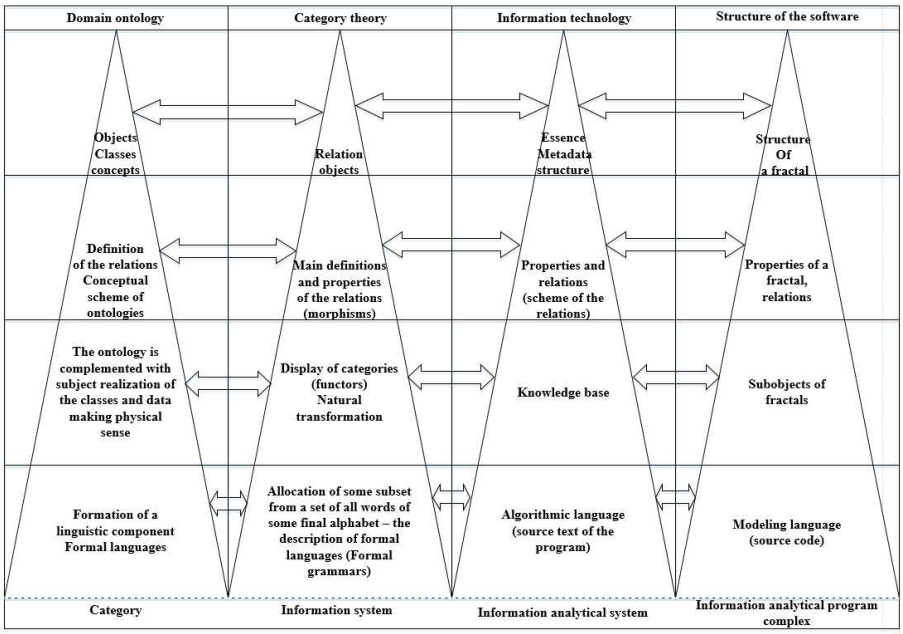


Figure 6

Diagram of relationships between different object categories

The software is built modulo; this is called a communication vector. We propose the following method: the method of building relationships for further software development and design, i.e. relationships between modules [7, 15].

3 Polycubic Model of Data

Based on the identified patterns and for ease of presentation of information, it is advisable to present it in the form of multidimensional cubes, the dimensions of which will be the values of attributes of subject areas. Given the constant expansion of the data set, together with the preservation of control, we come to the possibility of building a complex of related registers (in time), including a lot of data received from the outside, i.e. the constant appearance of new virtual objects, which at the next iteration of the system construction become internal. We can talk about the digital twin model control subsystem. The principle of multidimensional models is based on the OLAP concept. OLAP technologies are focused on complex analytical tasks. Figure 7 shows a polycubic olap model for a complex analytical task built by the user. The OLAP model consists of data sources, DataStore, OLAP cubes, and result presentation [24, 25, 28]. The characteristics of the OLAP model include: the power, dimension, volume of the hypercube, where the power of the hypercube is the product of the power of its measurements, and the volume of the hypercube is the product of the power of the cube and the number of indicators.

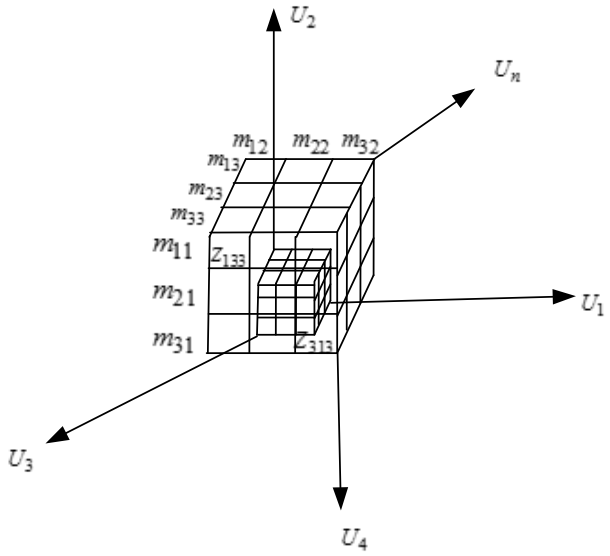


Figure 7
Building a Cube Based on Polycubic Data Organization

$U = \left\{U_1, U_2, U_3, U_4, ..., U_n\right\}$ - set of measurements of a polycube;

$M_{U_i} = \left\{m_{1i}, m_{2i}, m_{3i}, ..., m_{si}\right\} - i = 1, ..., n$ - set of tags of measurement U_i ;

$M = M_{u_1} \cup M_{u_1} \cup ... \cup M_{u_n}$ - set of tags of a polycube;

$Z = \left\{Z_{123}, Z_{223}, ..., Z_{mui}\right\}$ - a plurality of cells with measurements of a polycube.

In polycubic data organization, a multidimensional database is represented by several hypercubes, with different dimensions and dimensions. Each cell is a category, the relationship between them is functional. As a result, we can use different intersection points for different needs. For example, the time vector can be applied to call a module of this year or a previous one. This ensures traceability and repeatability of the output results. When considering the vector "competences," it is possible to pick, regardless of training, any new set of competences.

Using a polycubic data model will allow to define hypercubes with different dimensions and with different measurements as their faces [26].

Let's consider the hypercube construction option, for which we will produce the necessary conversion of the data stored in the database tables. For example, in order to improve performance when building a hypercube, we find unique elements stored in columns that will be measurements of the hypercube. Records that have the same dimension values are pre-aggregated. To create hypercube slices, you must define the coordinates (actual measurement values) for the table entries and define the entries. Actions are divided into two stages - according to the given classification of attributes. For identification and functional attributes, it is possible to create a table of unique values on the basis of which the measurement will be built. Depending on the type of business processes to be integrated, we define an integration rule. For formalized attributes, based on the set semantic dependencies between accounting object types, define an integration rule.

The transformations are schematically shown in Figure 8. We define keywords (phrases), draw up a table of the accuracy of variants of their combinations satisfying the semantic rule (by the function of belonging) and as in the case of formalized attributes we build a table of unique values. Instead of a single table, we get the equivalent of a normalized database. Define the coordinates for the measurement values by renumbered the element values. Lists of measurement values are pre-sorted alphabetically within a single dimension [15].

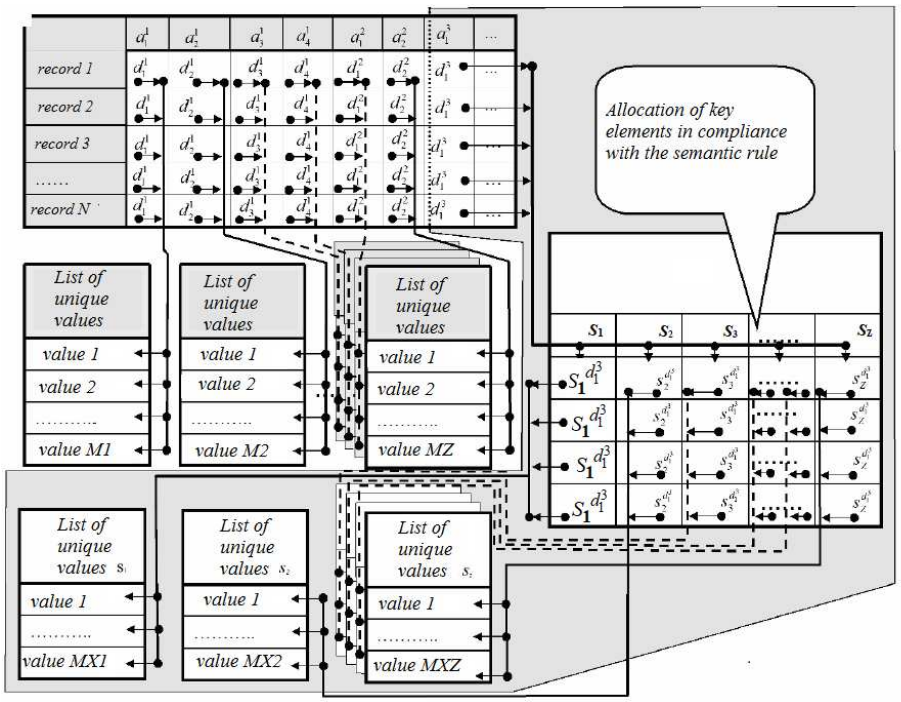


Figure 8
Example of formation Olap-cub

4 Information Support of the Subject Area under Consideration based on the Data Warehouse

Using theoretical-multiple models of agreement processes, an information and analytical system (IAS) scheme based on data repositories was built (Figure 5).

IAS is characterized by the following parameters: use of various database management systems (Oracle, MySQL, etc.), modern storage tools (Data Warehouse), application of Online Transaction Processing (OLTP), On-line Analytical Processing (OLAP), Data Mining, etc.

As a result of IAS use, users (decision-makers, analysts, experts) receive an effective tool for selection, analysis of candidates for personnel reserve, as well as generation of data on personnel reserve, presented in tabular and graphical form. This IAS provides centralized collection, storage, updating, structuring, systematization, integration, processing and consolidated analysis of diverse data of candidates of personnel reserve [16, 17].

In IAS subsystems comply with the following principles:

– Razvivayemost;

– Construction blochnost;

– Certain independence of individual subsystems and their databases (DB);

– Correction of functional part, etc.

The main data for the DB is the candidate data directory, which has full information about the candidate, the name of the table in which the key figures are stored, the key figure group code to which the key figure belongs, the key figure to the sub-indicators, the key figure unit code, the frequency of the key figure presentation and the field with the comment for the key figure. The knowledge base contains rules of the algorithm of the program operation, described by means of categories of sets.

External sources of data are: state authorities (Ministries of Labor and Social Protection, local administration of the city and district, etc.), cooperating organizations (employment centers, universities, enterprises, etc.), higher authorities (holdings, corporations, etc.).

The use of a formal description of information objects in the language of category theory helps to increase the formality of the subject area, opening up new opportunities to identify problems [26, 27].

When forming storage of data, the separate objects expressed by categories of users, categories of program modules can be considered, the relations between them will be defined by functors. This allows data and knowledge to be stored in a single place. Thus, the problem of data and knowledge separation in processing and transmission is solved.

Conclusions

Invention proposes a rule for synthesis of a self-organizing structure, of an agreement process, based on sequential decomposition - composition of process objects in the form of commutative triangles into Descartes squares and further into Cartesian polyhedron. We can say that for all elementary structures the rules of formal logic, parametric control of the state of processes are preserved. The structure of the agreement process itself is synthesized by the scheme above. This makes it possible to apply decentralization-based technologies - the use of blockchain, and eventually the development of a fully decentralized platform on which business processes are carried out. Taking into account the above-mentioned studies, built on a strict mathematical basis, the prerequisites for creating a single communication language are created. Moreover, all this is represented as a system of equations derived from a formal software model. This method of transition to a quantitative mono-variant description of information processes and the application of the requirements of ISO / IEC 15288 standard,

allows one to link each stage of the life-cycle of the developed information analytical system by introducing recursion, recursively searching for the optimal solution of subtasks, performing a similar sequence of actions. The Olap cube was built on the basis of polycubic data organization.

The Application of the method of semantic differentials of C. Osgood, for processing and presentation of information, in relation to stages of the process under consideration, provides the possibility for building the required functions of belonging.

Acknowledgement

The study is conducted with financial support from the Ministry of Education and Science of the Russian Federation as part of the basic part of the state assignment to higher education educational institutions # FEUE-2020-0007.

References

[1] Filatova E. V.: Investments in human capital at small businesses [Text]: thesis abstract of Cand. Econ. Sciences, M: 2010, 177 p.

[2] Bazarova T. Yu., Eremina B. L.: Personnel management. Textbook for high schools / Ed. T. Yu Bazarov, B. L. Eremina. - 2nd ed, changed and added. M: UNITY, 2007, 560 p.

[3] Kovács, G. L., Petunin, A. A.: An information technology view of manufacturing automation, Product life-cycle management, Pollack Periodica, Vol. 11, No. 2, 2016, pp. 3-14

[4] Zachman, J. A.: A Framework for Information System Architecture, IBM System Journal, Vol. 26, No. 3, 1987, 276 p.

[5] Lokir K., Gordon O.: Project Management: Stages of Excellence, Publishing House, 2008, 352 p.

[6] D. Gianni, A. D'Ambrogio, A. Tolk (eds.): Modeling and simulation-based systems engineering handbook, CRC Press, London, 2014, 513 p.

[7] Goldblatt R. Toposa. Category analysis of logic. Moscow: Mir, 1983, 486 p.

[8] Z. Zhang, W.-J. Xie, Y. I. Yang, G. Sun, Y.-Q. Gao: Simulation studies of the self-assembly of halogen-bonded sierpiński triangle fractals, Wuli Huaxue Xuebao. Acta Physico, Chimica Sinica, Vol. 30, No. 3, 2017, pp. 539-547

[9] Mabrok M. A., Ryan M. J.: "Category theory as a formal mathematical foundation for model-based systems engineering", Appl. Math. Inform. Sci., Vol. 11, No 1, 2017, pp. 43-51

[10] Uhlemann T. H.-J., Steinhilper C. L. R., Steinhilper R.: The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0. Procedia

CIRP, Part of special issue: The 24th CIRP Conference on Life-cycle Engineering. Ed. by S. Takata, Y. Umeda, S. Kondoh, Vol. 61, 2017, pp. 335-340, DOI: 10.1016/j.procir.2016.11.152

[11] Baumester D., Jekert A., Cajlinger A.: Fizika Kvantovoj Informacii. M.: Postmarket, 2002, 376 p.

[12] G. G. Kulikov, V. V. Antonov, D. V. Antonov, F. F. Shingarev: Method of subject-oriented classification and system modeling of weakly formalized information flows in industrial automation systems, Bulletin of SUSU. A series of "Computer technologies, management, radio electronics", Vol. 16, No. 2, 2016. pp. 116-130, DOI: 10.14529 / ctcr160212

[13] Rechkalov AV, Kulikov GG, Antonov VV, Artyukhov AV.: Development of a formal model of the production process for the organization of project and production management with the use of intellectual KIC, Bulletin of PNIPU. Electrical engineering, information technology, control systems, No 11, 2014, pp. 34-54

[14] Vasilyev N. A.: Imaginary logic. Selected works, M: Nauka, 1989, 264 p.

[15] Cordon, O., Herrera F.: General study on genetic fuzzy systems, Genetic Algorithms in engineering and computer science, 1995, pp. 33-57

[16] Artemyev, S. A., Zamay S. S., Pitenko A. A.: GIS the designer with means of the analysis of the information and analytical systems of, Computing technologies, Messenger TREASURY, Vol. 9, No. 3(42), 2004, pp. 188-192

[17] Massel L. V. Kovalyov: Fractal approach to structuring knowledge and examples of its application, design Ontology, Vol. 6, No. 2 (20), 2016, 149 p.

[18] Kovalev S. P.: Teoretiko-category approach to design of programm systems, base. and approx. maty., Vol. 19, No 3, 2014, pp. 111-170

[19] E. E. Akimkina, A. E. Abbasov.: Analysis of the tools of information systems for processing multidimensional data, Information and Technology Bulletin, 2015, pp. 61-75

[20] Kovalyov S. P.: Category of computing systems, International conference "Algebra and Logic: Theory and Applications". Theses of reports. Krasnoyarsk: SFU, 2013, pp. 64-66

[21] Kulikov, G. G. Antonov V. V., Antonov D. V.: Theoretical and application-oriented aspects of creation of models of information systems of, LAP LAMBERT Academic Publishing GmbH & Co.KG, Germany, 2011, 134 p.

[22] G. G. Kulikov, V. V. Antonov, M. A. Shilina, A. R. Fahrullin: Mathematical and software for the construction and implementation of object-oriented IUS from the conditions of identification and traceability, Journal of the UURSU. Series "Computer Technologies, Control, Radio

Electronics.", Vol. 16, No. 3, 2016, pp. 143-151, DOI: 10.14529/ctcr160316

[23] ISO/IEC/IEEE 15288:2015 System and software engineering – System life-cycle processes

[24] Kovács, G., Yussupova, N., Rizvanov, D.: Resource management simulation using multi-agent approach and semantic constraints, Pollack Periodica, An International Journal for Engineering and Information Sciences, Vol. 12, 2017, pp. 45-58 https://doi.org/10.1556/606.2017.12.1.4

[25] Abdulnagimov A., Ageev G.: Neural Network Technologies in Half-Time Modeling: Principles of Implementation of Digital Dual GTS, Bulletin USATU, Ufa, Vol. 23, No. 4 (86), 2019, pp. 115-121

[26] Kulikov, G. G., Antonov V. V., Savina A. A.: Information and Analytical System Based on Knowledge of Population Registration, UGATU: Scientific journal of Ufa state aviation technical University, Vol. 10, No. 2(27), 2008, pp. 60-67

[27] Antonov V. V. Kulikov G. G., Antonov D. V.: Formalization of subject domain with use of the tools supporting standards, UGATU: Scientific journal of Ufa state aviation technical University, Vol. 16, No. 3 (48), 2012, pp. 42-52

[28] Yussupova, Nafissa & Kovacs, George & Boyko, Maxim & Bogdanova, Diana: Models and Methods for Quality Management Based on Artificial Intelligence Applications. Acta Polytechnica Hungarica, No. 3, 2016, pp. 45-60

Identification of Critical Objects in Reliance on Cyber Threats in the Energy Sector

Aleksei Massel and Daria Gaskova

Information Technologies in the Energy Sector Melentiev Energy Systems Institute of SB RAS Irkutsk, Russia
amassel@isem.irk.ru, gaskovada@isem.irk.ru

Abstract: The article describes the identification of critical facilities being a significant trend in researching critical infrastructures, particularly in the energy sector. Cyber threats are believed to be important contemporary threats to energy security in Russia. The authors formulated the risk-based approach to support decision-making in the identification of critical facilities in the energy sector in terms of cyber threats. The novelty of the study lies in the development and application of a fractal stratified model of the relationship of risks, energy facilities, and technologies. This author's approach allows one to identify all the hidden relationships of the influence of cyberthreats on energy objects.The results of the proposed approaches are used and supported by an intelligent decision support system.

Keywords: intelligent system; risk management; cyber threats; risk-based approach; critical infrastructures; energy facilities

1 Introduction

Any military actions are primarily targeted at destructing civil infrastructure which might cause the damage comparable with the blows hitting military forces.

Scientific research and applied developments on complex problems on safeguarding of population, infrastructure units, and life activity environment were described by A. Kondratjev, N.A. Makhutov and S. A. Timashev.

In fact, the problems of ensuring the safety of population and technosphere objects are multidisciplinary [1], and their solution assumes the acquisition of new data on all components through fundamental research carried out in proper areas [2].

Critical infrastructures include complex technological systems, which are believed to be unique technogenous objects. Modeling and investigation of the safety of such systems are reported in the paper by [1]. As a rule, the study of the security of such systems is intended at developing the methods on system analyses,

representation and processing of information on the objects. Such information reflects knowledge of experts, collected data and knowledge, mathematic models and others by applying methodologies of investigations, as well as the proposed information environment.

Modeling is executed via recognizing alternative solutions for reducing risk of reliability and safety violations. To add, because such systems bear risks for population and environment, it is vital to consider malfunction of the systems, which might result in the origination of extreme situations and cascade accidents.

The conception of risks is based on the identification of the current status of the elements of the system and conditions of origination and development of extreme, accident and disastrous situations qualitative and quantitative description of the scenarios and consequences of achieving limit conditions causing accidents and disasters [2].

1.1 Management of Risks

The design of critical information infrastructure is aimed at ensuring reliability and safety, and the sources of risk are represented by catastrophic failures and equipment errors, natural impacts on the facility, or deliberate actions of operators. An important component of the information and technological process of modern production are information flows, making IT risk management a priority area of risk management.

IT risk management integrates technologies used to identify, analyze evaluate the incidents and threats, as well as implement security enhancements [3].

Four major components of risk management, indicated in the literature, are [4]:

- risk identification,
- risk analysis,
- risk-reducing measures,
- risk monitoring.

For the purposes of information security risk assessment, the qualitative risk analysis is usually carried out in the Information System [4]. Qualitative methods are provided as descriptions and recommendations; these approaches include scenario analyses, surveys, and audits.

IT risk management of corporate information technologies plays an important role in many aspects of the modern organization functioning, and the key task of such management is risk analysis [3], [4], [5]. This task becomes especially urgent in the context of energy security (ES), under conditions of preparing infrastructure projects both for reorganization and expansion of energy network to meet the

national electricity needs [6] and low-carbon energy transition projects. In most cases, they are based on renewable energy sources [7] requiring the introduction of new technologies. This problem is aggravated by the underestimation of the safety of introduced new technologies represented by intelligent information systems, technical equipment, and devices. Thus, the current tendency is to enter service-oriented and cloud-based corporate technologies in industrial production [8], [9]. Control systems based on these technologies include online data analysis and processing, real-time monitoring of devices; they are capable to be flexible and integrated. However, the absence of adequate management of cybersecurity risks can have serious consequences not only for the enterprise itself but also for the environment within a city, region or nationwide.

In this paper, cyberthreats are considered in the context of a strategic analysis of critical infrastructures. The introduction of new information technologies often carries significant risks and uncertainties, intangible benefits, but provides attractive long-term financial benefits, and can be considered in the framework of strategic project management [10]. Management and control, and possible ways of adapting, if necessary, are provided at the monitoring stage by the project and are favorable to identify potential risks and their adverse effects, as well as developing the ways to eliminate or minimize them.

1.2 Critical Infrastructures

Critical infrastructure is part of civil infrastructure, which makes up a combination of physical or virtual systems and means that are important for the country, as their malfunction or destruction can trigger disastrous consequences in the fields of defense, economy, health safety and nation security [11].

Awareness of the importance of ensuring cybersecurity came first in the west. For example, in 2015, the US Department of Defense prepared the final version of the «Cybersecurity Strategy» [12]. In the US, the Department of Homeland Security allocates 16 critical infrastructures [13].

The investigations of critical infrastructure and, in particular, identification of critical facilities (CF) are a focus area in many countries and primarily in the United States. It can be reasoned by progressively increasing the development of new information technologies and the capacity of modern simulation complexes. Critical facilities of the RF infrastructure are the key objects (or their combinations) of infrastructures, when being affected might violate (or terminate) their functioning, thus causing loss of control, destruction of infrastructure, irreversible negative changes (or failure) of the economy of RF or its subject, or its administrative-territorial unit. The energy systems are surely referred to as the critical infrastructures [11].

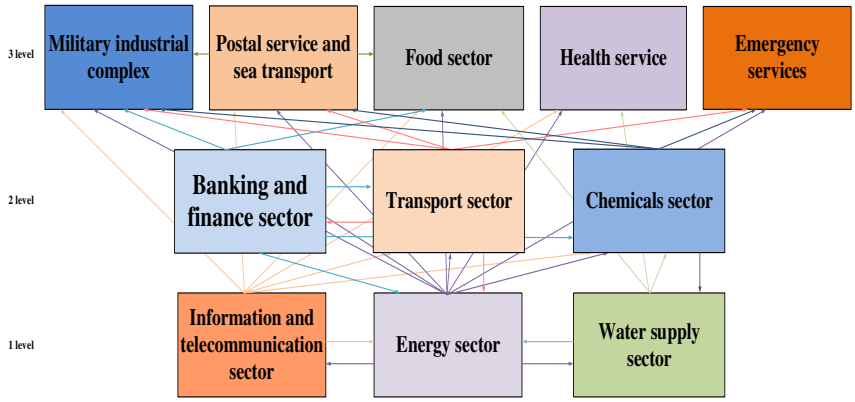


Figure 1
Critical infrastructures

As energy is regarded as critical infrastructure (Figure 1) [11], it is vital to identify critical facilities in the energy sector. In 2012 EMERCOM of Russia approved the «Methodology of attributing governmental and non-governmental proprietary objects to critical objects of the Russian Federation national security». Thus, the methodology for the formation of the list of gas transmission network CF has been provided [14]. It focuses on gas, being the energy system, and its transportation. The authors of the methodology proposed an indicator to locate critical facilities in the gas transportation network, which is determined by the relative gas shortfall to consumers (5% or more) because of the malfunction of these facilities. In this methodology ranking objects by the degree of significance for the country's economy in general and its individual regions is disregarded. The authors propose to perform ranking through the analysis of possible critical situations (CS) and taking into account the risks of the CS.

1.3 Energy Sector

The energy sector combines power plants and energy systems, including energy transporting main lines [15].

Energy security (ES) makes up a significant part in the Russian national security. ES threats are traditionally grouped as: (1) economic, (2) social-political, (3) technogenous, (4) natural and (5) managerial-legal. This threat list was supplemented with the cybersecurity threats [16], their implementation possibly provoking serious emergency situations in the energy fraught with a drastic reduction of energy resources to be provided to consumers [17]. The development of the Smart Grid conception in Russia exacerbates the problem of cybersecurity in energy.

Engineering systems are commonly designed, constructed, and operate under unavoidable conditions of risk and uncertainty; they are often expected to achieve multiple and conflicting objectives. The overall process of identifying, quantifying, and assessing risks should represent an integral and explicit component of the overall managerial decision-making process [18]. In the energy sector, the decision-making process is deteriorated by the complexity of the target area, difficulty to adequately model the energy facilities, availability of multiple criteria, and large-scale computational experiments [19].

For instance, the electric power facilities are complex structures with a multitude of equipment elements designed both for the main technical process and for ensuring its protection, the safety of maintenance personnel, and consumers of electrical energy. The operating capacity of generators, transformers, transmission lines, engines, electro-technological installations, and things is characterized by a variety of reliability indexes disregarding possible cyber-attacks onto the network or software or employee negligence.

In relation to these reasons, the authors propose to employ a risk-based approach to support decision-making in the identification of critical facilities in the energy sector subject to cyberthreats.

2 Methodology for Identifying the CF

Correct identification of critical facilities in the energy will reduce risks of financial losses in the event of damage or destruction of energy facilities, and also will facilitate the continuous supply of energy products to a consumer. The rapid spread of the computer environment, development of information technologies and the tendency of transition to intelligent energy make the cyber threats the most notable tactical threats of ES. Consequently, in June 2017 a series of cyberattacks have been undertaken in the Russian and international energy, telecommunications and financial companies and organizations. The cryptoware actions, among other things, caused failure in the functioning of the fuel station network. Serious consequences were avoided due to switching to the reserve management system.

In the energy sector, the validity of decisions is exacerbated by the need to process and comprehend large amounts of information belonging to several subject fields. To facilitate this task, the authors suggested developing an intelligent decision support system (DSS). Intelligent DSS is a complex of software tools for data analysis, modeling, forecasting, and decision-making.

In the energy sector, the identification of the CF is to be performed through several stages, as shown in Figure 2.

At the first stage (1) determine the criteria to affiliate the power facility to the CF. Criteria could be related to both the scale of fuel consumption and life quality within the region. A full set of objects is to be determined to solve the task of the CF identification within the country, region, municipality or some other territorial units.

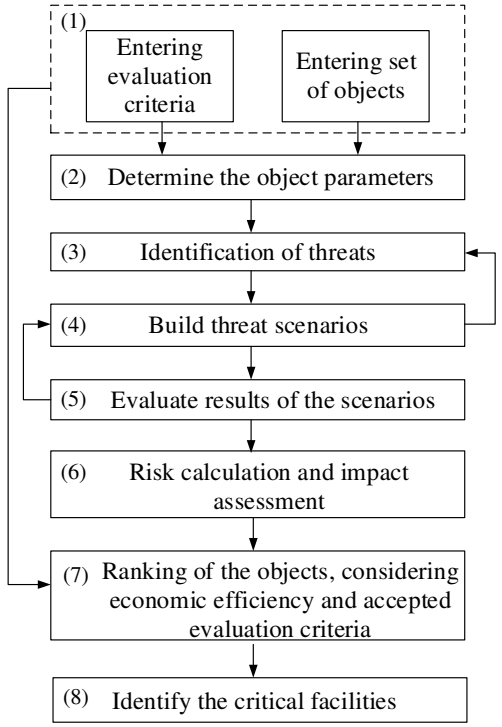


Figure 2
Identification of the CF stages

(2) Typify the object, make up its safety certificate and identify the vulnerability of cybersecurity via organizational, institutional, technical, and the operation procedures.

(3) Having information on the object and its vulnerability we establish the links between vulnerabilities and possible cyberthreats, considering the pre-existing preventive measures against security breaches.

(4) Build threat scenarios for each object or simulate some critical situations involving several objects, e.g. cascade failure.

(5) Evaluate results of the scenarios. An expert selects either most probabilistic scenarios or those meeting criteria for evaluation. It is also feasible to identify

frequently-involved objects in various scenarios and the most vulnerable assets in the scenarios for one object.

(6) Calculate and visualize the risks for each object on a risk map to be further analyzed by experts.

(7) Perform a ranking of energy facilities based on the calculated risks, considering economic efficiency and accepted evaluation criteria.

(8) Finally, identify the CF and make further decisions to ensure preventive countermeasures from the CS.

3 Risk-based Approach

The authors offer the risk-based approach to support the decision-making to ensure cybersecurity in the energy sector. This approach considers harm from damage or demolition of the object using quantitative and qualitative parameters, as well as probability for further damage or destruction of the object components, tailored to extent of damages and cascade failure. The formula of risks consists of three components (1),

$$R = \{T, V, D\}, \tag{1}$$

T – threat, V – vulnerability, D – damage by threat realization.

Threats are defined through the probability of event occurrence triggering critical situations. Cyberthreats might cause the subsequent implementation of the other ES threats. Cyberthreats might initiate scenario events; and final events are responsible for the consequences and determine the damage. The risk-based approach is aimed at developing scenarios for threat implementation leading to critical situations in the energy sector.

A scenario is represented by the set of conditions leading to a threat implementation and threats proper. In this context such conditions could be:

- an event is the threat realized, i.e. with a 100% probability of occurrence.
- consequences are threats to energy security or estimated losses of the energy facility assets.
- operating conditions of the information technology system affecting the threats involved in the scenario.

It is reasonable to apply the Bayesian networks to build cyber threat implementation scenarios using conditional probability.

The approach involves a description of the facility information technology system and further assets detailing. The asset contains vulnerabilities, which can be both

critical and not dangerous. The vulnerability criticality level is proposed to be determined following risk management standards [20], and using expert ranking. Vulnerabilities are classified in the work [21]. The authors offer in this work classification is made by subject area: general vulnerabilities, vulnerabilities related to information systems, and the ones specific to the energy sector.

Previously, the methodology for analyzing threats and assessing the risk of information technology security violations of the energy complexes was proposed [22]. In this methodology, the risk-based approach involves the risk analysis at various levels of the energy facility. The transition from the upper levels to the lower ones is accompanied by detailed elaboration and refinement of both the risk assessments and the representations about the object, its information technology system, and its domains. Having got the estimates received at detailed levels establishing feedback will allow adjusting the results obtained at the upper levels. This approach can be projected in the case of risk analysis and vulnerability assessment of several objects. The risk ranking of the objects is proposed to be implemented using risk scales, and on this basis to determine the critical objects. The solution of this task is specific and requires a lot of elaboration and detailing of the energy infrastructure under investigation. At this stage, we only assess risks of its facilities, rather than entire energy infrastructure.

This model is formulated through the fractal approach [19]; it will be described in the next section.

The asset vulnerabilities are determined from the databases and vulnerability banks, e.g. [23].

Damage is estimated for each consequence as the economic efficiency of the scenario. Economic efficiency is understood as the ratio between available risk assessments of the onset of critical situations, expressed in monetary units, and the cost of selected countermeasures with the evaluation criteria on hand.

4 Fractal Stratified Model of Interaction between Energy, Risks and Technologies

The fractal approach represents the subject field as a collection of information layers and their reflections from any layer to each one [19]. In this case, the energy (E) can be stratified into several levels and affiliate each level with the category of risk. Regarding ontology of the fuel and energy complex (FEC) of Russia such levels are: energy systems (Es), energy objects (Eo), and information technology systems (Ei), taking into account their subspecies (2).

$$E = \{Es, Eo, Ei\} \tag{2}$$

Therefore, as follows from (3), the risks (R) are stratified into groups (R_g), species (R_k), subspecies (R_s) and risks (r).

$$R = \{R_g, R_k, R_s, r\} \tag{3}$$

Also, the fractal approach is applied to information technologies implementing a risk-based approach. To identify the energy object vulnerabilities it makes sense to use an expert system and databank created for ontologies. The data is further transferred to the Bayesian Belief Networks to construct the threat implementation scenarios. The result of the expert work with the scenario is supplied to the risk assessment and risk maps. This approach is described as set of methods (4):

$$M = \{M_o, M_p, M_b, M_r\}, \tag{4}$$

where M_o is ontology modeling, M_p is structuring personal knowledge structure, M_b is Bayesian Belief Networks technology, M_r is the risk assessment method.

Every information technology is a specific layer. For the user of an intelligent system, work with layers is performed sequentially, but the feedback is possible with any layer passed to refine the data or to reassess them.

Figure 3 presents the fractal stratified model of energy sector risk assessment in terms of the proposed approach.

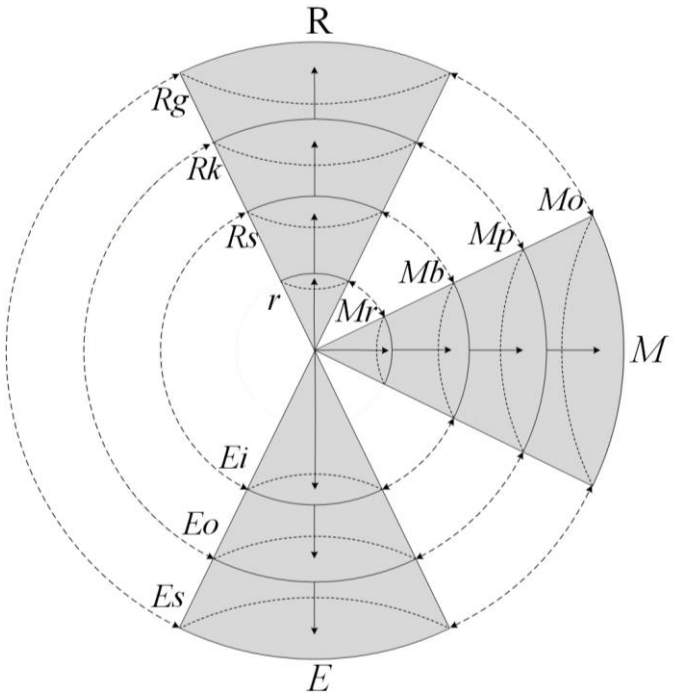


Figure 3
Fractal stratified model of interaction between energy, risks and technologies

5 Intelligent Decision Support System

At present, the structure of the intelligent system given in Fig. 4 is being designed. It consists of three interrelated components: (1) an expert system for supporting threats vector building, (2) the Bayesian network for modeling threat scenarios, and (3) the module for assessing the risk of cybersecurity violations, which includes visualization as a risk map. Besides, the researched prototype for the system described above has been made. The intelligent system is intended to support decision-making in the CF identification in the energy sector cyberthreats considered. It is being designed using a risk-based approach.

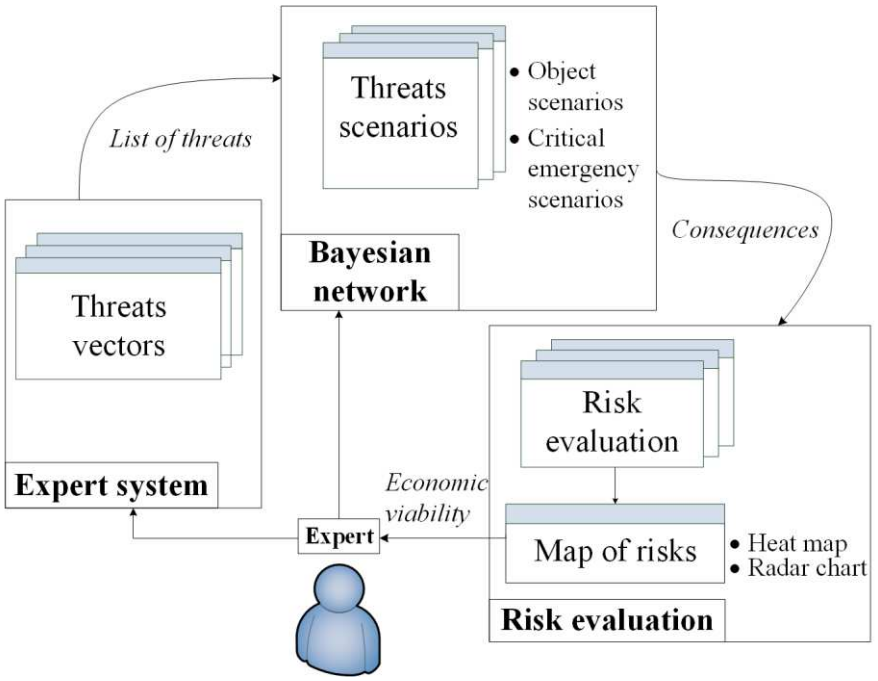


Figure 4
Structure of intelligent system

The expert system serves to obtain data on the energy facility and to make up threats vectors including penetration vector, attack vectors, and cybernegligence. For this purpose, an expert can use information about the information technology system, management, and legal internal standards, as well as the operational functioning of the facility. Having obtained the user information, the templates are filled to identify any relationships between the vulnerability of the object and the threat. The information in templates is wired to the Bayesian network component so that the user could construct the probabilistic scenarios of the extremal situation in energy. The work with a scenario results in locating the impact of cyberthreat

group implementation on the energy facility assets, leading to the implementation of other energy security threat groups. Information on the consequences of the implementation of the threat is transferred to the risk assessment component. This component consists of the calculated and visualizing components. The intelligent system work results in the determination of economic efficiency for each energy object considered, and also in the ability to rank objects.

The intelligent system means to support decision-making in the CF identification on the basis of constructing scenarios of possible extremal situations including critical situations and evaluating their significance by the groups of parameters. At present, the energy sector is very much concerned with emergency situations. The definition of a particular situation assumes assessment of the system state or objects through the scale: "norm", "pre-crisis" - a critical situation, "crisis" - an emergency situation. With this in mind, the critical situations are referred to the situations when something threatens a uninterrupted functioning of the technical objects and the objects of life support and/or the life or health threats of individuals or social (professional) groups [16].

The application of the intelligent system allows the expert to build up possible critical situations, based not only on his own experience, but also on those identified by the IDSS for the CF of energy determination, and regarding the scenarios provided by the other experts.

Conclusions

The article reports the energy sector as the critical infrastructure and important part of national security. Because of the lack of any approved methodology to identify CF in critical infrastructures and the trend of introducing modern information and communication technologies in the energy sector, the authors propose to apply the risk-based approach with modeling and analysis of critical situations arising from the implementation of cyberthreats. It is necessary to develop an intelligent system for the risk assessment of cybersecurity violations due to feasible cyberthreats with the risk-based approach and methodology of threat analysis and risk assessment applied.

Acknowledgment

This work was partially supported by RFBR grants №18-07-00714, №19-07-003514, №19-57-04003. The authors are grateful to this organization.

References

[1] Berman A. F., Nikolaychuk O. A., Yurin A. Yu., Pavlov A. I. A methodology for the investigation of the reliability and safety of unique technical systems // Journal of Risk and Reliability. Proceedings of the Institution of Mechanical Engineers. – Vol. 228, 2014 – pp. 29-38

[2] Makhutov N. A, Gadenin M. M. Ensuring the safe operation of technosphere and population facilities using risk criteria // Abstracts of the

XXI International Scientific and Practical Conference on the problems of protecting the population and territories from emergency situations. – 2016 – pp. 137-146 (in Russian)

[3] Rot A. IT Risk Assessment: Quantitative and Qualitative Approach // Proceedings of the World Congress on Engineering and Computer Science WCECS 2008, 2008, San Francisco, USA

[4] Rot A. Enterprise Information Technology Security: Risk Management Perspective // Proceedings of the World Congress on Engineering and Computer Science 2009, Vol. II WCECS 2009, 2009, San Francisco, USA

[5] Allodi L. Massacci F. Security Events and Vulnerability Data for Cybersecurity Risk Estimation // Risk Anal. – 2017 – T. 37 – № 8– 1606–1627pp.

[6] Battaglini A., Komendantova N., Brtnik P., Patt A. Perception of barriers for expansion of electricity grids in the European Union // Energy Policy 47 (2012) – pp. 254-259

[7] Schinko T., Komendantova N. De-risking investment into concentrated solar power in North Africa: Impacts on the costs of electricity generation // Renewable Energy 92 (2016) – pp. 262-272

[8] Lojka T., Bundzel M., Zolotova I. Service-oriented Architecture and Cloud Manufacturing // Acta Polytechnica Hungarica. - Vol. 13, No. 6, 2016 – pp. 25-44

[9] Chen J. Zhu Q. Security as a Service for Cloud-Enabled Internet of Controlled Things Under Advanced Persistent Threats: A Contract Design Approach IEEE Trans. Inf. Forensics Secur. – 2017 – T. 12 – № 11– 2736–2750pp.

[10] Jovanovic F., Milijic N., Dimitrova M., Mihajlovic I. Risk Management Impact Assessment on the Success of Strategic Investment Projects: Benchmarking Among Different Sector Companies // Acta Polytechnica Hungarica. Vol. 13, No. 5 (2016) – pp. 221-241

[11] Kondrat'ev "The current trends in research of Critical Infrastructure in foreign countries". №1. Foreign Military Review, 2012, pp. 19-30 (in Russian)

[12] Cyber Strategy // The Department of Defense, 2015 [Online] URL: https://defence.ru/document/61/

[13] Critical Infrastructure Sectors // Official website of the Department of Homeland Security [Online] URL: https://www.dhs.gov/critical-infrastructure-sectors

[14] A. V. Edelev, S. M. Senderov, I. A. Sidorov "The use of distributed computing to identify critical objects of the Russian gas transmission

network". №1. Information and Mathematical Technologies in Science and Management, 2016, pp. 55-62 (in Russian)

[15] N. I. Voropai. Reliability of power supply systems. - Novosibirsk: Science, 2015 -208 p.

[16] Massel L., Massel A. Cyber security of Russia's energy infrastructure as a component of national security. / Proceeding of the International Conference on Problems of Critical Infrastructures, 6th International Conference on Liberalization and Modernization of Power Systems. Edited by Z. A. Styczynski and N. I. Voropai. 2015. C. 66-72

[17] Nazir S. Assessing and augmenting SCADA cyber security: A survey of techniques / S. Nazir, S. Patel, D. Patel // Comput. Secur. – 2017 – T. 70– 436–454c.

[18] Yacov Y. Haimes. "Systems-based risk analysis". In: "Global Catastrpphic Risks" Nick Bostrom, Milan M. Cirkovic (ed), Oxford, 2008, pp. 146-163

[19] Massel L. V. "Fractal approach to he structuring of knowledge and examples of its application". Vol.6, № 2 (20). Ontology of designing, 2016, pp. 149-161(in Russian)

[20] Draft International Standard ISO/DIS 31000 «Risk management – Principles and guidelines on implementation», ISO, 2008

[21] Young A. L., Yung M. Cryptovirology // Commun. ACM – 2017. – Vol. 60 – № 7 (2017) – P. 24-26c.

[22] Massel A., Massel L. The current state of cyber security in Russia's energy systems and the proposed activities for situation improving / Proceeding of the International Conference on Problems of Critical Infrastructures, 6th International Conference on Liberalization and Modernization of Power Systems. Edited by Z. A. Styczynski and N. I. Voropai. 2015, C. 183-189

[23] Data Security Threats Database of the FSTEC of Russia [Online] Available: http://bdu.fstec.ru/

Complex Analysis of Medical Data with Data Mining Usage

Nafisa Yusupova[1], Gyuzel Shakhmametova[1], Rustem Zulkarneev[2]

[1]Ufa State Aviation Technical University, Computer Science & Robotics Dept., Karl Marks str. 12, 450008 Ufa, Russia, yussupova@ugatu.ac.ru, siit@ugatu.su

[2]Bashkir State Medical University, Faculty of General Medicine, Lenin str. 3, 450008 Ufa, Russia, pvb@bashgmu.ru

Abstract: the new approach to the medical, in particular, the toxicological data analysis is considered. For the data processing multilevel system realization, the three-stage technique for data analysis with data mining usage is offered. The results of the research are discussed.

Keywords: medical data processing; data mining; complex analysis

1 Introduction

The most important field of modern IT application for medical purposes is processing data for solving specific tasks in such areas as diagnostics, treatment and prevention of diseases. It is of paramount importance especially for those medical professionals who must manage particularly vast amounts of entrance information or execute a complex algorithm of data processing which at times represents serious difficulties for the decision-maker [1]. The importance of this task increases with the avalanche-like accumulation of information due to the improvement of technologies for collecting and storage of information today. Proceeding from understanding the need for high-quality changes in the approaches to processing of the growing volumes of medical information, in recent years the increasing development is gained by methods of the data analysis which include statistical methods as well as methods of data mining [2]. Introduction of information technologies in medical practice allows to radically turn around a situation though there is still a set of problems that need to be solved by joint efforts of information scientists and physicians.

The main idea of research consists in the realization of the new approach to the analysis of medical, in particular, toxicological data. In this article, the technique

for the analysis of medical data is offered and examples of the analysis of the results yielded with interpretation on the example of the toxicological data collected in the Republic of Bashkortostan for 2015-2016 are given. For organizing a multilevel system of information processing, it is offered to create a technique of the three-stage data analysis that allows the researcher to gain an impression about the structure of data, to understand the main regularities, to take new, earlier unknown knowledge on the basis of the considered selection, and also to increase the efficiency of the information analysis process.

2 State of Art

Today, one of the most popular medical data in the field of toxicology analysis methods is the elementary visual analysis utilizing diagrams and charts. In general, for the medical data study, statistical and intelligent analysis techniques are used. In [3], the medical data analysis methodology is considered, the research is aimed at increasing the efficiency of doctors'/clinical physicians' activities. The preliminary analysis was carried out using the charts of dispersion and density, and the applied analysis methods included correlation analysis, logistic regression, and the accidental trees nonparametric qualifier. In [4], several selection methods that may be used in medical studies with various scenarios and problems are considered. The main applied methods are sampling and randomization. A large number of researches is devoted to Data Mining applications for the hidden patterns recognition in the medical data analysis tasks. For example, in [5], the Data Mining use for obesity disease detection in children is discussed. The cluster analysis used for the definition of children's groups with similar results after the executed treatment is considered. The methods applied in the research are dispersion analysis, cluster analysis (K-averages method), and the limited search algorithm. In [6], the analysis of data on acute peroral poisonings based on the REACH data is considered, the main research direction is the separate chemicals impact definition on an organism. The data analysis is carried out by the following methods: training at examples, neural networks, and k-nearest neighbors. In [7], the research of acute exogenous poisonings in Altai Region during 1997-2013 is given. The research allows us to assess a toxicological situation in the region and includes the systematization of acute poisonings in a section of gender and age and social groups of the population. Research was carried out in MS Excel using data visual analysis methods. In [8], the research of possible risks in patients with acute alcohol poisoning is conducted using descriptive statistics and visual analysis methods.

All data analysis research considered by authors use a limited number of methods, i.e., only the visual analysis, or only the statistical methods, or only the Data mining methods. Today there is no complex technique for the medical

toxicological data analysis which could allow us to process data most fully and to get the maximum quantity of new knowledge and hidden patterns.

Medical data, from the point of view of the analyst, have the next features:

- the data are retrospective;
- as a rule, they are diverse and have quantitative and qualitative indicators;
- they are semi-structured;
- as a rule, such data do not submit to normal distribution.

In this regard, there are difficulties with processing and analysis of such data because of the heterogeneity of the data that demands the application of various methods and approaches for data analysis.

Several methods are used today for processing medical data, i.e., descriptive and inductive statistics; correlation, regression, multiple-factor discriminant, cluster analysis; method of artificial neural networks, and others. Comparative characteristics of these methods are provided in Table 1.

Table 1

Contains the result of comparing in pairs with the final result

Method	Goal	Advantages	Shortcomings
Descriptive statistics	Processing of empirical data, systematization, quantitative description by means of main statistics	Effective and rather easy way of data consideration and description; convenient way of information representation [9]	It does not make conclusions about population based on results of special cases research [10]
Inductive statistics	Check of statistical hypotheses for the law of distribution	Simplicity of method application [11]	Low level of reliability; considerable errors for small size samplings [12]
Correlation analysis	Detection of existence and level of communication between two and more variables for predicting possible value of one of them if another one is known	Possibility of creating new rules for interaction of functions and also an assessment of functions interaction [13]	The results may be used only in the immediate research field or in one close to it [12]
Regression analysis	Detection of dependence	It allows to define dependence (linear,	Application for processing of

	between independent variable and one or several dependent variables	nonlinear) quantitatively, in the form of a mathematical formula [14]	qualitative data is impossible [15]
Multiple-factor analysis	Detection of latent variables or the factors causing multiple correlative communications	Possibility of smaller number of data utilization therefore leading to more expedient model generation [15]	The subjectivity of results interpretation, complexity of the procedure; requires several cycles of conducting a procedure for obtaining qualitative result [16]
Discriminant analysis	Classification of objects, i.e., reference to one of several set groups (classes)	It allows to make the multidimensional analysis of data [15]	By improperly conducted research, the developed models will not work with new data [17]
Cluster analysis	Method of classification analysis; its purpose is splitting a set of the studied objects and signs into uniform groups	It does not impose restrictions for a type of the objects considered, allowing to work with large volumes of data and to classify objects by a number of signs [18]	Subjective interpretation of results [19]; at different introduction can give different results [20]
Neural networks	Generalization and allocation of hidden dependences between the input and output data	No need of knowledge formalization; orientation to parallel processing; possibility of multidimensional data and knowledge processing without increase in labor input [21]	Difficulties in explanation of neural network functioning results [22]; impossibility to guarantee repeatability and uniqueness of obtaining results [23]

Each of these methods has its advantages and shortcomings, and many of them have restrictions in the character of the analyzed data. The problem is that a single method may only solve a narrow task of the data analysis which is not enough for decision-making. Authors offer a complex technique for the analysis of medical data in the field of toxicology including methods of mathematical statistics as well as methods of data mining, allowing to carry out a comprehensive analysis of the data and to benefit from the largest possible amount of knowledge, interrelations, and patterns. The research novelty consists of the new approach to the toxicology data analysis which represents the complex three-stage analysis with visual,

statistical and Data Mining methods use and allows us to study the data comprehensively.

3 Data Analysis Stages

The technique of medical data analysis in the field of toxicology including complex data analysis and interpretation of results, and consisting of the next main stages is suggested as follows (Figure 1):

1) Primary statistical data analysis by the means of MS Excel, visualization of raw data for understanding their quantitative and qualitative structure, making hypotheses about patterns existence in data.
2) Statistical analysis of the data using multidimensional analysis and nonparametric methods for confirmation or denial of the hypotheses made at the first stage. This stage is also the stage of the "prospecting" analysis for the purpose of making new hypotheses and assumptions.
3) Data mining assumes search of hidden regularities and patterns in the data with utilizing Data mining, knowledge discovery, confirmation or denial of the hypotheses made at the previous stages.

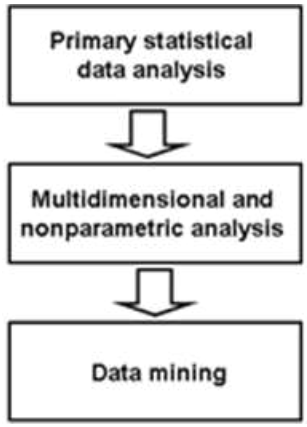


Figure 1
Stages of medical data analysis in the field of toxicology

For the analysis, the sampling of the toxicological data on the Republic of Bashkortostan for 2015-2016 has been taken. The volume of the database for the analysis is 6 338 records. The structure of the data is provided in Table 2.

Table 2

Data structure

Parameter name	Data type	Data kind
Gender	character	qualitative
Age	integer	quantitative
Social group	character	qualitative
Address (city, region)	character	qualitative
City/village	character	qualitative
Place of poisoning	character	qualitative
Date of poisoning	date	qualitative
Diagnosis	character	qualitative
MKB10 code	character	qualitative
Who has made the	character	qualitative
Health facility	character	qualitative
Number of victims	integer	quantitative
Lethal outcome	logic	qualitative
Purpose	character	qualitative
Place of obtaining poison	character	qualitative
Other	…	…

The data are diverse, and only a part of the data is quantitative (numerical); for the most part, the data are qualitative (symbolical) which, on one hand, complicates the statistical analysis, but on the other hand this creates prerequisites for data mining application.

A feature of biomedical data is primarily the heterogeneity of the data itself, they can be both quantitative and qualitative, in most cases the data is not subject to normal distribution. Medical data is generally semi-structured. Toxicological data are generally the same as medical data. This is information about the state of a person - his age, social group, objective and subjective signs of disease (poisoning), etc.

Our data can be divided into the following kinds:

1) Quantitative data - parameters; they can be characterized by discrete values: the age of the patient, the number of victims.

2) Qualitative data - characteristics; do not give in to exact assessment, though can be ranged (for example, are systematized on conditional points: one point, two points, etc.). The social group, the address, the place of poisoning, date of poisoning (data type - date), the diagnosis, the MKB10 code, who established the diagnosis, where there took place treatment, the place of poisoning, the poisoning purpose, the place of acquisition of poison, etc. Qualitative characteristics can be classified into only two categories - sex, individual and group poisoning, lethal outcome (logical data type).

Thus, the data are presented in different kinds, in different types and formats, not subject to normal distribution. It is necessary to analyze not only within each kind and type of data, but also to identify patterns between data belonging to different kinds and types.

3.1 Primary Statistical Data Analysis

The primary research of raw data is the first stage of the analysis and is carried out for the detection of the most general regularities and tendencies, character and properties of the analyzed data, and laws of the analyzed data distribution [25]. The results of the initial prospecting analysis are not used for making decisions, their purpose is to help in the development of the best strategy for the profound analysis, hypotheses making, specification of mathematical methods, and model feature application. The prospecting analysis helps to concisely describe the structure of data in a visual form, and then to research it in more detail by statistical analysis and data mining. The purpose of this stage is to visualize data and to collect the maximum quantity of hypotheses for possible interrelations and regularities in data.

To carry out the initial analysis of data, it is necessary to process data and to output them the quantitative indices. For this purpose, selections have been divided into groups on the grouping indicators. We will consider the dynamics of quantity for cases of acute poisonings with various poisons on the example of age groups in Figure 2.

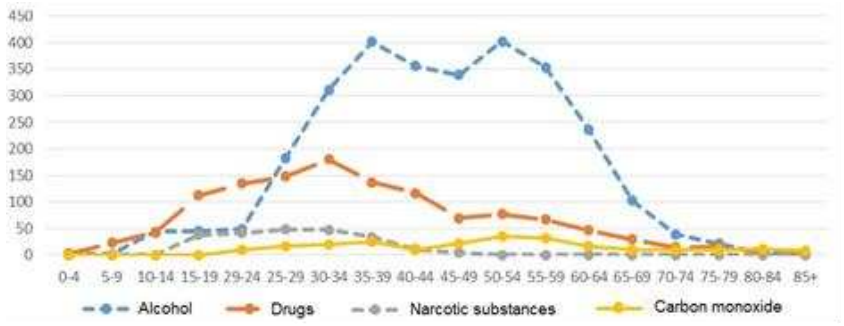


Figure 2

Dynamics for quantity of acute poisoning cases with various poisons in different age groups

The most frequent cause of acute poisonings in the Republic of Bashkortostan is alcohol poisonings. They are most prevalent in the group of 20 to 70 years of age. The maximum peak is reached for the groups from 35 to 60 years of age. Narcotic poisonings are registered in age groups from 15 to 40 years. Medicinal poisonings and also cases of poisoning with carbon monoxide are characteristic for all age

groups. The percentage ratio for the quantity of acute poisonings depending on age and gender is presented in Figure 3.

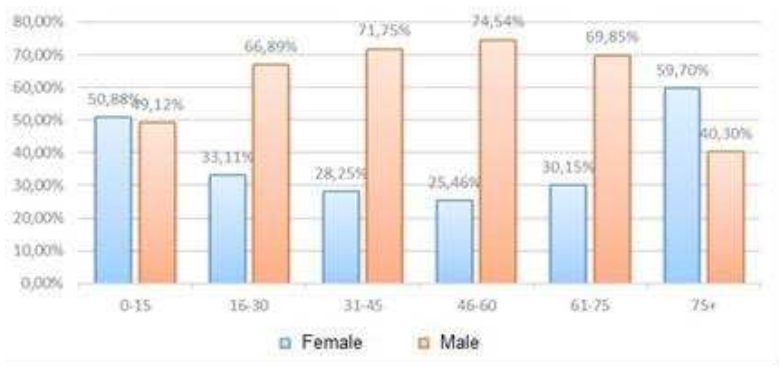


Figure 3

Structure of acute poisonings depending on age of men and women

From the chart, it is possible to see that the peak of the acute poisonings in men is reached at the age of 46-60 years, and in women, at the age of 75 and over, which, in turn, is not evident knowledge and demands further research.

For a clearer understanding of poisoning distribution in different age groups, dynamics of mortality (Figure 4) have been studied.

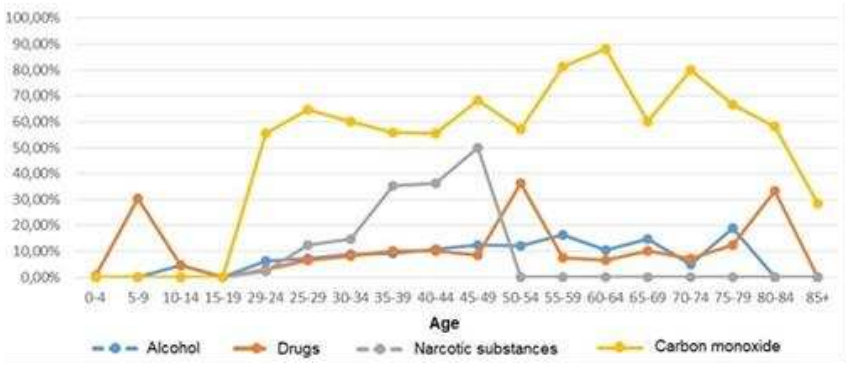


Figure 4

Mortality rate from acute poisonings with various poisons by age groups

The greatest indicator of mortality is present at poisonings with carbon monoxide. Alcohol poisonings mortality rate around 20 years of age is rather low but steadily growing. Medicinal poisonings also have an average level, with 3 peak age groups in which the death rate considerably exceeds the general. Risk groups are at the age of 5-9 years, 50-54 years, and also after 80 years. For the specification of

results, the structure of lethal cases of acute poisonings depending on the cause for men and women (Figure 5) is considered.

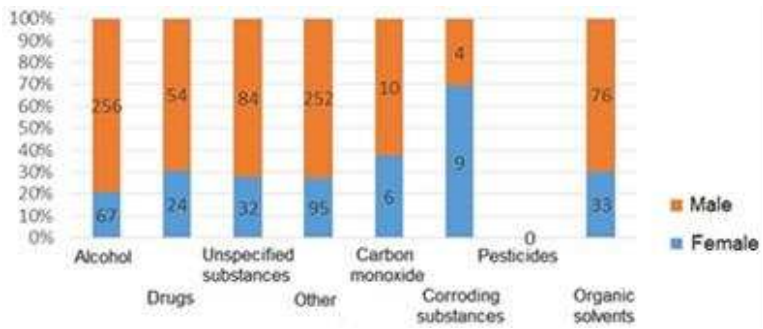


Figure 5

Structure of lethal cases of acute poisonings depending on the cause for men and women

Thus, already at the first stage of the prospecting analysis, it is possible to draw the following conclusions: the greatest number of poisonings occurs from 25 up to 60 years, this index is also high for people aged from 1 up to 3 years; the number of cases of acute poisonings for men is much higher than for women; the greatest death rate is for people aged from 25 up to 65 years; the death rate for men is higher than for women in all basic causes of poisonings, except the ones from corroding substances.

3.2 Data Statistical Analysis

The next stage of analysis is deeper data studying by means of the statistical analysis methods.

According to statistical principles used in the basis, the methods are subdivided in [12]:

- parametric - applied mainly to the analysis of normally distributed quantitative signs;
- nonparametric - applied to the analysis of quantitative signs irrespective of their distribution and to the analysis of qualitative signs.

Nonparametric methods are developed for those situations when the researcher knows nothing about any parameters of the analysed data [12]. Owing to features of the medical data that is listed above it is more effective to apply methods of the nonparametric analysis to their processing.

The examples of data processing results with nonparametric analysis methods for the main causes of poisoning executed in version 13.2 STATISTICA package are given below. In this analysis, dynamics of number of acute poisonings for the

main causes and quantity of cases of poisonings with a lethal outcome for 2015-2016 have been analysed.

From sample data, the following indicators for the number of cases of acute poisonings depending on the main causes for men and women (Table 3) have been received.

Table 3

Number of cases of acute poisonings by main causes for men and women

Poisoning cause	Number of poisonings	
	Female	Male
Alcohol	553	2346
Other	152	203
Drugs	888	811
Unspecified substances	254	727
Organic solvents	20	33
Pesticides	29	38
Corroding substances	49	72
Carbon monoxide	60	101

Then the Kruskal-Wallis test has been applied to compare the number of poisonings according to the main diagnoses (Figure 6).

Kruskal-Wallis ANOVA by Ranks; Var2 (Spreadsheet18)
Independent (grouping) variable: Var1
Kruskal-Wallis test: H (7, N= 137) =73,03182 p =,0000

Depend.: Var2	Code	Valid N	Sum of Ranks	Mean Rank
Alcohol	101	18	1803,500	100,1944
Other	102	18	1428,500	79,3611
Drugs	103	18	1875,000	104,1667
Unspecified substances	104	18	1703,500	94,6389
Organic solvents	105	15	374,500	24,9667
Pesticides	106	14	447,000	31,9286
Corroding substances	107	18	827,500	45,9722
Carbon monoxide	108	18	993,500	55,1944

Figure 6

The results of Kruskal-Wallis test for comparing the number of poisonings according to the main diagnoses

The significance value of $p<0.05$ signifies the statistical importance of results. The value for the criterion of H is exceeded the tabular one. Therefore, the statistical importance of distinctions is high. The greatest contribution to poisonings, from the largest to the smallest: alcohol, drugs, unspecified substances, carbon monoxide.

The indicators numbers of acute poisonings with lethal outcome depending on the main causes for men and women are presented in Table 4.

Table 4
Number of acute poisoning cases with a lethal outcome depending on their main causes for men and women

Poisoning cause	Number of poisonings	
	Female	Male
Alcohol	67	256
Other	24	54
Drugs	32	84
Unspecified substances	95	252
Organic solvents	6	10
Pesticides	9	4
Corroding substances	33	76

The results of Kruskal-Wallis test (Figure 7) have shown that the significance value lies also below 0:05 which signifies the statistical importance of results of this testing. Therefore, it is possible to say that separate indicators make significant contributions to the death rate from acute poisonings.

Kruskal-Wallis ANOVA by Ranks; Var3 (Spreadsheet21)
Independent (grouping) variable: Var2
Kruskal-Wallis test: H (6, N= 92) =33,45563 p =,0000

Depend.: Var3	Code	Valid N	Sum of Ranks	Mean Rank
Other Quantity	101	15	646,500	43,10000
Drugs Quantity	102	16	631,500	39,46875
Carbon monoxide Quantity	103	16	782,000	48,87500
Unspecified substances Quantity	104	16	1035,000	64,68750
Alcohol Quantity	105	13	869,000	66,84615
Organic solvents Quantity	106	10	166,000	16,60000
Corroding substances Quantity	107	6	148,000	24,66667

Figure 7
The results of Kruskal-Wallis test for comparison of lethal outcome according to the main diagnoses

Ranging of the leading causes of death, decreasing in-order: alcohol, unspecified substances, carbon monoxide, drugs, corroding substances, organic solvents.

The analysis of dynamics for numbers of acute poisonings for 1981-2016 (Fig. 8) had been carried out.

Kruskal-Wallis criterion has shown the statistical importance of differences in indications for selection of poisoning levels by main causes. The greatest contribution is made by alcohol poisonings, the level of poisonings with carbon monoxide and unspecified substances is also high. Indicators of narcotic and medicinal poisonings which had considerably high rates by consideration of dynamics of poisonings in the last two years throughout the long period have shown the lowest level.

Kruskal-Wallis ANOVA by Ranks; Quantity (Spreadsheet1)
Independent (grouping) variable: Diagnosis
Kruskal-Wallis test: H (7, N= 288) =187,7026 p =0,000

Depend.: Quantity	Code	Valid N	Sum of Ranks	Mean Rank
T51 Alcohol	101	36	8869,000	246,3611
T58 Carbon monoxide	102	36	8062,000	223,9444
T40 Narcotic substances	103	36	3721,500	103,3750
T54 Corroding substances	104	36	4013,500	111,4861
T52-T53 Organic solvents	105	36	2969,000	82,4722
T36-T50 (without T40) Drugs	106	36	2762,500	76,7361
T60 Pesticides	107	36	3341,500	92,8194
T65 Other and unspecified substances	108	36	7877,000	218,8056

Figure 8

The results of Kruskal-Wallis test for comparison of death rate according to the main diagnoses during the period for 1981-2016 in RB

At the second stage analysis of data, deep interrelations between data that can be used for decision-making are detected. At the same time, both stages are intended for the prospecting analysis, the best understanding of data, and hypotheses making which are preliminary steps for data mining.

3.3 Data Mining Stage

The technology of data mining allows to discover such patterns among large volumes of data which cannot be found by statistical ways of data processing but are objective and practically useful. Using these methods. the researcher can observe five main patterns in data [2], namely:

- association – several events are connected with each other;
- sequence – a chain of the events connected in time;
- classification – reference of an object to one of the classes with known characteristics;
- clustering – allocation of uniform groups of objects;
- temporary templates - dynamics of behavior of target indicators.

All listed patterns are applicable to medical data.

Results of the data processing with use of decision trees has been executed in the Deductor Studio Academic 5.3.0.88 package are given below.

Decision trees are a method of representation governed by a hierarchical, consecutive structures where the only unique knot giving the decision corresponds to each object. This method is already actively applied in medicine and biomedical studies and allows to make the diagnosis and to predict possible consequences of treatment.

Having set the MKB10 Code as an output parameter, and all other available parameters as entrance, we have received model which will allow us to estimate at the initial stage the importance of a contribution of separate parameters to a resultant indicator – the "MKB Code" describing the type and causes of poisoning. The greatest contribution to the definition of the poisoning diagnosis is made by such indicators as "the purpose of toxic agents intake ", age of the patient, and lethality; an insignificant contribution is made by "the number of victims", date of poisoning and the patient gender (Figure 9).

Target attribute: MKB10 Code

№	Number	Attribute	Certainty, %
1	9	Purpose	51,605
2	2	Age	22,032
3	7	Letal_outcome	16,630
4	8	Number_of_victims	4,746
5	5	Date_of_poisoning	2,009
6	1	Gender	1,945
7	4	City/village	0,812
8	6	Who_has_made_the_diagnosis	0,221
9	3	Social_group	0,000

Figure 9

Analysis for the importance of separate factors influence on the diagnosis

From the received model, the following results were obtained:

1) At the age of more than 80 years, the most frequent causes of poisonings are an erratic drug intake or the corroding substances and also poisoning with carbon monoxide.
2) Up until the age of 36, men in public places become poisoned with narcotic substances, after this age, more often with drugs.
3) Group poisonings are most often caused by carbon monoxide.

Also, the contribution of separate parameters to the probability of a lethal outcome in case of acute poisoning was evaluated. The largest contribution is made by the Who Set the Diagnosis parameter; however, upon studying the initial selection it becomes clear that in the majority of lethal cases, the diagnosis is made by the forensic scientist. Other contributors to the probability of a lethal outcome are such indices as the cause of poisoning and the age of the patient (Figure 10).

Target attribute: Lethal_outcome			
№	Number	Attribute	Certainty, %
1	9	Who_has_made_the_diagnosis	95,168
2	8	MKB10_Code	2,896
3	2	Age	1,936
4	7	Date_of_poisoning	0,000
5	11	Purpose	0,000
6	10	Number_of_victims	0,000
7	3	Social_group	0,000
8	1	Gender	0,000
9	4	Address_(city,_region)	0,000
10	6	Place_of_obtaining_poisoning	0,000
11	5	City/village	0,000

Figure 10

Analysis of the importance of influence of separate factors on the probability of a lethal outcome

In case of rendering medical assistance (both by the doctor and the paramedic), it is possible to avoid a lethal outcome in almost all cases except for those caused by the influence of narcotic substances in people over 30 years of age (Figure 11).

Antecedent	Consequent	Support	Reliability
IF		1997	1755
Who_has_made_the_diagnosis = Doctor	False	1608	1587
Who_has_made_the_diagnosis = Unknown		74	61
MKB10_Code	False	0	0
MKB10_Code = Drugs	False	39	39
MKB10_Code = Narcotic substances		25	13
Age < 30.5	False	10	10
Age >= 30.5	True	15	13
MKB10_Code = Organic solvents	False	3	3
MKB10_Code = Pesticides	False	0	0
MKB10_Code = Corroding substances	False	6	6
MKB10_Code = Carbon monoxide	False	1	1
Who_has_made_the_diagnosis = Forensic scientist	True	210	206
Who_has_made_the_diagnosis = Paramedic	False	105	103

Figure 11

Decision tree denoting medical assistance and lethal outcome

For specification of results, all non-significant parameters have been excluded from selection, and the value of a contribution of the most significant parameters to rgw probability of a lethal outcome (Figure 12) was estimated once more. The poisoning diagnosis parameter defines the probability of a lethal outcome by 80%. A small contribution is made by the age of the patient (18%). The influence of the patient belonging to a certain social group was revealed in this research in a very insignificant form.

To verify the results received above, the impact of the cause of poisoning, age, and social group on the probability of a lethal outcome has been considered (Fig. 13). The indicator of a lethal outcome is equal to "False" (false) in all groups, except for patients with the diagnosis of poisoning with narcotic substances who were older than 42.

Target attribute: Lethal_outcome			
№	Number	Attribute	Certainty, %
1	3	MKB10_Code	81,521
2	1	Age	18,026
3	2	Social_group	0,452

Figure 12

The analysis of the importance of the influence of the set factors on the probability of a lethal outcome

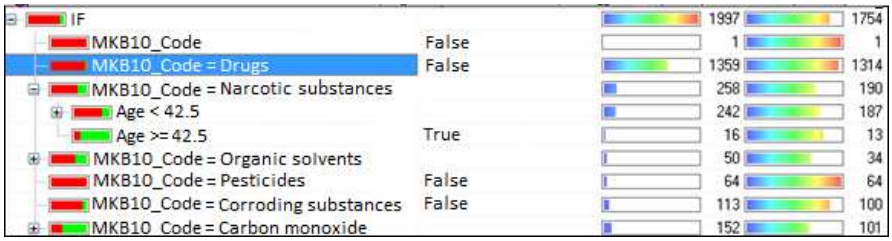


Figure 13

Decision tree with use of the set parameters

Thus, it is possible to draw a conclusion that the influence of narcotic substances becomes more considerable with age and with high probability leads to lethal cases after the age of 40. Let us check the assumption of the influence of age on the cause of poisoning. The analysis of the decision tree denoting social groups and main causes for toxic agent intake has shown that in women, cases of suicide poisonings are the most frequent cause for all main social groups. We will introduce amendments to the model and look at the influence of age on the cause of poisoning (Figure 14).

The analysis of the received results shows that the most frequent cause is drug intake for the purpose of suicide; the age, except for childhood, does not play an essential role, and the number of suicides for women is much higher than for men.

Results of the third analysis stage with the application of data mining allow us to draw the following main conclusions:

1) In the analysis of the main causes for poisonings in standard age groups, it is revealed that mortality from narcotic substances considerably increases after the age of 30, and in most cases, acute poisonings from drugs lead to a lethal outcome.

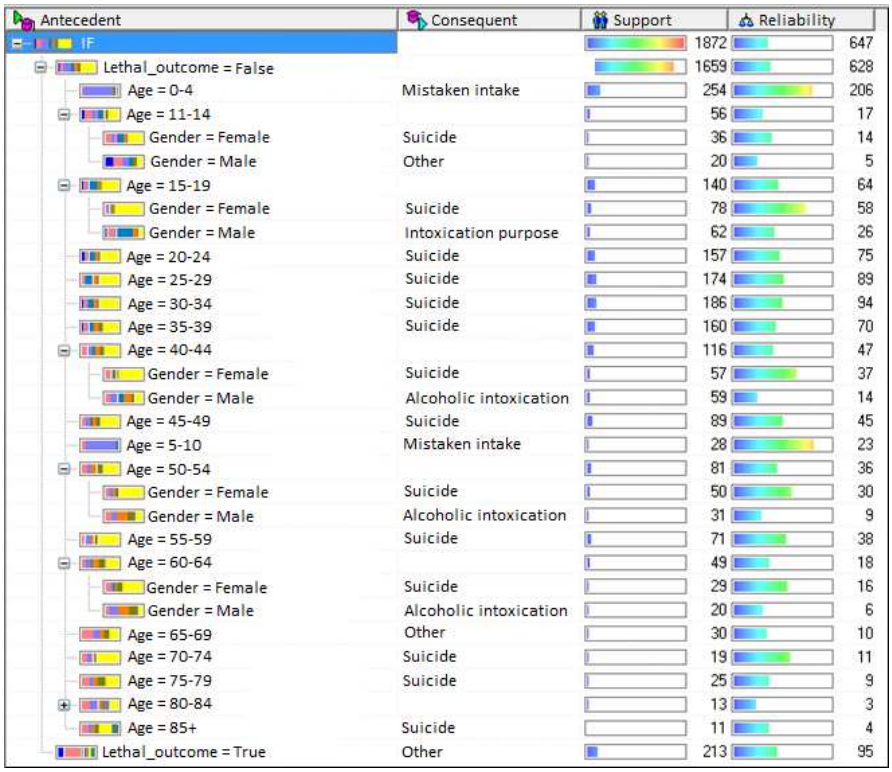
2) At this stage, alcohol poisonings have been excluded from the research. The most frequent cause of acute poisonings, as well as lethal outcome, was medicinal poisonings. For women, acute medicinal poisonings most often result from suicide intentions, but extremely seldom lead to a lethal outcome. For men, the level of suicide poisonings cases with medicinal substances is

much lower. However, mortality is much higher. Most often, medicinal poisonings are caused in men seeking to receive alcoholic intoxication.

3) At the retirement age, medicinal poisonings are most often caused by intake errors.

4) Timely assistance by the doctor or paramedic in most cases allows to avoid lethal outcome, except for cases of acute narcotic poisoning in people over 30.

The received regularities represent a new, unevident knowledge taken from the data and suitable for use in decision-making.



Antecedent	Consequent	Support	Reliability
IF		1872	647
Lethal_outcome = False		1659	628
Age = 0-4	Mistaken intake	254	206
Age = 11-14		56	17
Gender = Female	Suicide	36	14
Gender = Male	Other	20	5
Age = 15-19		140	64
Gender = Female	Suicide	78	58
Gender = Male	Intoxication purpose	62	26
Age = 20-24	Suicide	157	75
Age = 25-29	Suicide	174	89
Age = 30-34	Suicide	186	94
Age = 35-39	Suicide	160	70
Age = 40-44		116	47
Gender = Female	Suicide	57	37
Gender = Male	Alcoholic intoxication	59	14
Age = 45-49	Suicide	89	45
Age = 5-10	Mistaken intake	28	23
Age = 50-54		81	36
Gender = Female	Suicide	50	30
Gender = Male	Alcoholic intoxication	31	9
Age = 55-59	Suicide	71	38
Age = 60-64		49	18
Gender = Female	Suicide	29	16
Gender = Male	Alcoholic intoxication	20	6
Age = 65-69	Other	30	10
Age = 70-74	Suicide	19	11
Age = 75-79	Suicide	25	9
Age = 80-84		13	3
Age = 85+	Suicide	11	4
Lethal_outcome = True	Other	213	95

Figure 14

Decision tree denoting age groups and the main causes for the intake of toxic agents in specified groups

Conclusions

The analysis of researches in the field of toxicological data processing has shown that the majority of researches is conducted with the use of the simplest statistical methods. At the same time, traditional mathematical statistics uses the concept of

averaging on selection, i.e., operates with average characteristics which often are nonexistent whereas methods of the intellectual analysis can find unevident regularities in data. However, applying all of these methods, it becomes possible to study data in its entirety. Methods of mathematical statistics are useful mainly for check of hypotheses formulated in advance and for the prospecting analysis, representing an effective base for the subsequent application of methods of the intellectual analysis.

On the basis of the obtained information about existing methods of data analysis, the technique for medical data complex analysis in the field of toxicology has been developed. The technique includes three main stages: prospecting analysis of data with the use of visual methods of the analysis, statistical nonparametric, and also data mining.

The developed technique is applied to the analysis of data on acute poisonings in the Republic of Bashkortostan in 2015-2016. The results received during such analysis can help the leaders of medical institutions and chief specialists of governing bodies in the analysis of the indicators characterizing dynamics of tendencies in population health, distribution resource planning for health care of an area, and management of specialized health services.

The authors plan to conduct further researches in the area of data mining for the extraction of implicit regularities.

Acknowledgement

The reported study was funded by RFBR according to the research projects No. 19-07-00780, 19-07-00709, 18-07-00193 and state task No. FEUE-2020-0007.

References

[1] E. H. Shortliffe, J. J. Cimino, *Biomedical Informatics*. Springer-Verlag, London, 2014

[2] Herland et al., "A review of data mining using big data in health informatics", *Journal of Big Data*, 2014

[3] Tsanas, M. A. Little, P. E. McSharry, "A methodology for the analysis of medical data" in *Handbook of Systems and Complexity in Health*, Springer, New York, 2013

[4] Karthik Suresh, Sanjeev V. Thomas, Geetha Suresh, "Design, data analysis and sampling techniques for clinical research", *Ann Indian Acad Neurol*, 14(4), pp. 287-290, 2011

[5] O. V. Marukhina, E. E.Mokina, E. V. Berestneva, "Using Data Mining for revealing hidden regularities in the task of analyzing medical data" in *Fundamental Research*, Vol. 4, pp. 107-113, 2015 (in Russian)

[6] Thomas Luechtefeld et al., "Analysis of Public Oral Toxicity Data from REACH Registrations 2008-2014", *Alternatives to Animal Experimentation*: ALTEX, Vol. 33 (2), pp. 111-122, 2016

[7] P. Saldan, A. A. Ushakov, T. N. Karpova, "The analysis of the situation on chemical etiology acute poisonings in the Altai Region administrative center city Barnaul in 1997-2013", *Fundamental and application-oriented aspects of risk analysis for the population health*: Proc. of the All-Russian scientific and practical Internet-conference of young scientists. Perm: Book format, pp. 121-128, 2014 (in Russian)

[8] Joachim Gruettner, Thomas Walter, Siegfried Lang, Miriam Reichert, Stephan Haas, "Risk Assessment in Patients with Acute Alcohol Intoxication" in *Vivo*, Vol. 29, No. 1, pp. 123-127, 2015

[9] Ross A., Willson V. L., "Descriptive Statistics" in: *Basic and Advanced Statistical Tests*, SensePublishers, Rotterdam, 2017

[10] Igual L., Seguí S., "Descriptive Statistics" in: *Introduction to Data Science*, Undergraduate Topics in Computer Science, Springer, 2017

[11] Le Roux et al., "Inductive Data Analysis" in: *Geometric Data Analysis*, Springer, Dordrecht, 2005

[12] I. Kobzdar, *Applied mathematics and statistics*, Moscow: FIZMATLIT, 2012 (in Russian)

[13] Kirch W. (eds), "Canonical Correlation Analysis" in: *Encyclopedia of Public Health*, Springer, Dordrecht, 2008

[14] Igual L., Seguí S., "Regression Analysis" In: *Introduction to Data Science*. Undergraduate Topics in Computer Science, Springer, 2017

[15] K. Adachi, *Matrix-Based Introduction to Multivariate Data Analysis*, Springer Nature Singapore Pte Ltd., 2016

[16] E. Filatov, *Methods of the determined factor analysis*, LAP Lambert Academic Publishing, 2012

[17] H. K. Ramakrishna, *Medical Statistics*, Springer Science+Business Media Singapore, 2017

[18] T. J. Cleophas, A. H. Zwinderman, *Machine Learning in Medicine - a Complete Overview*, Springer International Publishing Switzerland, 2015

[19] Bonet et al. "Clustering of Metagenomic Data by Combining Different Distance Functions", *Acta Polytechnica Hungarica*, Vol. 14, No. 3, pp. 223-236, 2017

[20] G. Gosztolya et al. "Application of Fuzzy and Possibilistic c-Means Clustering Models in Blind Speaker Clustering", *Acta Polytechnica Hungarica*, Vol. 12, No. 7, pp. 41-56, 2015

[21] Jaeger D., Jung R. *Encyclopedia of Computational Neuroscience*, Springer, New York, NY, 2015

[22] José Neves, Adriana Cunha et al., "Artificial Neural Networks in Diagnosis of Liver Diseases", *Proc. of the 6th International Conference on Information Technology in Bio- and Medical Informatics ITBAM 2015*, Valencia, Spain, pp. 71-80, 2015

[23] H-Y Li et al. "Medical Sample Classifier Design Using Fuzzy Cerebellar Model Neural Networks", *Acta Polytechnica Hungarica*, Vol. 13, No. 6, pp. 7-24, 2016

The New Combined Method of the Generation of a 3d Dense Map of Evironment based on History of Camera Positions and the Robot's Movements

Aleksander Vokhmintcev[1,2], Michail Timchenko[3]

[1] "Intelligent information technologies" Research laboratory, Chelyabinsk State University, 129 Br. Kashirinyh, 454001, Chelyabinsk, Russia, vav@csu.ru

[2]Institute of Digital economy, Ugra State University, 16 Chekhova street, 628012, Khanty-Mansiysk, Russia vav@ugrasu.ru

[3]School of Electrical Engineering and Computer Science, South Ural State University, 76 Lenin prospect, 454080, Chelyabinsk, Russia timchenkoms@susu.ru

Abstract: A combined solution is proposed for solving the point-plane variational problem in a closed-form based on fusion visual features and 3d point clouds. An accurate method for reconstructing 3d scene is introduced, and the closed-form solutions of registration task for orthogonal transformation are presented. This method is used to reconstruct a three-dimensional model of the environment from a set of images and depth map. The given method is used to match and register point clouds with an arbitrary spatial resolution and an arbitrary scale with respect to each other. The suggested method improves the convergence and accuracy of reconstruction methods for dynamic, bulky scenes.

Keywords: 3d mapping; registration problem; three-dimensional map; feature-matching; iterative closest point algorithm

1 Introduction

In order to promptly involve in the development of oil reserves of the YNAO in the conditions of high laboriousness of mapping filtering channels and large amounts of seismic data, it is necessary to develop information technology for mapping filtering channels based on machine learning methods and artificial intelligence. In this paper, we consider an approach that allows us to automate the procedure for selecting fracture zones in 2D images and 3D models and a present new method for the identification of neotectonic structures from digital maps of

the terrain based on a 3d reconstruction methods of environment and SLAM (Simultaneous Location and Mapping). This method is used for reconstruction of seismic tracks in a three-dimensional seismic cube. We will use the suggested approach to monitor the mineral deposits in the region using a mobile robot platform. The development of an adaptive dynamic system for robot mapping and navigation in real-time is one of the key tasks in modern robotics and machine vision. Over the last decade numerous methods of SLAM have been suggested [1]. The fast methods of SLAM and the methods based on orienting points for localization in an unknown environment were the first to be worked out. Subsequently, the main direction of investigations was connected with the development of intellectual SLAM methods which use various multi-sensors and Kalman filtering to estimate the motion trajectory of a robotic system in unknown space [2]. To date, the following solutions to the SLAM problem are known: graphic SLAM, particle filter SLAM, Extended Kalman filter (EKF), and visual SLAM [3]. In addition to these approaches of the SLAM the following methods can be used: extreme methods of navigation based on comparison of two consecutive scans by optimizing the cross-correlation function; the method of recursive filtering; method Normal Distributions Transform is a transformation of normal distributions. The disadvantages of this algorithm should include the strong dependence of convergence from the initial approximations and discretization of space.

The performance of SLAM methods directly depends on the accuracy of the 3d mapping method which allows to create shape model for example based on the registration task of the 3d point clouds obtained from depth sensors. The most popular registration method is the Iterative Closest Point (ICP) algorithm [4]. A key stage in the ICP algorithm is a search for an orthogonal or affine transformation that imposes two point clouds with a given matching of points in the best mean. The most commonly used methods for searching correspondence between a pair of 3d clouds are ICP algorithm with point-to-point metric and ICP algorithm with point-to-plane metric: the second metric is superior to the first metric in terms of accuracy and convergence [5]. The point-to-point task can be solved by applying the Levenberg–Marquardt algorithm. It has typical shortcomings of iterative methods, namely, the accuracy of the registration depends on the chosen initial approximation [6]. In the class of orthogonal transformations, a closed-form solution of the point-to-point task was produced with the help of quaternions in or orthogonal matrices in [7]. Horn's method was used in [8] to formulate an ICP algorithm in the point-to-plane version. For the variational problem of the point-plane for the class of affine transformations, the exact solution was proposed [9]. An approximate solution of the point-plane problem for the class of orthogonal transformations was obtained in [10]. A registration algorithm ICP of point clouds of no rigid objects was proposed [11]. The ICP algorithm is characterized by two main problems: first, the algorithm does not use the local surface shape around each point, and secondly, the search of nearby points has a high computational complexity [12, 13]. The approach based

on point-to-plane has more constraints of the structure of the reconstructed scene [14]. The ICP algorithm using the point-to-plane metric has a poor convergence for scenes with a small number of geometric constraints [15]. The convergence of the ICP algorithm may be considerably improved [16]. Since modern RGB-D cameras provide color and depth information, a data fusion from two sources are desirable. We suggest applying visual (color) features to significantly improve the initial point of the ICP algorithm, and the alignment between key-frames is computed by joint optimization of appearance and shape matches. [17]. The most successful feature descriptors such as SIFT (Scale Invariant Feature Transform) [18] and SURF (Speeded-Up Robust Features) [19] are robust to scale and viewpoint changes. However, other visual features could be also useful for localization and global optimization. For example, histograms of oriented gradients (HOGs) possess attractive invariance to viewpoint changes [20]. Within the framework of this project a closed-form solution of the point-plane problem for the class of orthogonal transformations will be found, and an iterative registration method was suggested. The fusion ICP method is used for reconstruction of seismic tracks in a three-dimensional seismic cube.

The paper is organized as follows. In Section 2, the proposed effective fusion algorithm for registration 3d data is presented. In Section 3, the robot mapping algorithm is considered. Computer simulation results are provided in Section 4. Section 5 summarizes our conclusions.

2 Fusion Iterative Closest Point Algorithm

This project has two main objectives:

- the development of new accurate registration methods for reconstruction of three-dimensional maps in an unknown dynamic environment;
- the development of a new path planning method of a mobile robot in an unknown dynamic space which have theoretically approved estimates of computational complexity.

Modern practical solutions of a SLAM task for large-scale scenes are based on metric-topological and visual approaches. In this case, the SLAM task can be described in the form of the following scheme:

- the local reconstruction of a map using metrics to assess a trajectory and the creation of local maps;
- the recognition of places in order to solve the problem of "loop closure" or to fulfill the task of relocalization;
- the optimization of a constructed map and global optimization.

Let us define some requirements for input data and for objects of a scene:

- the robot determines its position in an environment relative to the target point;
- objects on a scene have a rigid shape (rigid objects);
- the degree of freedom of a robot' movement is 6D;
- it is necessary to consider constructive and dynamic limitations of the robot for the path planning;
- no odometrical information is used;
- no plan of the environment for a mobile robot is preset;
- the environment contains obstacles, including dynamic ones;
- the accuracy of the evaluation of a trajectory is determined by a type of sensors employed.

The algorithm for constructing a three-dimensional map of a surface consists of the following basic steps:

Step 1. Measure a depth map and compute a normalized surface.

Step 2. Estimate the camera position by applying the ICP method between the predicted and measured surfaces.

Step 3. Refine the global map of the surface by incorporating the measured surface into the map constructed at the preceding steps.

Step 4. Predict the surface at the next step by applying modified Kalman filtering. Go to Step 2.

A scheme of the proposed method is presented on Fig. 1.

Suppose that $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_m\}$ be an initial point cloud and a target point cloud in R^3 accordingly. Then each iteration in the variational problem of the classical ICP algorithm for an arbitrary affine transformation can be represented as follows:

$$RSx_i + T, \tag{1}$$

where R is a rotation matrix, T is a translation vector, $i = 1, \ldots, n.$, S is a scaling matrix.

The matching function becomes [5]:

$$E(R,S,T) = \frac{1}{N}\sum_{i=1}^{N} \left\| (RSx_i + T) - y_i \right\|^2, \tag{2}$$

where l_m is the number of corresponding points pairs.

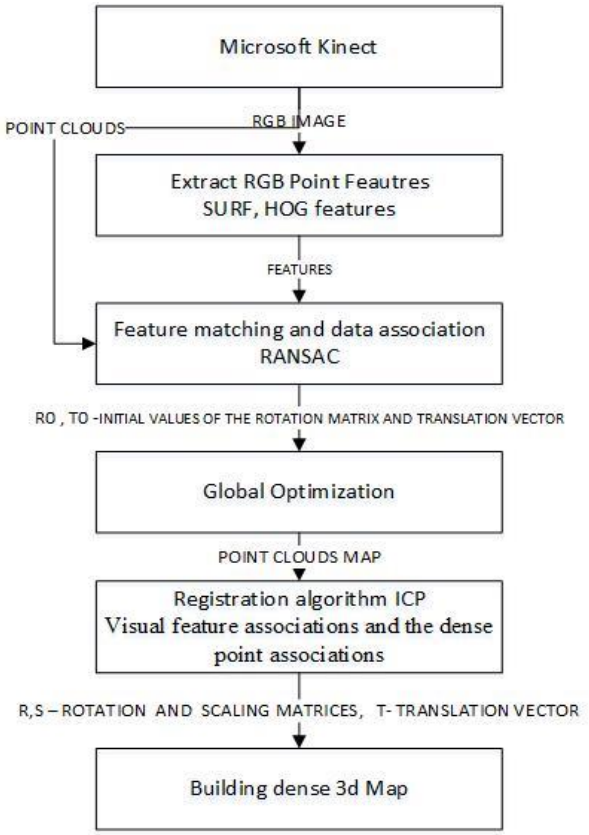


Figure 1
A scheme of the fusion registration method combining visual features and depth information

If (2) has minimum it will satisfy the following equation $\frac{dE(T)}{dT} = 0$.

$$\frac{dE(T)}{dT} = \frac{2S}{N}\sum_{i=1}^{N}\left(\left(RSx_i + t\right) - y_i\right) = 0 \tag{3}$$

So, we can make the following conclusion $T = \frac{1}{N}\sum_{i=1}^{N} y_i - \frac{1}{N}\sum_{i=1}^{N} RSx_i$. Therefore, matching function is as follows in order to achieve the minimum

$$E(R,S) = \frac{1}{N}\sum_{i=1}^{N}\left\| RS\left(x_i - \frac{1}{N}\sum_{i=1}^{N} x_i\right) - \left(y_i - \frac{1}{N}\sum_{i=1}^{N} y_i\right)\right\|^2. \tag{4}$$

Let $a_i = \left(x_i - \frac{1}{N}\sum_{i=1}^{N} x_i \right), b_i = \left(y_i - \frac{1}{N}\sum_{i=1}^{N} y_i \right)$, taking into account this equation, the matching function can be represented as follows:

$$E(R,S) = \frac{1}{N}\sum_{i=1}^{N} \left\| RSa_i - b_i \right\|^2 . \quad (5)$$

Then

$$E(R,S) = \frac{1}{l_m}\left(\sum_{i=1}^{l_m} a_i^T S^2 a_i - 2\sum_{i=1}^{l_m} b_i^T RSa_i + \sum_{i=1}^{l_m} b_i^T b_i \right). \quad (6)$$

Then we find out the minimum of this matching function:

$$\frac{\mathrm{dE(R,S)}}{\mathrm{dR}} = 0 \quad , \quad \frac{\mathrm{dE(R,S)}}{\mathrm{dS}} = 0 \quad . \quad (7)$$

Then we can get scale matrix from (7):

$$\frac{dE(R,S)}{dS} = 2\sum_{i=1}^{l_m} a_i^T SE_J a_i - 2\sum_{i=1}^{l_m} b_i^T RSE_J a_i = 0 \quad , \quad (8)$$

where $E_j = (0,\cdots,0,1,0,\cdots,0), j = 1,2,3$ is s-diagonal scale matrix. Then the scale matrix can be computed as follows:

$$S_j = \frac{\sum_{i=1}^{l_m} b_i^T R E_j a_i}{\sum_{i=1}^{l_m} a_i^T R E_j a_i} . \quad (9)$$

Let A is the transformation matrix including the components of the rotation matrix R; and T is the translation vector. Let *J(A,T)* be the following function:

$$J(A,T) = \sum_{i=1}^{n} \| A x_i + T - y_i \|^2 . \quad (10)$$

Then new project solutions to 3d total variation regularization will be obtained [21]. The function *J(A,T)* can be rewritten:

$$J(A) = \sum_{i=1}^{n} \left(< A p_i - q_i, n_i > \right)^2 = \sum_{i=1}^{n} \left(< A p_i, n_i > - < q_i, n_i > \right)^2 = \quad (11)$$

$$= \sum_{i=1}^{n} < A p_i, n_i >^2 - 2 < A p_i, n_i >< q_i, n_i > + < q_i, n_i >^2$$

where p_i is a point from the point cloud P, n_i is the unitary normal for the tangent plane $T(q_i)$.

Let us consider the partial derivatives with respect to a_{ij}:

$$\frac{\partial J}{\partial a_{ij}} J\left(A,T\right) = \sum_{k=1}^{n} 2 < A\, p^k, n^k > p_j^k, n_i^k - 2 p_j^k, n_i^k < q^k, n^k >= 0 , \tag{12}$$

where $i,j = 1,\ldots,3$, $k = 1,\ldots,n$.

$$\sum_{i=1}^{n} < A\, p^k, n^k > p_j^k, n_i^k - p_j^k, n_i^k < q^k, n^k >= \sum_{i=1}^{n} < A\, p^k, n^k > p_j^k, n_i^k - \sum_{i=1}^{n} p_j^k, n_i^k < q^k, n^k >= 0 \tag{13}$$

The partial sum $\sum_{k=1}^{n} p_j^k, n_i^k < q^k, n^k >$ is a constant with respect to A:

$$\sum_{k=1}^{n} p_j^k, n_i^k < q^k, n^k >= c_{ij}, i, j = 1,...,3 . \tag{14}$$

Taking into account Equations 14 can be rewritten as

$$\sum_{k=1}^{n} p_j^k, n_i^k tr\left(A \square (\mathrm{PN})^k\right) = c_{ij}, i, j = 1,...,3 . \tag{15}$$

Let $\hat{M}$ is the transformation matrix including the components of the rotation matrix R; and T is the translation vector. In the first iteration, a matrix $\hat{M}$ that is generated by the matrix A and vector T can be initialized by the visual RANSAC transformation. In the proposed formulation of the problem we find a solution of variation problem based on the combination of data on feature points (the color scenes), and a dense three-dimensional point cloud (depth data). The considered functionals are composed of terms that measure the average of the squares of the distances to visual-associated characteristic points with the normalizing factor (the variation of metric characteristics of a function of two variables), terms that measure the average of the squares of the distances for a dense point cloud based on metrics point-to-plane. The solution of the variation problem obtained using various iterative methods. Next, we minimize the alignment error of both visual feature and dense point associations.

Let A_f is a set of matches consisting of pairs of singular points in the image that are related to the best transformation; similarly; A_d is a set of matches in the cloud of points associated with the best transformation; α is an empirically chosen weighting coefficient; w_i and w_j – weighting coefficients for features and depth data respectively. The joint error function is given as

$$E(R,T) = \operatorname{argmin}_{\hat{M}} \left[\begin{array}{c} \alpha \frac{1}{W} \left(\frac{1}{|A_f|} \sum_i w_i \left| \hat{M} F_s^i - F_d^i \right|^2 \right) + \\ (1-\alpha) \frac{1}{W} \left(\frac{1}{|A_d|} \sum_j w_j \left| \left(\left(\hat{M} p_j - q_j \right) n_j \right) \right|^2 \right) \end{array} \right]. \quad (16)$$

The functions F^I_S and F^I_d provide the projection of a feature point from its Euclidean 3d position into the image space of the camera [2, 22].

3 Robot Mapping Algorithm

In this project, we use approach to the implementation of the task is based on the application of Extended Kalman Filter, EKF. We use the term visual feature and term semantic landmark synonymously. The robot platform is equipped with a multi-sensor cameras (Kinect and Beward camera) which take measurements of the relative location between any semantic landmark and the robot platform itself. The absolute locations of the semantic tags are not available. The implementation of the suggested visual SLAM algorithm uses nonlinear kinetic models of robot platform and non-linear asynchronous observation models in controlled/ uncontrolled indoor conditions). The state of the system of interest consists of the position and orientation of the robot platform together with the position of all semantic tags. The general algorithm of 3d mapping is shown in Fig. 2.

The updating of a history of camera positions and the robot's movements, as well as the configuration of surveyed characteristics are carried out at a certain interval of time. As a rule, data is added not at every step, but in case the position of the camera (the robot) has significantly changed, for example, when a considerable displacement or rotation has taken place as compared to the previous memorized position. A mathematical model of the scene was saved in the form of a graph the vertices of which correspond to certain moments of time. In the present work, a new combined adaptive method was developed for the generating a three-dimensional combined dense map of an accessible environment and determining the position of the robot in a relative coordinate system based on the history of camera positions, on symbolic (semantic) tags, on the robot's movements and on the matching of obtained three-dimensional depth maps which account for the accuracy of their superimposition, as well as geometric relationships between various images of one scene [23].

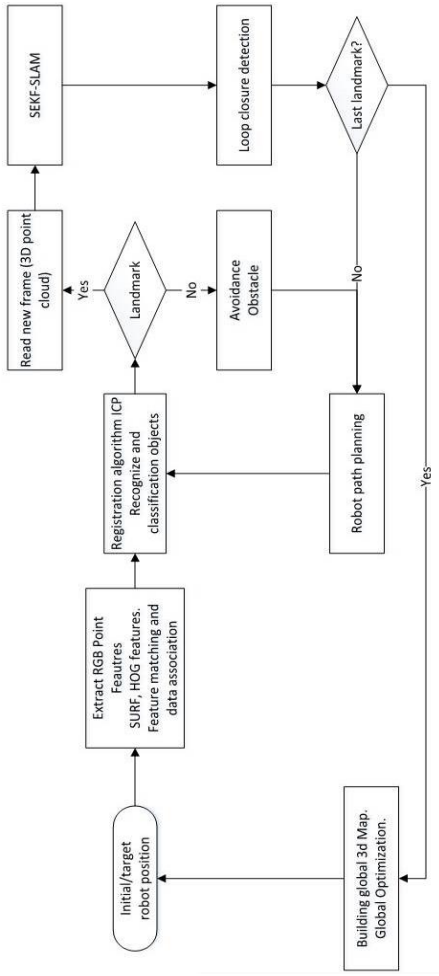


Figure 2
A block diagram of the proposed 3d mapping method

4 Computer Simulation

In this section, computer simulation results are presented and discussed. We used NYU Depth Dataset for computer simulation. In this section, the proposed fusion registration method is compared with the Fast ICP and point-to-plane ICP algorithms. The performance of the registration method is shown in Table 1 compared with that of the tested algorithms. The second row shows the

performance of the algorithms in terms of computational complexity. Each row has two values: first value show matching accuracy (%), the second value shows time in sec. Fig. 3 shows the performance of the tested algorithms in terms of mean square error. One can observe that the fusion registration method has the best performance.

Table 1

Matching accuracy (%) and computational complexity (sec) depending on rotation angle

Algorithms	Rotation angle (degrees)					
	5	10	15	20	25	30
Fast ICP	91/0,06	80/0,1	73/0,11	66/0,13	56/0,15	47/0,17
ICP	97/1,56	95/1,86	91/2,34	86/3,8	81/4,17	78/6,78
Fusion algorithm	98/1,45	99/1,78	98/2,4	96/2,95	91/3,2	92/3,63

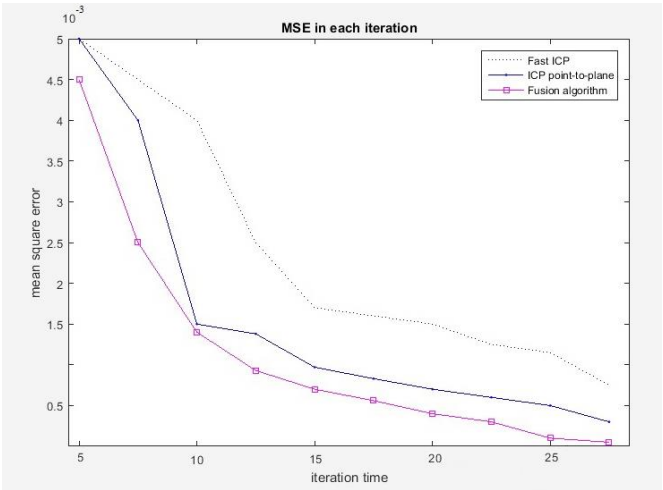


Figure 3

The performance of the algorithms in terms of mean square error

To process visual scene characteristics, we use an image-matching algorithm based on the recursive calculation of oriented gradient histograms for several circular sliding windows and a pyramidal image decomposition [20, 24]. The algorithm yields initial values for a fusion registration algorithm. State of art image matching algorithms (SIFT, SURF, ORB) and suggested algorithm were tested in a series of different experiments:

- transformation "image rotation in the scene plane" (Table 2);
- transformation "rotation outside the scene plane" (Table 3);
- conversion of scale changes (Table 4);

The matching algorithms were probed in different conditions such as in-plane/out-of-plane rotations and scaling changes. The performance matching algorithms were estimated in terms of the number of accurate matches. The suggested algorithm gives a more stable matching performance than other algorithms for out-of-plane rotations (Table 3) and small scaling (Table 4). In the future, we plan to test the suggested registration method using Netherlands Dataset: A New Public Dataset for Machine Learning in Seismic Interpretation.

Table 2
Accuracy of matching (in %) of various algorithms vs.in-plane rotation angle

Matching algorithm	Angle of rotation, □					
	45	90	135	180	225	270
SIFT	100	98	99	98	97	95
SURF	74	69	74	69	74	69
ORB	87	85	83	86	85	87
Proposed algorithm	100	98	96	98	97	95

Table 3
Accuracy of matching (in %) of various algorithms vs. off-plane rotation angle

Matching algorithm	Angle of rotation, □					
	5	10	15	20	25	30
SIFT	98	91	78	64	58	47
SURF	82	77	64	55	38	29
ORB	96	84	77	61	58	54
Proposed algorithm	83	78	74	72	76	72

Table 4
Accuracy of matching (in %) of various algorithms vs. small scaling changes

Matching algorithm	Scaling factor				
	0.8X	0.9X	1.0 X	1.1X	1.2 X
SIFT	92	95	100	98	91
SURF	79	90	99	97	92
ORB	78	79	90	83	89
Proposed algorithm	84	94	100	99	91

Conclusions

In this paper, the accurate method for reconstructing 3d scene and the closed-form solutions of registration task for orthogonal transformation combining visual features and the depth information were proposed. The obtained results were illustrated with the help of computer simulation. In the proposed formulation of the problem we find a solution of variational task point-to-plane based on the

combination of data on feature points (the color scenes), and a dense three-dimensional point cloud (depth data). The considered functionals are composed of terms that measure the average of the squares of the distances to visual-associated characteristic points with the normalizing factor (the variation of metric characteristics of a function of two variables), the terms that measure the average of the squares of the distances for a dense point cloud based on a metrics point-to-plane. The given method is used to match and register a point clouds with an arbitrary spatial resolution and an arbitrary scale with respect to each other and provide accurate estimates for complex dynamic large-scale scenes. Within the framework of the project, the variational problem of the ICP algorithm is extended to a class of affine transformations. The computational experiments carried out showed a significant increase in the accuracy of the suggested registration method. The results of the project have shown that the use of an approximate 3d map of the surrounding environment significantly improves the quality of recognition and localization of objects in dynamic, bulky scenes, especially with partial or complete occlusion of an object by other scene objects.

Acknowledgement

The reported study was funded by RFBR and Yamalo-Nenets Autonomous Okrugr, project number 19-47-890001.

References

[1] C. Hertzberg, R. Wagner and O. Birbach “Experiences in building a visual slam system from open source components”, In. Proc. IEEE International Conference on Robotics and Automation, pp. 2644-2651, Shanghai, China, May 09-13, 2011

[2] P. Henry, M. Krainin and E. Herbst “RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments”, In. Proc. 12th International Symposium on Experimental Robotics, pp. 477-491, Delhi, India, Dec 18-21, 2014

[3] F. Endres, J. Hess, N. Engelhard and J. Sturm “An evaluation of the RGB-D SLAM system”, In. Proc. IEEE International Conference on Robotics and Automation, pp. 1691-1696, Taipei, Taiwan, Oct 18-22, 2012

[4] P. Besl and N. McKay “A Method for Registration of 3-D Shapes Trans”, IEEE Transaction on Pattern Analysis and Machine Intelligence, IEEE, Vol. 14(2) pp. 239-256, 1992

[5] S. Ying, J. Peng, S. Du and H. Qiao “A Scale Stretch Method Based on ICP for 3D Data Registration”, IEEE Transactions on Automation Science and Engineering, IEEE, Vol. 6(3) pp. 559-565, 2009

[6] G. Tam, Z.-Q. Cheng, Y.-K. Lai, F. Langbein, Y. Liu, D. Marshall, R. Martin, X.-F. Sun, and P. Rosin “Registration of 3D point clouds and meshes: A survey from rigid to nonrigid”, IEEE Transactions on Visualization and Computer Graphics, IEEE, Vol. 19 (7), pp. 1199-1217, 2013

[7] B. Horn B., H. Hilden and S. Negahdaripour S. "Closed-form Solution of Absolute Orientation Using Orthonormal Matrices", Journal of the Optical Society of America, Vol. 5 (7), pp. 1127-1135, 1988

[8] K. L. Low "Linear least-squares optimization for point-to-plane ICP surface registration", Technical Report TR04-004, Department of Computer Science, University of North Carolina at Chapel Hill, 2004

[9] S. Du, N. Zheng, S. Ying and J. Liu "Affine iterative closest point algorithm for point set registration", Pattern Recognition Letters, Elsevier, Vol. 31(9), pp. 791-799, 2010

[10] K. Khoshelham "Closed-form solutions for estimating a rigid motion from plane correspondences extracted from point clouds", ISPRS Journal of Photogrammetry and Remote Sensing, Elsevier, Vol. 114, pp. 78-91, 2016

[11] S. Cheng, I. Marras and S. Zafeiriou "Active nonrigid ICP algorithm", In Proc. 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, pp. 1-8, Ljubljana, Slovenia, May 4-8, 2015

[12] Y. Chen and G. Medioni "Object Modeling by Registration of Multiple Range Images", In. Proc. IEEE International Conference on Robotics and Automation, Vol. 10(3) pp. 145-155, Sacramento,USA, April 17-19, 1991

[13] C. Langis, M. Greenspan and G. Godin "The parallel Iterative closest point algorithm", In. Proc. IEEE Third International Conference 3-D Digital Imaging and Modeling, pp. 195-204, Quebec City, Canada, 28 May-1 June, 2001

[14] S. Thrun, W. Burgard and D. Fox "A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping", In. Proc. IEEE International Conference on Robotics and Automation, pp. 321-328, San Francisco, USA, April 24-28, 2000

[15] A. Segal, D. Haehnel and S. Thrun "Generalized-ICP", In. Proc. of Robotics: Science and Systems, Vol. 2(4) pp. 645-653, Seattle, USA, June 6-9, 2009

[16] S. Rusinkiewicz and M. Levoy "Efficient variants of the ICP algorithm", In Proc. International Conference on 3-D Digital Imaging and Modeling (3DIM) pp. 145-152, Quebec City, Canada, May 28-June 1, 2001

[17] M. Fischler and R. Bolles "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Graphics and Image Processing, pp. 381-395, 1981

[18] D. G. Lowe "Object recognition from local scale invariant features", In. Proc. 7th International conference on Computer Vision, Vol. 2, pp. 1150-1157, Kerkyra, Greece, September 20-27, 1999

[19] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding, Journal Elsevier, Vol. 110, pp. 346-359, 2008

[20] D. Miramontes-Jaramillo, V. Kober, V. H. Diaz-Ramirez and V. Karnaukhov V. "A novel Image Matching Algorithm Based on Sliding Histograms of Oriented Gradients", Journal of Communications Technology and Electronics, Springer, Vol. 59(12) pp. 1446-1450, 2014

[21] A. Vokhmintcev, A. Melnikov, K. Mironov, V. Burlutskiy "Reconstruction of three-dimensional map based on closed form solution of variational problem of multi-sensor data registration", Doklady Mathematics, Springer, Vol. 99 (1) pp. 108-112, 2019

[22] A. Vokhmintsev and K. Yakovlev "A Real-time Algorithm for Mobile Robot Mapping Based on Rotation invariant Descriptors and ICP", In Proc. 5th Analysis of Images, Social Networks and Texts, Communications in Computer and Information Science, Vol. 661, pp. 357-369, Ykaterinburg, Russia, April 7-9, 2017

[23] I. Sochenkov and A. Vokhmintsev "Visual duplicates image search for a non-cooperative person recognition at a distance", Procedia Engineering, Elsevier B. V., Vol. 129, pp. 440-445, 2015

[24] A. Vokhmintcev, I. Sochenkov, V. Kuznetsov V. and D. Tikhonkikh "Face recognition based on matching algorithm with recursive calculation of local oriented gradient histogram", Doklady Mathematics, Springer, Vol. 93(1) pp. 37-41, 2016

Analysis of the GRID's Basic Topological Structure Reliability

Vladimir E. Gvozdev, Lilia R. Chernyakhovskaya, Darya V. Blinova, Dinara R. Akhmetova

Department of Computer Science and Robotics, Ufa State Aviation Technical University, 12 K. Marks St, 450007 Ufa, Russia
e-mail: gvozdev.ve@net.ugatu.su, chernyahovskaya.lr@net.ugatu.su, blinova.dv@net.ugatu.su, ahmetova.dinara-r@net.ugatu.su

Abstract: Approaches to the investigation of the reliability of the basic topological structures of GRID systems, which are minimal enough set of constructive elements for constructing such systems, are considered in this work. Methods used for the quantitative assessment of the reliability of basic topological structures under conditions of different interpretations of the concept of "full failure of the basic structure" are considered. The construction of structural reliability models, based on series-parallel reliability schemes, are described in this work. The application of these approaches creates the basis for a multi-dimensional analysis of the GRID-systems reliability characteristics, at the design and development stages. This, in turn, creates the prerequisites for assessing the functional safety of such systems.

Keywords: GRID; topology; reliability; failure

1 Introduction

The tendency to improve the efficiency of computing and storage resources has served as the basis for the emergence of a focus in the field of information technologies, especially, GRID systems.

In the work [1], with references to other literature [2, 3], three key indicators of the effectiveness of GRID-systems are listed, including reliability. The problem of GRID systems reliability management is a priority for world-wide research efforts. There exists the need of development of methodological, methodical bases and tools of reliability management, at different stages system's components life cycle.

The main feature of GRID-systems is the possibility of redistribution of computing resources and resources, for data storage, it follows that the reliability

analysis of such systems depends on the reliability of information processing nodes and the reliability of communication facilities [4-6].

Reliability is the characteristic of the systems, retaining in time, within the established limits and values for all parameters, that characterize the ability to perform the required functions in the specified modes and conditions of use.

In the works [1, 7], basic topologies are distinguished, called by the author, the basic fractal architectures as follows:

- Complete cellular topology
- Ring topology
- Topology "star"
- Linear topology
- Mixed topology

Allocated components of GRID systems play the role of typical aggregates, that consist of two types of elements: data processing (storage) unit and equipment that provides communication of nodes. In the mentioned work it is emphasized, that the selected aggregates are the basic constructions from which GRID-systems of any complexity can be built. Due to this, the task of GRID-systems aggregates reliability, can be singled out as an independent task. In the work [1], an approach to assessing the reliability of basic topological structures (BTS) is described. This approach based on the analysis of graphs connectivity, where nodes are computers, workstations, concentrators, and edges are communication channels.

In this paper approaches to quantifying the reliability of basic topological structures are considered. In the present work, an approach to quantitative assessment of the basic topological structures, sensitivity indicators to the change in the reliability of nodes and connections under various definitions of the concept of "failure of BTS GRID systems" on the basis of the well-known apparatus of series-parallel logic circuits [8]. Building the reliability models of BTS is the basis for assessing the reliability of GRID systems at the design and development stage, which, creates prerequisites for assessing the functional safety of such systems.

2 Characterization of GRID Systems Components Reliability Studies

Analysis of published sources allows for identification of the following areas of research, related to managing the reliability of the components of GRID-systems:

- Methods for software testing (components, interfaces, stress, security, system, etc.), verification and validation. A significant number of publications, manuals and standards are devoted to this theme [9, 10, 11]
- Development of theoretical and methodical bases of the analysis of reliability and safety of info-communication systems of various scale and complexity [1, 12-16]
- Development of theoretical basis and tools for designing of error-steady systems [17-21]
- Development of theoretical foundations, manuals and decision support tools at the design stage according to the reliability criteria [10, 22-24]
- Development of methodological, methodical and model bases of a new scientific direction of software defectology [15, 25]

Analysis of the references allows to make a conclusion, that the task of managing reliability of GRID-systems belongs to the class of tasks of managing complex systems.

3 The Approaches to Assess the Reliability of Basic Topological Structures of GRID-Systems

The basis of the analysis of the reliability of complex technical systems, an integral component of which is software, is the concept of "failure." In the well-known literature [1] it is shown that the concept of “failure” is treated differently depending on the functional architecture of the systems [10], the point of view of the target users group [26], the system’s conditions of functioning [27]. The interpretation of the concept of "failure of a part of the GRID system" by stakeholders and target groups of users is based on their subjective idea of the desired consumer properties of the systems [25].

As an independent problem of reliability of technical systems, the problem of reliability of the elements making up this system is singled out in the classical theory of reliability. In the framework of this study, the basic topological structures are considered as a minimal sufficient set of aggregates from which a GRID system can be built. Taking into account that GRID-systems are related to complex systems, it is possible to adapt the provisions of the reliability theory for the reliability analysis of GRID.

In the theory of reliability, approaches to the study of repaired and non-repaired systems are distinguished. In the construction of the models presented below, postulated the independence of failures of nodes (computers, workstations, routers, communication devices) and edges (communication channels) and the

provision for unrepairable components. This assumption is valid, firstly, when the emphasis in research is on the hardware component of the BTS, and secondly, when the current repair of a technical object is technically impossible at the site of its operation or is economically inexpedient. In the framework of this paper, it is assumed that the element can be either in the working or in the inoperative state, the situation of partial loss of efficiency is not considered.

The approach to estimate the reliability of non-repaired systems using serial-parallel logic schemes is known. Below are given examples of structural-logical schemes for a linear basic architecture, corresponding to the case when it consists of four nodes.

4 Evaluation of Reliability of the Basic Topological Structure "Linear" depending on the Content of the "System Failure" Concept

4.1 Evaluation of the Basic Topological Structure Reliability by the Criterion of the Isolation of the Components of the Graph

Described in this section models are based on the following interpretation of the concept of "full failure": a complete disruption of the GRID network (all nodes of this network are isolated) due to non-repairable failure of nodes (computers, workstations, routers, communication devices) or edges. Taking into account the main feature of GRID-systems, this designation is formulated as follows: There was a full failure of the system, if there is no any possibility of reallocation of resources.

In accordance with this interpretation of the concept of failure, a serial-parallel scheme is constructed to evaluate the reliability of a typical aggregate of the basic architecture "linear" consisting of 4 nodes (Figs. 1-2). In this case, the communication channel between all the elements is a bus.

This paper considers the characteristic of BTS reliability as a time between failures and also assumes that all blocks on the above structural serial-parallel scheme have the same value of probability p.

Figure 1
Logical structure "linear"

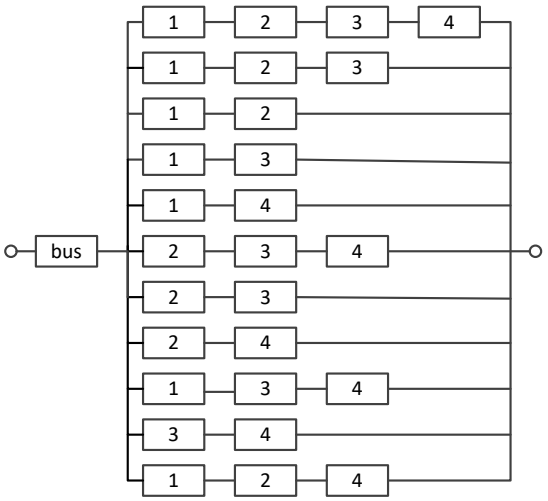


Figure 2
Serial-parallel scheme of reliability (all nodes are isolated)

Proceeding from the aforementioned, for the basic topological structure "linear" consisting of four nodes, the following relation can be put in correspondence for calculating the reliability of the system:

$$p_{BTS} = p_{bus} * (1-(1-p^4) * (1-p^3)^4 * (1-p^2)^6) \quad (1)$$

Analogous dependencies were obtained for the topological structure "linear" consisting of five and six nodes, the plots of the dependence of p_{BTS} on the value of p are shown in Fig. 3.

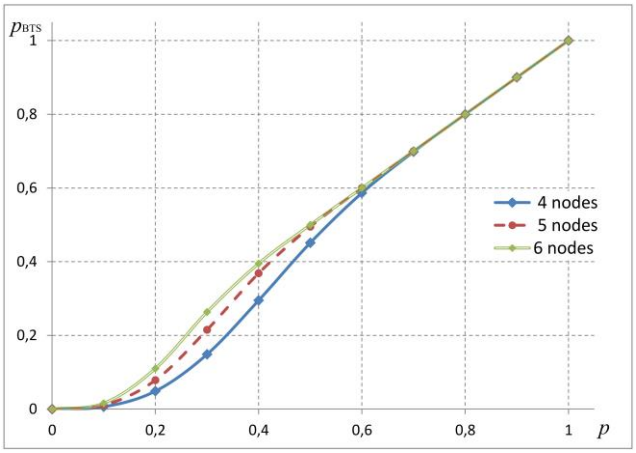


Figure 3
Dependence the system's reliability on the elements reliability (all nodes are isolated)

From the dependencies shown in Figure 3 it follows that within the framework of the introduced "full failure" concept, the reliability of the component increases with the increase in the number of nodes.

4.2 Evaluation of the Basic Topological Structure Reliability by the Criterion of the Absence of a Connection between Any Two Components of the Graph

In this section, the concept of "total failure" is interpreted as: violation of the GRID network workability (there is no connection between any two nodes) non-repairable failure of nodes (computers, workstations, routers, communication devices) or edges (communication channels). This interpretation of the concept of failure is given in [1]. A serial-parallel logic scheme on the basis of which the reliability of a linear BTS is evaluated will look like it is shown in Fig. 4. Taking into account the main feature of GRID-systems, this definition is formulated as follows: there was a full failure of the system, if there was a failure of any element, which leads to the impossibility of redistributing resources between all elements.

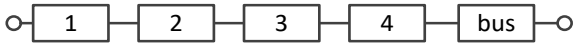


Figure 4
Serial-parallel scheme of reliability (there is no connection between any two nodes)

For the considered serial-parallel logical scheme of the BTA "linear", the following relationship for calculating the reliability of the system can be used:

$$p_{BTS} = p^5 \qquad (2)$$

The dependencies of p_{BTS} on the value of p for "linear" BTA with four, five and six nodes are given on Fig. 5.

From the dependencies shown in Figure 5 it follows that within the framework of the introduced "full failure" concept, the reliability of the component decreases with the increase in the number of nodes.

As can be seen from the given examples, even in the framework of one basic topology, the results of the construction of serial-sequential schemes differ depending on the interpretation of the concept of "full failure". This, in turn, leads to obtaining of various calculation formulas for assessing the reliability of the BTS and the various dependencies of reliability of the elements on the reliability of the unit.

Building the reliability assessments is a complex task, at the center of which is a failure, the concept of which is treated differently by each target group of users with their valuable concepts.

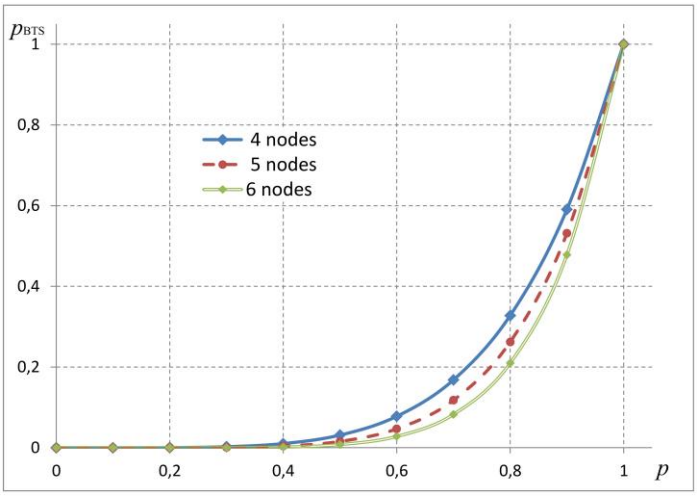


Figure 5
Dependence the system's reliability on the elements reliability
(there is no connection between any two nodes)

5 Analysis of the Reliability Sensitivity of a Basic Topological Structure "Linear" to the Reliability of Its Elements

The sensitivity characteristic allows to make an informed conclusion about how the increase in the reliability of the elements is affects the change in the reliability of the structure as a whole.

Figs. 6-7 show the sensitivity of the reliability characteristics of the system p_{BTS} from the reliability characteristics values of architecture elements, obtained on the basis of the ratio:

$$S = \frac{\Delta p_{BTS}}{\Delta p} \tag{3}$$

As can be seen from the figures, the formulation of the concept of "failure", affects a change in the reliability characteristics of the unit, depending on the reliability characteristics of its elements.

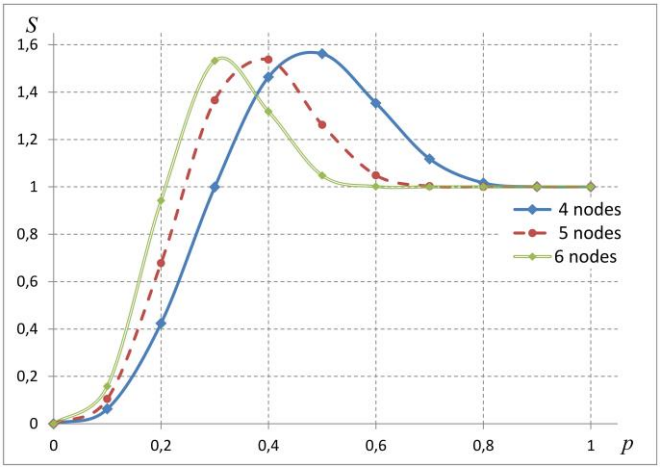


Figure 6
Sensitivity of system reliability to reliability of elements (all nodes are isolated)

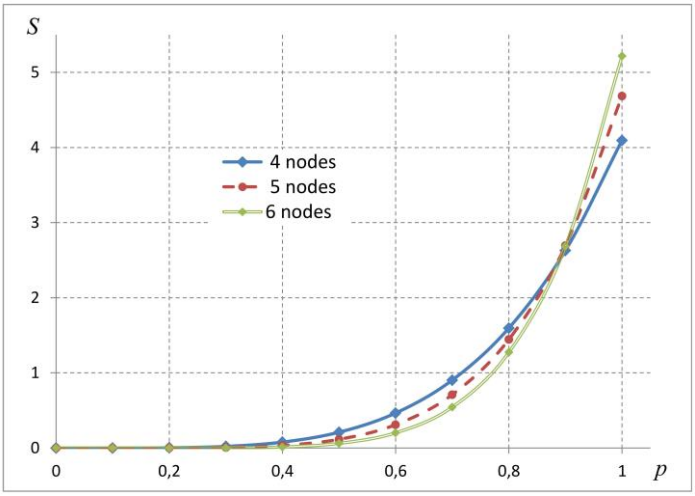


Figure 7
Sensitivity of system reliability to reliability of elements
(there is no connection between any two nodes)

Conclusion

The selection of basic topological structures makes it possible to represent any distributed GRID systems in the form of a set of standard aggregates. In this regard, as a scientific problem with applied value, it is necessary to single out the task of analyzing the reliability of the basic topological structures that make up the basis for the analysis of design solutions in the construction of GRID systems.

The GRID reliability analysis is a multi-criteria task and a system approach should be the methodological basis for its solution.

GRID systems are used to meet the information needs of different users with completely different system requirements, that affects the interpretation of the concept “failure”.

Due to the fact that the basic units are hardware-software complexes, different contents of the concept "failure" can be put in correspondence with them, which effects on the reliability characteristics of the basic topological structures and GRID-systems with different interpretation of the concept of reliability.

Depending on research objectives, it is possible to construct a set of models of failures (models of operational state) GRID. The choice of a model is determined by the specific features of the tasks within the framework of ensuring functional safety (survivability, storability, etc.).

As an example of a model that supplements the previously described models, we give the failure model of the linear topology GRID shown in Fig. 1 by the criterion "the possibility of interaction of any nodes of the system." Failure in this case will be understood as the absence of communication between any two nodes at any time due to non-recoverable failure of nodes (computers, workstations, routers, communication devices) or edges (communication channels, bus, terminators).

In this interpretation of the "failure" concept, such condition of GRID is possible, when all nodes are in working order, but are isolated from each other. In this case, it is impossible to reallocate resources, which is a failure of GRID. The failure model is shown in Fig. 8.

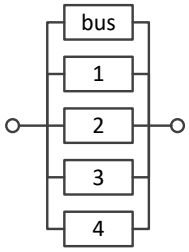


Figure 8
Serial-parallel scheme of failure (there is no connection between any two nodes)

As future research, the authors plan to develop a bank (matrix) of models for different formulations of reliability and failure, depending on the level of the GRID interaction protocols (physical, logical, applied).

Acknowledgments

This research was supported by grant 19-08-00177 of Russian Foundation for Basic Research.

References

[1] Timofeev, A. V.: Adaptive control and intellectual analysis of information flows in computer networks, Anatoliya, Moscow, 2016

[2] Halabi, S., McPherson, D.: Internet routing architectures, second edition, Cisco Press, Indianapolis, 2001

[3] Walrand, J., Varaiya, P.: High-Performance Communication Networks, 2nd Edition, Morgan Kaufmann, 1999

[4] Katic, N.: Performance analysis of smart grid solutions in distribution power systems, Acta Polytechnica Hungarica Vol. 15, No. 6, 2018, pp. 79-97

[5] Thierion, V., Ayral, P., Jacob, G.: Grid technology reliability for flash flood forecasting: end-user assessment, Journal of Grid Computing, Vol. 9, 2011, pp. 405-422

[6] Islam, T. Z., Bagchi, S., Eigenmann, R.: Reliable and efficient distributed checkpointing system for grid environments, Journal of Grid Computing, Vol. 12, 2014, pp. 593-613

[7] Timofeev, A. V., Dimitrichenko D. P.: Multi-criteria evaluation for network topological structures for GRID-systems modeling and design, Proc of SPIIRAN, 2009; Vol. 11, pp. 193-212

[8] Kapur, K., Lamberson, L.: Reliability in engineering design, Wiley India Pvt. Limited, 2009

[9] Myers G. J. "Software reliability: Principles and practices". Wiley, New York, 1976

[10] ESA PSS-05-10 Guide to software verification and validation, March 1995

[11] IEEE STD 1012 Approved Draft Standard for System, Software and Hardware Verification and Validation, 2016

[12] IEC 60300-3-1:2003 "Dependability management – Part 3-1: Application guide – Analysis techniques for dependability – Guide on methodology"

[13] IEC 61165:1995 Application of Markov techniques (MOD)

[14] Lee D. Y., Vouk M., Williams L. "Using Software Reliability Models for Security Assessment – Verification of Assumptions", In proc. of ISSRE 2013, pp. 23-24

[15] Borodakij, Yu. V., Yusupov, R. M., Palchun, B. P.: Problem of imitation modelling of computing infosphere defectoscopy features, Proc. of the 3rd Russian conf. «Imitation modelling. Theory and Practice» Sankt-Peterburg, 2007, pp. 87-92

[16] Gvozdev, V. E., Blinova, D. V., Akhmetova, D. R.: Statistical analysis of time of establishing steady phases of functioning of complex hardware-

software systems, 2018 International Russian Automation Conference (RusAutoCon), 2018, pp. 1-5

[17] Morozov, A., Janschek, K., Yusupova, N.: On the influence of control from properties to software error location, Proc. of the International Workshop "Innovation of Information Technologies – Theory and Practice", Dresden, Germany, 2010, pp. 86-91

[18] Michael, C. C., Jones, R. C.: On the uniformity of error propagation in software, Proc. of the Annual Conference on Computer Assurance (COMPASS'97), 1997, pp. 68-76

[19] Ammar, H., Nassar, D., AbdelMoez, W., Shereshevsky, M. "A framework for Experimental Error Propagation Analysis of Software Architecture Specifications". In: FastAbstract ISSRE, 2002

[20] Aguilar, J. A., Garrigós, I., Mazón, J.-N.: Requirements Engineering in the Development Process of Web Systems: A Systematic Literature Review, Acta Polytechnica Hungarica Vol. 13, No. 3, 2016, pp. 61-80

[21] Nethravathi, B., Kamalesh, V. N.: Computation of maximal connectivity for fault tolerance: graph theoretical model and algorithm, International Journal of Grid and Distributed Computing, Vol. 13, No. 1, 2020, pp. 412-417

[22] Pentti, H., Atte, H. "Failure mode and effect amalysis of software-based automatic systems". STVK-YTO-TR130, 2002

[23] Munson, J. C., Khoshgoflaar, T. M. "The detection of faultprone programs". IEEE Transactions on Software Engineering, 1992; 18(5): 423-433

[24] Kan-Hin Chai, Jun Zhang, Kay-Chuan Tan: A TRIZ-Based Method for New Service Design". Journal of Service Research, National University of Singapore, 2005; 8(1): 48-66

[25] Gvozdev, V. E., Blinova, D. V., Chernyakhovskaya, L. R.: Prevention of defects on the early stages of designing hardware-software complexes on the basis of the positions of the theory of intersubjective management, Ontology of designing, 2016, 4(22), pp. 452-464

[26] Yussupova, N., Kovács, G., Boyko, M., Bogdanova, D.: Models and methods for quality management based on artificial intelligence applications, Acta Polytechnica Hungarica, Vol. 13, No. 3, 2016, pp. 45-60

[27] Gvozdev, V. E., Guzairov, M. B., Blinova, D. V., Davlieva, A. S.: Control of component alterations according with the target efficiency of data processing and control system, CEUR Workshop Proceedings, 2017, Vol. 1903, pp. 11-16

Energy-Efficient Technologies in the Educational Programs of the Architectural Higher Education Schools

Galina B. Zakharova[1,2], Alexandr I. Krivonogov[2], Sergey V. Kruglikov[2], Alexandr A. Petunin[2]

[1]Ural State University of Architecture and Art, K. Libknekht str. 23, 620075 Ekaterinburg, Russia, zakharova@usaaa.ru

[2]Ural Federal University, Mira str. 19, 620002 Ekaterinburg, Russia {g.b.zakharova, a.i.krivonogv, s.v.kruglikov, a.a.petunin}@urfu.ru

Abstract: The world trends in modern construction correspond to the ecological, rational design, according to "green" standards and applications of the building information modeling – BIM. Herein, we give a short review and some examples of Green, BIM technologies, in different countries. For more effective implementation of "green" building, it is necessary to introduce corresponding thematic in the higher education. Based on a comprehensive analysis of the articles, we described the competencies that a University Graduate must possess, for successful implementation in the industry AEC. Despite the absence of common methods, we developed a practice-oriented approach for teaching students Green BIM technologies and here integration these technologies into the educational process is considered. Some projects of energy-efficient objects, created by students of University of Architecture and Art, meeting requirements of "green" standards, are presented.

Keywords: energy-efficient technologies; BIM; "green" standards; Green BIM; education

1 Introduction

In recent years, Russia has been paying increasing attention to environmental issues and sustainable development. The Ministry of Natural Resources and Environment of the Russian Federation has developed the strategy for the environmental safety of the Russian Federation for the period up to 2025 and the plan for its implementation. The document determines the level of environmental safety in the territory where lives the most of the Russian population at the present time as unsatisfactory. Fundamental changes in this situation, improving the quality of life and health of the population require joint actions of the state, business, public organizations and the population.

The Strategy notes that one of the important measures to overcome environmental problems is to foster environmental literacy among the young generation through teaching methods and means of "green" designs, in relevant universities.

In 2018, by decree of the President of the Russian Federation, 12 national projects were adopted in all areas of the country's strategic development, and one of these projects was Ecology. The goal of the project is to improve the environmental situation in Russia and create comfortable conditions for life in the country.

In 2009, The Russian Green Building Council (RuGBC) was created as a member of the World Green Building Council. It is a not-for-profit industry organization dedicated to accelerating development and adoption of market-based «green» building (sustainable building) practices. In accordance with its mission, RuGBC promotes the modernization of the construction industry in accordance with the principles of environmental construction through the creation of standards, methodologies, design solutions, and the union of supporters of the integrated promotion of green technologies and solutions. The instruments of «green» design are passive and active house technologies [1], traditional energy-efficient power generation, use of wind and solar energy, heat pumps, energy efficient controlled lighting, water treatment, solid waste processing, etc.

The world trends in modern construction correspond to the ecological rational design in accordance with green standards and building information modeling – BIM [2]. In a review of numerous publications on Green BIM [3], it was noted that the attention of research is focused on environmental indicators at the design stage (more than half of the articles reviewed) and the construction stage (about a third). Few studies concentrated on the development of BIM-based tools for managing environmental performance during the building maintenance, retrofitting, and demolition stages. The authors suggest that BIM for monitoring and managing environmental sustainability over the full life cycle of a building should be developed in future research. Future green BIM tools should include the concept of three R (reduction, reuse and recycle) in sustainability analysis for new development and modernization projects. It was also noted that the lack of computer tools and the complications of the BIM models are hindering the adoption of green BIM.

Based on an analysis of about 400 articles published from 1999 to 2016 and the 12 most common types of BIM software, the Survey [4] concludes that there is an urgent need for a nexus between BIM and Green Buildings. The study proposed the Green BIM Triangle taxonomy, which indicates that the relationship between BIM and green buildings should be understood based on three dimensions: project phases, green attributes, and BIM attributes. In accordance with this, the following are shown:

1) BIM applications to support the design, construction, operation and retrofitting processes of green buildings.

2) BIM functions for the analysis of green buildings, such as analysis of energy, emissions and ventilation analysis.

3) The use of BIM to support green building assessments.

4) Research gaps and future research directions in this area. This issue contributes to better alignment BIM development with green building development in the future.

In order for end-users to be involved in the building design process in the early stages to obtain adequate information, the Green2.0 online system was developed [5]. It is an open platform that leverages advances in information model building (BIM) and energy efficiency modeling tools. The system is integrated with a social network, which allows participants (end users or professionals) to share their views on building design. Social analysis and semantic modeling tools are then used to extract information from these interactions. At the same time, it connects BIM to energy analysis software, which allows users to select various products from the catalog and evaluate the impact of each of them on energy consumption. The platform is aimed at promoting the modern level of technology due to fundamental changes in the ways of joint work of AEC specialists, end users and government politicians throughout the life cycle of the building.

The authors of [6] propose the tool for assessing green building, which extracts the necessary data from BIM models to calculate the green rating and helps the development team to create documentation for obtaining a green certificate (for example, BREEAM) and provides feedback for further assessment. Article [7] assesses the applicability of BIM to optimize building sustainability assessment (BSA) methods. Although BIM is not designed to support sustainable construction, it has great potential: BSA can be simplified with BIM. The results show that using the developed application, it is possible directly and indirectly evaluate most of the criteria using BIM.

Here are a few more specific examples of the effective use of BIM and Green BIM technologies. In [8], a Green BIM case study was developed and Nordic experience was discussed as a source of good practice regarding the implementation of Green BIM. The experience of Finland, which is now the undisputed leader in innovation, is noted. One of the leading investments in Sweden, in Stockholm, implemented in accordance with the BIM concept and sustainability criteria – the largest shopping center in Scandinavia is mentioned. BIM and sustainability are two symbiotic forces that currently span the construction industry around the world. Study [9], using the design of the Hungarian National Sports Center as an example, shows the great potential for dynamic modeling of climate and energy in buildings. A set of calculations was performed, while energy modeling and architectural planning were performed synchronously. Study [10] contains a detailed analysis of the application of BIM technology, energy modeling and life cycle assessment in a project of the Helsinki Music Center.

The article [11] made a fundamental review of more than 1,500 BIM publications that have been published over the past 25 years by BIM researchers from 65 countries. The growth dynamics of publications, the leading countries (USA, UK, China, Australia), the main trends in global research areas are shown. Of the 12 BIM fields, 3 areas have the most significant contribution: process simulation and monitoring (22%), building information services (16%) and standardization (14%). In our context, we note that the direction of Education and training is noticeable - 8%.

An important role in ensuring the further more effective development of green technologies in construction is assigned to the education system. Fostering environmental responsibility among youth through teaching Green BIM design methods and tools at architectural and construction universities will improve the environment in the future and ensure a transition to an environmentally-friendly model of economic development. In higher education, there are currently no common standards and approaches. Moreover, most universities in Russia are at the initial stage of introducing Green BIM technologies into the educational process. For the wide spread of the Green BIM technologies, it is necessary to develop educational standards for teaching in universities. But this process is not fast, therefore an architectural and construction universities should find their own forms and methods for the rapid introduction of modern technologies in the educational process.

Aware of all the responsibility to future generations, we have continuously introduced a number of practice-oriented methods of teaching students at the Ural State University of Architecture and Art (USUAA) under the unique interdisciplinary educational program "Applied Informatics in Architecture." Next, we will describe the techniques that we developed and applied in the process of teaching students majoring in computer science-architect. Our graduates, who have the most modern knowledge in the field of environmental design and information modeling, organically introduce these technologies into the practice of building facilities that meet the principles of sustainable development. We will also present a number of projects carried out by students of the Department of Applied Informatics USUAA using BIM and Green BIM technologies and meeting the requirements of "green" standards.

Before that, we present the results of our study on the subject of what competencies a graduate of a modern architectural and construction university should have in order to successfully integrate into a new dynamically developing digital environment and what methods and approaches to teaching have been developed in foreign and Russian universities.

2 Overview of BIM Technologies Teaching Methods

A modern BIM specialist, based on the main property of the BIM model, focused on teamwork in a common information space, should have the main quality - the ability to solve professional problems. Nevertheless, soft skills are equally important, such as, people management, coordination skills, emotional intelligence, focus on customers and negotiation skills. This is part of the list of key competencies of the future for 2020, which was compiled at the Davos World Economic Forum in 2016 as part of the discussion "Competences of the future: what to learn and how to teach." Supplemented by personal qualities, such as analytical abilities, initiative, responsibility, sociability, motivation and focus on effective work, this model of BIM specialist meets modern requirements.

BIM makes fundamental changes in the architectural, engineering and construction industries and therefore affects the requirements for educational process in the relevant areas. Foreign sources devoted to various aspects of BIM show significant experience in teaching, which corresponds to the mature level of BIM development in Europe, USA, Singapore, Australia, etc.

The review [12] systematically presents research methods, data collection, information about disciplines, and educational literature on BIM. Such approaches and teaching methods as joint and active are described; the most popular are project and problem methods. Realistic modeling of projects is used, the nature of BIM as a way of working together pushed the teachers of AEC to implement interdisciplinary models. The conceptual classification of the efforts of a BIM teacher and researchers in higher education systems is given. Six categories of tasks have been developed that reflect a systematic approach to the implementation of BIM education: determining the needs for BIM in universities; identification of basic skills for teaching BIM; development of educational structures BIM; curriculum development; experiments with BIM courses; development of strategies to overcome educational problems.

Non-technological skills of a BIM specialist are studied in the article [13]. It is shown that problems associated with people and processes can impede the success of BIM even more than the technology itself. This points to a new additional set of BIM skills that teachers need to develop in preparing students for a successful future career. Based on an analysis of the literature on non-BIM skills, it is shown how problem-based learning can improve these types of skills and to what extent it can be beneficial in applying BIM. The article uses an interesting abbreviation for assessing the quality of students' answers to poorly structured and open problems, which are usually included in problem education – S.M.A.R.T. (specific, measurable, assignable, realistic, time-based).

Requirements for a BIM specialist on the part of customer companies in the UK construction industry are described in [14]. Customer demand is recognized as a significant motivation and encourages the construction industry to begin the

transformation and implementation of BIM. Based on numerous case studies and data collection through interviews, several types of competencies are identified as critical success factors. Any BIM-based project should start with EIR (Employer Information Requirement), customer information requirements. EIR should be designed in accordance with the supply chain throughout the project life cycle. According to UK government standards, customers must develop their requirements in three main areas: technical, managerial and commercial.

The study [15] analyzed and compared BIM maturity models. The results showed that there is no holistic model that includes definitions of processes that encompass the entire life cycle of an object and contain measures to evaluate all AEC/FM (facility management) processes. A reference model has been developed to evaluate BIM capabilities in AEC/FM processes, which has been developed iteratively through expert reviews and preliminary research. The results showed that the model is able to identify the BIM capabilities of various AEC/FM processes.

The experience of developing countries [16], which is shown by the example of the Nigerian construction industry, is interesting. As a result of this study, it was found that there is a great need for adequate training of professional specialists in the skills and competencies of using BIM. Professional organizations can apply training methods such as organizing a regular seminar and conference on BIM concepts and applications, creating a BIM academy where professional members can receive proper training on BIM. The educational community is exploring the best ways to integrate BIM into study programs.

In Russia in recent years, the number of publications devoted to the experience of implementing BIM in the educational process has increased significantly. For example, the Samara State University of Architecture and Civil Engineering uses the BIM methodology and the accompanying methodology for integrated project implementation in construction IPD (Integrated Project Delivery) [17]. It is proposed to model such an IPD feature as early involvement in the design and construction process of the project team and general contractor in the educational process through cross-disciplinary course design.

At the Saint-Petersburg State University of Architecture and Civil Engineering, training in collaboration in a BIM project was conducted in the optional format [18]. The course is not compulsory, therefore it allows to improve the balance in the number of participants, is out of schedule and makes it possible to assemble students of different specialties into the study group: architect, specialist in constructions, water supply, heat and gas supply and ventilation, electric networks, estimator and programmer responsible for automation processes. Specialists in techno sphere safety, organization of the construction process, etc., can be involved.

Problems and prospects in Russian universities from the perspective of a software integrator company are discussed in the article [19]. The authors offer practice-

oriented BIM education and engineering training at architectural universities. All aspects of BIM are shown throughout the entire life cycle of the capital construction project. In accordance with this, requirements and approaches to training are formulated. We draw attention to the noted importance of cooperation between universities with leaders in the construction industry for the development of software and equipment, with integrator companies with practical experience in the development and implementation of BIM.

We also implemented the introduction of new technologies in the educational process of students of the architectural university [20-22]. In the next section, our methods are described in more detail.

3 Green BIM in the Educational Program "Applied Informatics in Architecture"

3.1 Practical-oriented Methods for Teaching Green BIM

In this section, we describe a comprehensive approach that allowed us successfully teach students the principles of environmental design and information modeling using active and practice-oriented methods. This approach was applied at the USUAA on the Department of Applied Informatics for multidisciplinary specialists in the field of “Applied Informatics in Architecture” and consists of 6 sections described below.

1) *The introduction to the CAD course of modern BIM and Green BIM tools*

In the multi semester CAD discipline we were free to include a large number of modern automation tools for the architectural and construction industry. We teach them Autodesk Revit and Graphisoft ARCHICAD for creation of BIM projects, in Navisworks students learn principles of 4D и 5D modeling.

We explain students that the information model is a multidimensional concept. If it is structured as in the work of Pennsylvania University on recommendations for BIM implementation [23], then a lot of BIM scenarios (or BIM Uses) are highlighted, these are components of BIM technologies tied to the planning, design, construction and operation stages. There are 25 scenarios at this resource, some of them are: modeling of existing conditions, cost estimation, design, construction analysis, light analysis, energy analysis, mechanical analysis, LEED compliance assessment, compliance, service planning, building system analysis, planning in case of emergency.

Autodesk 3DS Max and AutoCAD packages are also used in the learning process. Students are able to import 3D models into the Unity 3D and Unreal Engine software environment to organize VR interactive project management in real time.

Projects in the technique of AR (augmented reality) are also elaborated. To work with a relief and to design infrastructure objects, AutoCAD Civil 3D is used. The development of project concepts is carried out using the InfraWorks 360 platform in order to improve the efficiency of data exchange and collaboration. Data bases such as Autodesk's Vault and CSoft's TDMS for the cataloguing and integration of construction objects and for collective project development were mastered.

2) *Discussions of the best world practices in the field of environmental design*

For interactive lessons students prepare topics including the best world examples of «green» architecture. Thus, the projects of architects who work by the criteria of sustainable development were analyzed. These architects were awarded the Pritzker Prize, which is an analogue of the Nobel Prize in the architecture. It should also be noted the international competition LafargeHolcim Awards, which is held since 2005. The jury evaluates all works according to five main criteria: a holistic, integrated view on the sustainable development, an innovation and ability to replicate simultaneously; ethical standards and social inclusion; resource consumption and environmental indicators; economic feasibility, universality and applicability; contextual and aesthetic impact of the project.

3) *Excursions to objects certified according to «green» standards*

We have visited with student's green objects in our city and study features of facilities in other regions. One of the objects was the cottage village Ekodolia near Ekaterinburg to study the experience of building the house of class A+. The house was realized in accordance with the concept of Active House on the basis of the European experience. The main objective of this project is to find a solution for a modern Russian house that meets high requirements of resource saving, ecology and quality of habitat. At the same time, the cost of building and operating a house must be in keeping with capabilities of people.

4) *Collaboration with innovative companies*

Throughout the entire period of training students, we established contacts and regularly invited specialists from leading enterprises in the field of building automation "Smart Home", energy-efficient housing construction, developers of Russian green standards Green Zoom and many others.

5) *The organization of conferences and roundtables*

We organize scientific conferences and roundtables with the participation of our students and graduates on advanced technologies in the urban environment, practical seminars organized by companies that produce new materials and technologies ("Knauf", "Izover", "Teplit"), in major international exhibitions and conferences in Ekaterinburg ("Innoprom", "Forum 100+"). Within the annual exhibition "The Building Complex of the Big Urals" we have organized the round table on the application of BIM and Green BIM technologies in the Urals region:

difficulties, effects, prospects; BIM in the education: who teaches and how, what competencies are needed for employers.

6) *The participation of students in international and Russian competitions*

Over several years, students of our department took part in the international competition of the Saint-Gobain company for the design of multi-comfort buildings in different countries. In the Autodesk international competition our student Alexander Ivlev in 2012 received the main prize in the nomination "Give the shape to the future" with the project of an eco-house controlled by the automatic system Smart House. In the next section, this project is shown.

3.2 Innovations and IT in the Construction

New information technologies make up the basis of the current techno-economic paradigm. Its scientific basis is achievements in artificial intelligence and interdisciplinary approaches. We include innovative technologies in the educational process, explain, show and discuss different cases.

Cloud technologies, with their application the concept of Connected BIM appeared, when all processes, people and technologies are interconnected in the information environment. This technology allows you to quickly and efficiently manage the entire life cycle of an object.

The Smart Home and Smart City technologies are also based on cloud services and Internet of Things (IoT) wireless technology. IoT allows you to organize an independent interconnection of various objects using sensors and wireless networks, as well as control all parameters.

Virtual and Augmented Reality (VR/AR). The user can view the building inside and out, have additional information in infographics and interact with the environment. We have implemented a number of such projects for the innovative promotion of architectural projects.

3D scanning of the terrain, combining data with GIS and formation of 3D model of the object in relation to the terrain. The use of unmanned aerial vehicles to obtain relevant information about the current state of construction. 3D printing of buildings, small architectural forms, etc.

Artificial intelligence has unlimited possibilities for solving optimization problems based on models and algorithms of machine learning and neural networks.

3.3 Diploma Projects with Green BIM Technologies

Graduate projects of students of the Applied Informatics department which demonstrate skills in designing of energy-efficient facilities using Green BIM

technologies are shown here. In some cases, our diploma projects correspond to real objects – industrial, public and apartment buildings. In the Fig. 1 we see the futuristic project of a Smart house (Alexander Ivlev), designed to meet the requirements of environmental friendliness and energy efficiency.

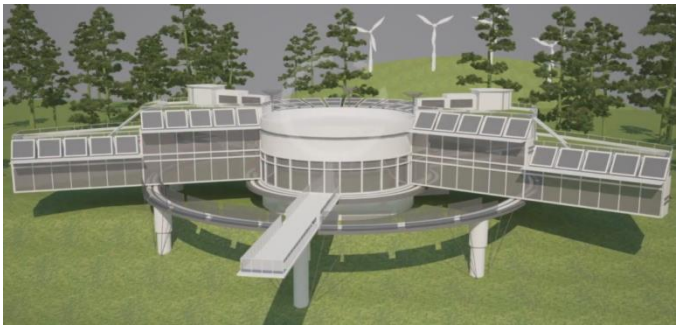


Figure 1
3D-visualization of the "smart" eco-house

For the least impact on the environment, the house is located on five columns, which also facilitate the placement of the building on uneven terrain, which will ensure the stability for the moving structure of the house. The work of the student was directed to a high-quality visualization and he developed a great video clip in After Effects program on the basis of 3DS Max models both the environment and the interior.

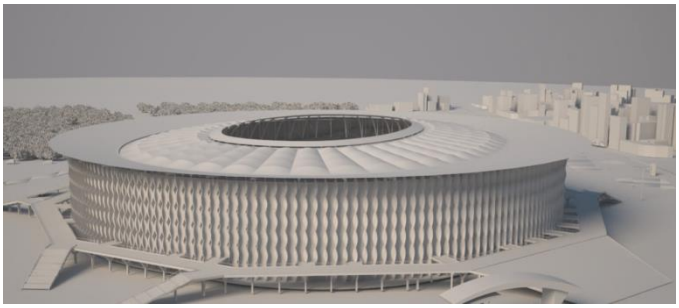


Figure 2
The stadium project in Tokyo

In the diploma work on Fig. 2 (Nadezhda Klyukina), a multifunctional stadium in Tokyo was designed using energy-saving architectural and construction technologies and energy efficient engineering solutions. When the stadium was designing, the attention was paid to the environmental optimization and in particular to the Green Goal program, which is being developed by the initiative of FIFA. In this work, the stadium project has been certified according to the LEED standard.

The architectural solution of the stadium is made using kinetic architecture, in particular, dynamic facades are offered as an alternative to sunscreens with low thermal conductivity, and also a way to organize the active effective ventilation. The facade was modelled in the Autodesk Revit program with the plug-in Dynamo.

The Autodesk Revit was also used for the information modeling. Along with the digital 3D model for joint development at all stages of the life cycle, the calculation of wind load, the analysis of external flow, the energy audit, the calculation of the heating, ventilation and air conditioning systems was carried out. Autodesk Green Building Studio software was used to conduct energy modeling and evaluate the energy efficiency of the building.

For more in-depth analysis, the following programs were used: Energy Plus for the analyzing of energy expenditure for heating, cooling, ventilation and lighting; Lighting analysis for Revit – for a more accurate analysis of the illumination of the object in accordance with the LEED standard; Green BIM Engineering uses Autodesk Simulation CFD 2014 to analyze the efficiency of ventilation and air conditioning systems, to develop a passive heating/cooling/ventilation, to study the influence of the shape of the building and obstacles on the external wind load; Flow Design – for modeling air flows. To calculate the financial model and investment attractiveness of the project, the software product Alt-Invest was used to prepare, analyze and optimize investment projects of various industries, scope and focus.

In the thesis work "The use of energy-efficient standard Green Zoom for design of residential buildings" (Ekaterina Kutishenko, Fig. 3) the conceptual design of the high-rise, apartment building was developed, its evaluation with the Russian system of energy efficiency and sustainability Green Zoom for civil and industrial construction projects has been made.

Figure 3
Visualization of the "green" apartment building

The strategic goal of the Green Zoom standard is to create a favorable living environment with affordable and comfortable housing. The evaluation of buildings according the Green Zoom system is carried out at 9 divisions: architectural and planning solutions, the project location and the organization of transport provision, environmental sustainability territory, water efficiency, energy efficiency, reduction of harmful emissions into the atmosphere, environmentally rational choice of building materials and waste control, the ecology of the internal environment of buildings, innovation in design and regional characteristics.

In the next diploma work (Maria Shevelyova, Fig. 4), the possibilities and advantages of using BIM in a reconstruction of architectural objects are studied using the example of the residential group in Madrid. A joint team of students of the architectural university and the Ural Federal University presented this project at the International Student Contest "Multi-comfort from Saint-Gobain 2017".

At present, many large urban centers are very tight, so architects pay tremendous attention, not to the construction of new facilities, but to the improvement and transformation of existing architectural fund. In this case, BIM as a tool, can bring great effect. BIM is especially important for the restoration of historical architecture objects. It is important to model the general condition of the object, as well as the degree of its wear and tear. In this case, it is better to use not modeling from plans, but scanning of a building with special equipment, since some parameters that may seem insignificant at first glance may turn out to be a destructive factor during reconstruction.

Figure 4
The reconstruction of the residential group in Madrid

The use of BIM in the reconstruction is a fairly new and unexplored direction. Usually, the building information model is applied at all stages of design and construction, including operation and demolition. For planning the reconstruction using BIM, you can also get all the necessary information. The advantages of BIM in the process of building reconstruction are as follows:

- The reconstructed object already possesses an architectural form and various systems, from which changes begin. In this case, BIM tools allow you to check automatically the object for various collisions, inconsistencies between old and new subsystems. Specifications are generated automatically; various calculations are performed, including cash estimates.
- The ability to model many reconstruction options, including mixing details of different versions.
- The possibility to develop a modification of the object by stages, for example, dismantling of a part of the building, installation of new architectural forms, installation of new systems and so on.
- Thanks to the information model, it is also possible to efficiently make a partial replacement of engineering systems.
- In the BIM model, you can specify the timing of the reconstruction, lay the time schedules and monitor the current state of the structure.
- Computer experimentation. Many worn objects are falling under reconstruction, so their innovating may be disastrous. That is BIM programs allowing reconstruction of various situations for optimizing systems and design solutions.
- Information modeling is also relevant for seismically unstable areas in the restoration of buildings, since bearing structures are most often only reinforced, and not completely reworked.
- BIM allows you to design a reconstruction of any age-related building, taking into account new standards for environmental and energy-saving requirements, thanks to which the facility can get a green certificate.

Conclusions

The world trend, in the modern construction, corresponds to the ecological, rational design, according to green standards and the application of BIM.

A modern BIM specialist, based on the main property of the BIM, is ready for teamwork in a common information space and should have a sufficient number of competencies to solve professional problems.

To widespread application of BIM and Green BIM technologies, it is necessary to develop the educational standards for teaching them at Universities.

Many architecture and Civil Engineering Universities, in different countries, introduce active and practice-oriented methods for teaching BIM. Herein, we have shown some of the features of these processes.

At the USUAA we teach interdisciplinary specialists, Informatics-Architects, and develop systems allowing us to comprehensively impart the principles of

environmental design and information modeling, through active practice-oriented methods.

Some results of Green BIM application, in diploma projects related to real architectural and construction objects, in accordance with green standards, are presented in this paper.

References

[1] Feist, W.: Gestaltungsgrundlagen Passivhäuser, Verlag Das Beispiel, 2000, 143 p. Russian translation: W. Feist, The main provisions on the design of passive houses, Moscow: ACB, 2011, 144 p.

[2] Krygiel, E., Nies, B.: Green BIM: Successful Sustainable Design with Building Information Modeling, Wiley Publishing, 2008, 268 p.

[3] Johnny Kwok, Wai Wong, Jason Zhou: Enhancing environmental sustainability over building life cycles through green BIM: A review, Automation in Construction, Volume 57, September 2015, pp. 156-165

[4] Yujie Lu, Zhilei Wu, Ruidong Chang, Yongkui Li: Building Information Modeling (BIM) for Green Buildings: A Critical Review and Future Directions, Automation in Construction, Vol. 83, Nov. 2017, pp. 134-148

[5] Tamer El-Diraby, Thomas Krijnen, Manos Papagelis: BIM-based Collaborative Design and Socio-technical Analytics of Green Buildings, Automation in Construction, Volume 82, October 2017, pp. 59-74

[6] Bahriye Ilhan, Hakan Yaman:. Green Building Assessment Tool (GBAT) for Integrated BIM-based Design Decisions, Automation in Construction, Volume 70, October 2016, pp. 26-37

[7] José Pedro Carvalho, Luís Bragança, Ricardo: Mateus Optimising building sustainability assessment using BIM, Automation in Construction, Volume 102, June 2019, pp. 170-182

[8] Araszkiewicz, K. Green BIM Concept – Scandinavian Inspirations, Archives of Civil Engineering, Vol. LXII, ISSUE I, 2016

[9] Baranyai Bálint, Bachmann Bálint, Kistelegdi István: Simulation-supported Design of a Hungarian National Sports Center, An International Journal for Engineering and Information Sciences, Volume 11, No. 1, 2016, pp. 113-127

[10] Shen Jiang, Wang Lei: The Application of BIM in Green Building Energy Saving: Take Helsinki Music Center as an Example, Advanced Materials Research, Volume 935, pp. 3-7

[11] Amarnath Chegu Badrinath, Yun-Tsui Chang, Shang-Hsien Hsieh: An Overview of Global Research Trends in BIM from Analysis of BIM Publications, Proceedings of 16th International Conference on Computing in Civil and Building Engineering (ICCCBE2016), July 6-8, 2016, Osaka,

Japan, http://www.see.eng.osaka-u.ac.jp/seeit/icccbe2016/Proceedings/Full_Papers/106-140.pdf

[12] Puolitaival Taija, Kestle Linda: Teaching and Learning in AEC Education – the Building Information Modelling Factor, Journal of Information Technology in Construction (ITcon), Volume 23, pp. 195-214, http://www.itcon.org/2018/10

[13] Rahman Rahimi A., Ayer Steven K.: Enhancing the Non-technological Skills Required for Effective Building Information Modeling through Problem-Based Learning, Journal of Information Technology in Construction (ITcon), Volume 24, pp. 154-166, http://www.itcon.org/2019/9

[14] Dakhil Ammar, Underwood Jason, Shawi Mustafa Al: Critical Success Competencies for the BIM Implementation Process: UK Construction Clients, Journal of Information Technology in Construction (ITcon), Volume 24, pp. 80-94, http://www.itcon.org/2019/5

[15] Yilmaza Gokcen, Akcameteb Asli, Demirors Onur: A Reference Model for BIM Capability Assessments, Automation in Construction, 101, 2019, pp. 245-263

[16] Oyewole Oluwatosin, Dada Joshua: An Evaluation of Training Needs of the Nigerian Construction Professionals in Adopting Building Information Modelling, Journal of Construction in Developing Countries, 24, № 2, 2019, pp. 63-81, http://web.usm.my/jcdc/current.html; doi.org/10.21315/jcdc2019.24.2.3

[17] Spryzhkov A. M., Privorotsky D. S., Privorotskaya E. V., Yashina N. A.: Information Modeling and Integrated Implementation of Projects in Cross-disciplinary Course Design for Students of Construction Specialties, Bulletin of the Samara Scientific Center of the Russian Academy of Sciences, Volume 18, No. 4, 2016

[18] Semenov A. A.: An Interdisciplinary Approach to the Training of BIM Specialists, New Information Technologies in Architecture and Construction: Materials of the II All-Russian Scientific and Practical Conference with International Participation. Yekaterinburg: USUAA, 2019, p. 43

[19] Grishina N. M., Chaly Yu. Yu.: Problems and Prospects of BIM in Universities: Development Management in Construction, News of KazGASU, 2017, No. 3 (41)

[20] Zakharova G., Krivonogov A., Petunin A.: The Need for Teaching of Green BIM Technologies in Higher School of XXI Century, Russian Journal of Construction Scince and Technology, Ural Federal University, Institut of Civil Engineering, 2017, Vol. 3, № 1, pp. 74-79

[21] Zakharova, G., Krivonogov, A.: Building Information Modeling on the Basis of Green Standards for the Purpose of Sustainable Development of Territories, Izvestiya Vuzov, Physics, № 8/2, 2016, pp. 37-40

[22] Krivonogov A., Zakharova G., Kruglikov S., Plotnikov S.: Implementation of BIM-Technologies in the Educational Program of the Architectural University: Matec Web Of Conferences 2018, C. 01001. Conference Proceedings, 9th International Scientific Conference Building Defects 2017, pp. 14-19

[23] BIM Project Execution Planning Guide [Online] Available: http://bim.psu.edu/ (current May 2020)

Design of Multi-Level Intelligent Control Systems for Complex Technical Objects on the Basis of Theoretical-Information Approach

Bary Ilyasov, Vladimir Vasilyev, Sagit Valeev

Ufa State Aviation Technical University, K. Marx Str. 12, 450000 Ufa, Russia
e-mail: ilyasov@tc.ugatu.ac.ru, vasilyev.vi@net.ugatu.su, valeev.ss@net.ugatu.su

Abstract: The problem of designing multi-level intelligent control systems for complex technical objects on the basis of the theoretical-information approach is discussed. The proposed approach allows us to make well-grounded project decisions when designing control algorithms for advanced control systems. The problem statement of applying the proposed approach to the design of intelligent control systems for aviation gas-turbine engines is considered.

Keywords: multi-level control system; intelligent control design; entropy; aviation engine

1 Introduction

When designing automatic control systems (ACS) for complex technical objects some main common criteria are usually used: the global goal of control process; the final level of achieving this goal; the computational resources required to achieve the delivered goal with a given performance; the complexity and the time terms of the design process. In addition, we need to reduce the total cost of designed ACS, guarantee the reliability of the system operation under conditions of coordinate-parametric and structural disturbances action. The requirements of network interaction in the framework condition of group control and cooperation in ACS as a multiagent system are also to be kept in mind. All these problems lay in the table of designing three-aspect components of ACS (so called problem C3 - Control, Computation, and Communication). Since the requirements of ACS characteristics for complex technical objects are constantly growing, the complexity of the control systems and the complexity of their design process become the key factors determining the quality of design solutions and the efficiency of ACS.

One of perspective directions in the modern control theory is the development of intelligent control systems (ICS) on the basis of IPDI (Increasing Precision with

Decreasing Intelligence) principle [1], the essence of which consists in decomposition of control objectives and tasks depending on the requirements to precision and intelligence of control problem solution. Here, the higher is the required control precision at a hierarchical level, the lower are requirements to the control intelligence level; and on the contrary, the higher is the required level of intelligence, the lower is the necessary precision of control [2].

The application of such an approach gives a possibility to estimate and minimize the complexity of design decisions under construction of control systems for complex technical objects (CTO).

The base architecture of a hierarchical ICS consists of three vertical interacting levels: the organizational, the coordination, and the executive control level.

The first who offered the notation of control system complexity lying in the base of law of requisite variety was W. Ashby [3]. This approach was then developed by V. Solodovnikov as the principle of minimal complexity [4]. According to the given principle, the design of CTO control system is reduced to the selection of the controller structure having minimal complexity under the specified requirements to control performance.

The problem discussed below is how to formalize the procedure of designing the multi-level intelligent control system, how to set the specifications for different control levels, and how to choose the control algorithms on each of control levels having minimal structural complexity. The universal measure of the structural complexity being used here is the entropy of control processes.

The indicated tasks of designing the control systems for complex technical objects are related to the class of multicriteria optimization problems. To reduce the number of iterations during the design procedure and to ensure the required performance of control processes, it is necessary to evaluate and minimize the complexity of the project solutions at the preliminary design stage.

Let us consider below the formulation of the problem of designing the control system for the complex technical object on the example of control algorithms design for the aviation engine (AE). The prospects of such systems development are associated with the development of multi-level intelligent control systems (MICS) that ensure the fulfillment of the given set of requirements to performance, robustness, and fault tolerance of AE control processes [5].

At the same time when designing AE MICS it is necessary to take into account the rigid constraints imposed on the computational processes implementing the control algorithms [6].

The tasks of analysis and design of ACS in the class of AE MICS have a number of special features:

- the use of control algorithms based on soft computing methods (fuzzy logic, neural networks, genetic algorithms, etc.) mostly is based on the trial and error

method. Theoretically, only the existence theorems of these algorithms are known in the literature, there are no formalized methods and techniques for analyzing and synthesizing algorithms for the intelligent control system of AE [7-17];

- it is necessary to guarantee the stability of control processes in the AE MICS both under conditions of the nominal modes and the influence of uncertainty factors (for example, in case of subsystems failures or appearance of possible structural defects in distributed databases and knowledge bases) [18].

These circumstances make it very difficult to solve the problem of design and analysis of MICS that significantly limits the use of achievements in the field of artificial intelligence theory for solving practical problems in this subject area [19-20].

The paper considers the formalized approach to solving the problem of designing AE MICS based on the theoretical information approach. The proposed approach allows us to make well-founded design decisions while designing the control algorithms of advanced MICS for the new generation of AE based on the fundamental principle of minimum complexity.

2 The Statement of the MICS Design Problem

As an analysis shows, the conception of AE MICS design must be based on the use of the following additional system principles:

1) the principle of the functional integration which supposes a unification of different (heterogenous) subsystems;

2) the principle of the hierarchical organization, meaning the construction of the control system in a class of multi-level hierarchical control systems with a division (decomposition) into several control levels differing by a choice of control goals and implementation methods:

3) the principle of integrating the various models, methods, and algorithms of analysis and design of control system with the use of both the classical methods of multivariable control and intelligent control of complex dynamic objects;

4) *the principle of the minimal complexity*, supposing a choice of the simplest structure of control algorithms with the account of the change of AE operation conditions at the optimum use of information and computing resources of the control system;

5) the principle of the open system construction, being a basis of intellectualization and standardization of information processing technologies at the various stages of MICS life cycle.

Let us suppose that any possible control situation can be defined as a triade including the control object state, the environment conditions and the control goal. The mapping $\Pi = (\pi 1, \pi 2, \pi 3)$ characterizes then a procedure of designing the control algorithms of MICS on the basis of their decomposition to three interacted control levels:

$$< \text{Control object} > \xrightarrow{\pi_1} < \text{Algorithms of executive level (EL)} >;$$

$$< \text{Environment} > \xrightarrow{\pi_2} < \text{Algorithms of coordination level (CL)} >;$$

$$< \text{Control goal} > \xrightarrow{\pi_3} < \text{Algorithms of organizational level (OL)} > .$$

The complexity of design decisions here should correspond to the complexity of the control situation. As a measure of control situation complexity, in this case, one can use the entropy estimates:

$$\left.\begin{array}{l} \pi_1 : H(Y/U,F) \text{ Complexity of EL algorithms} \\ \pi_2 : H(F) \text{ Complexity of CL algorithms} \\ \pi_3 : H(G) \text{ Complexity of OL algorithms} \end{array}\right\} \text{Entropy as a measure of design decision complexity}$$

Here $H(Y/U,F)$ is the entropy of control processes on the MICS EL-level;

$H(F)$ is the entropy of the environment conditions change;

$H(G)$ is the entropy of the control goal change;

Y, U, F, G are the vectors of the measured output parameters of the control object, the control actions, the external disturbances, and the control goals respectively.

The entropy of the output parameters vector Y is defined in this case as

$$H(Y/U,F) = -\int_{\Gamma_Y} p(Y/U,F) \ln p(Y/U,F) dY, \tag{1}$$

where $p(Y/U,F)$ is the distribution density function for the output vector Y. The entropy of the environment conditions change is estimated by the expression

$$H(F) = -\int_{\Gamma_F} p(F) \ln p(F) dF, \tag{2}$$

where $p(F)$ is the distribution density function for the vector F. The entropy of the control goal $H(G)$ can be estimated similarly.

With an account of the above mentioned, the problem of optimum design of three-level MICS on the basis of entropy approach is formulated as follows.

In the general case, the description of the AE behavior as a multivariable dynamic control object in the state space can be represented as a system of non-linear differential and algebraic equations

$$\dot{X} = F(X,U,V)$$
$$Y = H(X,U,V) \tag{3}$$

where $X \in R^n$ - is the state variables vector of AE; $Y \in R^m$ - the vector of the measured coordinates (outputs) of AE; $U \in R^l$ - the vector of the control actions (inputs) of AE; $V \in R^k$ - the vector of disturbances characterizing the external environment conditions.

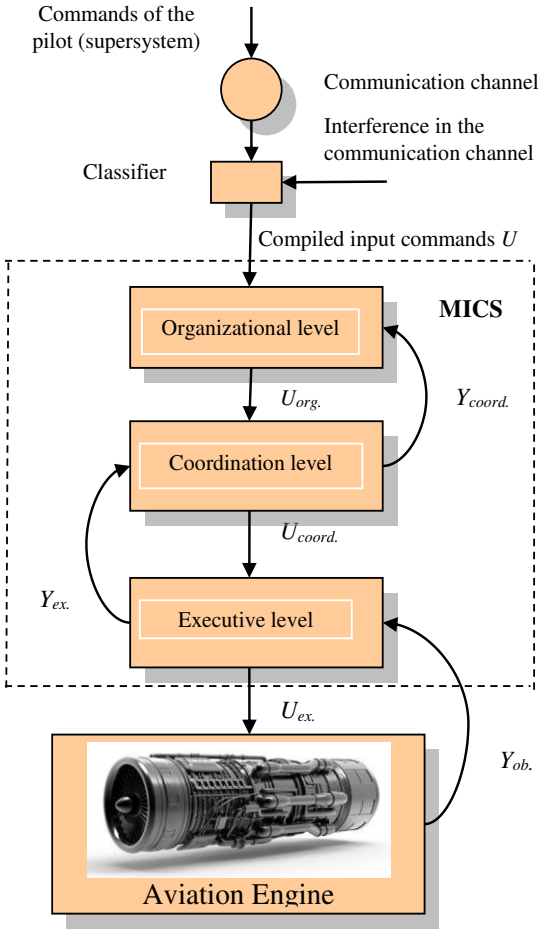


Figure 1
Multi-level control system of aviation engine

In Figure 1 the block diagram of the multi-level intelligent control system of AE consisting of the executive level, the coordination level, and the organizational level is presented.

To achieve the settled goal of control at each level, a set of interconnected algorithms is applied, for implementation of which the following information about the current state of the control situation is used:

- $X \in \Gamma_X$ - the control object state vector, $\Gamma_X = \bigcup_i \Gamma_{X_i}$ - the range of the AE possible states (correspondingly regular, potentially dangerous and emergency operating modes);

- $Y \in \Gamma_Y$ - the vector of the state variables values available for the measurement, $\Gamma_Y = \bigcup_i \Gamma_{Y_i}$ the range of the AE values for indicated possible states;

- $U \in \Gamma_U$ - the control actions vector, Γ_U - the area characterizing the possibility of the goal-directed change of the AE state vector;

- $F \in \Gamma_F$ - the vector of disturbances from the environment: $F(1)$ – the parametric disturbances; $F(2)$ – the signal disturbances; $F(3)$ – the structural disturbances (changing the configuration of the control object, failures, etc.). The range of the values for each of these disturbances classes contains, in turn, some nested regions.

Under the control goal we will understand the solution of the problem of approximation of the output parameters vector Y concerning the desired result, i.e. the target vector Y^*:

$$\left\| Y - Y^* \right\| \le \varepsilon \tag{4}$$

where ε is the given admissible control error; Y^* - the desired value of the AE output variables vector at the given operating mode.

Let us assume that the optimal control reduces the entropy of the control vector concerning the desired value, i.e. ensures the following condition fulfillment:

$$U_{opt} : H_{\max}(Y / U_{opt}, F) \le H_{adm.} \tag{5}$$

where $H_{adm.}$ - is the admissible entropy determined by the given requirements to the control processes performance:

$$H_{adm.} = \max_{F \in \Gamma_F} \{ -\int_{\Gamma_Y} p_{des.}(Y) \ln p_{des.}(Y) dY \}, \tag{6}$$

where $p_{des.}(Y)$ is the desired (given) distribution probability density, chosen on the base of specified accuracy of achieving the goal (specified for the control system in the form of the error distribution law);

$H_{\max}(Y/U_{opt},F)=\max_{F\in\Gamma_F} H(Y/U,F)$ is the entropy of the output vector Y, corresponding to the optimal control $U_{opt.}$.

The evaluation of the desired entropy (i.e. the entropy of the target vector Y^*) of the ACS can be defined as follows:

$$H(Y^*)=-\int_{\Gamma_{Y^*}} p(Y^*)\ln p(Y^*)dY^*, \tag{7}$$

where $p(Y^*)$ is the distribution probability density of the vector Y^*.

The purpose of the hierarchical AE MICS design is the choice of the set of such control system algorithms that provide the specified performance of control processes under minimal complexity of control algorithms.

3 Solution of Optimal Design Problem of AE MICS

The following problem statement is considered: it is necessary to find such method of constructing AE MICS and its hierarchical control levels, i.e. to determine such composition of control algorithms, database structure (DB) and knowledge bases (KB) of MICS, under which the requirement (5) is fulfilled, provided that the following condition is fulfilled:

$$H_{\Sigma}(A)\to\min \tag{8}$$

where $H_{\Sigma}(A)$ is the total entropy of the set of control algorithms for all three levels of AE MICS, which can be calculated as

$$H_{\Sigma}(A)=H_{ex.}(A)+H_{coord.}(A)+H_{org.}(A) \tag{9}$$

here, $H_{ex.}(A)$, $H_{coord.}(A)$, $H_{org.}(A)$ are respectively the entropy (computational complexity) of the algorithms of executive-level, the coordination level and the organization level of MICS.

It is assumed that in addition to condition (5) the following restrictions must be provided also:

$$H_{\max}(Y,U_{opt},F)\le H_{adm.};\forall X\in\Gamma_{X_1};Y\in\Gamma_{Y_1};U\in\Gamma_U;\ \forall F\in\Gamma_F, \tag{10}$$

where $H_{adm.}$ is the specified (maximum admissible) entropy of MICS control processes.

The solution of the stated design problem is carried out on the basis of the principle of sequential decomposition ("bottom-up"). In accordance with it at each of the control level, a selection of algorithms with minimal complexity satisfying the given requirements (4), (5), (8), (10) to control processes at the executive, coordination, and organization levels MICS are to be provided.

3 Examples of AE MICS Design

Let us consider the examples of designing the intelligent control system for the AE.

Example 1. In [6] the model of two-shaft multi-mode gas-turbine engine is presented as the set of piece-wise linear dynamic equations of the following form:

$$\begin{aligned} \dot{X} &= A(\eta)(X - X^{ss}(\eta)) + B(\eta)(U - U^{ss}(\eta)), \\ Y &= Y^{ss}(\eta) + C(\eta)(X - X^{ss}(\eta)) + D(\eta)(U - U^{ss}(\eta)). \end{aligned} \tag{11}$$

where

$$A_i \in R^{2x2}, B_i \in R^{2x2}, C_i \in R^{3x2}, D_i \in R^{3x2},$$
$$X \in R^2, U \in R^2, Y \in R^3;$$

$X = (n_L, n_H)^T$ is the vector of the state variables (which are defined as the AE rotation velocities of the low and high-pressure compressor rotors respectively); $U = (G_f, A_N)^T$ - the vector of control actions (the fuel flow rate and the jet nozzle area); $Y = (p_C^*, p_T^*, T_T^*)^T$ - the vector of controlled variables (the air pressure behind the compressor, the pressure and the gas temperature behind the turbine); $X_i^{ss}, U_i^{ss}, Y_i^{ss}$ are the values of the vectors *X*, *U*, *Y* at the steady-state operation modes of the gas-turbine engine (i = 1 ÷ 4); $\eta = \alpha \cdot \bar{n}_L + \beta \cdot \bar{n}_H$ - the parameter determining a selection of the i-th point of AE operation mode (or respectively the i-th piece-wise linear model); $\bar{n}_L, \bar{n}_H$ - the relative (dimensionless) values of the variables n_L and n_H.

If we solve the interpolation task with the use of a 3-layered neural network (NN), then the nonlinear multi-mode dynamic model of gas-turbine engine will have the following form:

$$\begin{aligned}
\dot{X} &= (\sum_{i=1}^{z} W_i^1 f_i(X) + B^1)(X - X^{ss}) + (\sum_{i=1}^{z} W_i^2 f_i(X) + B^2)(U - U^{ss}),\\
Y &= (\sum_{i=1}^{z} W_i^3 f_i(X) + B^3)(X - X^{ss}) + (\sum_{i=1}^{z} W_i^4 f_i(X) + B^4)(U - U^{ss}),\\
X^{ss} &= \sum_{i=1}^{z} W_i^5 f_i(X) + B^5,\\
U^{ss} &= \sum_{i=1}^{z} W_i^6 f_i(X) + B^6,
\end{aligned} \tag{12}$$

where $f_i(\cdot)$ is the activation function of a neuron; W_i^j - the tuned weights of NN; B_j – the biases in the separate layers of NN, $i = 1...N$, (N – the number of the neurons in the hidden layer).

The model (12) accuracy will depend in this case on the number of the neurons in the hidden layer. An obvious advantage of such an approach is the possibility of obtaining the required interpolation accuracy for coefficients and variables by means of NN learning.

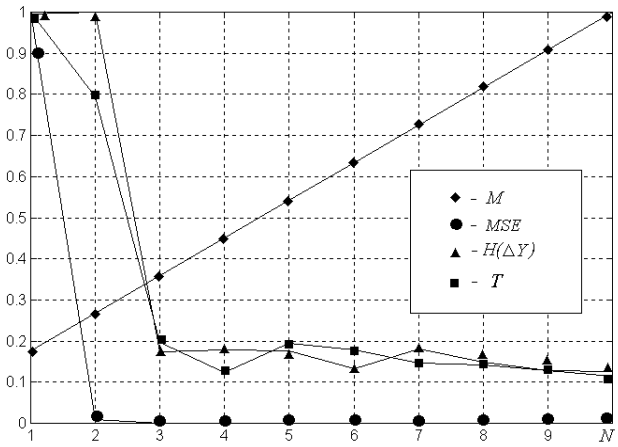


Figure 2

NN-model characteristics depending on the number of the neurons in the hidden layer

In Figure 2 the results of AE NN-model design for the considered case are presented, depending on the number of the neurons N in the hidden layer (note that the scaling procedure concerning the maximal values of the variables was used). Here: MSE is the mean square error of NN learning; T - the time of NN training; M – the necessary volume of the memory for storage of NN coefficients; $H(\Delta y)$ – the entropy of NN output signals, determined with an account of (12). For estimation of NN algorithmic complexity it is possible to make an assumption that

the complexity of NN-model in our case is proportional to M, i.e. it grows with increasing N.

It follows from Figure 2 that at N =3 the optimum ratio between the complexity of NN model and its accuracy is reached.

In the process of training AE NN, there is a decrease in the entropy of the output reactions of the network due to additional information on the goals and actual learning outcomes obtained from the outside. Thus, in the process of adjusting the weights of the AE NN, it accumulates a certain amount of knowledge aimed at reducing the a priori uncertainty inherent in the beginning of the learning process.

The level of a priori uncertainty is determined in this case by the magnitude of entropy, and the ability of AE NN to assimilate the necessary knowledge is determined by the magnitude of the entropy of the weights of its interneuronal connections.

Example 2. Let us consider below the control system based on the application of the dynamic inversion method. The characteristics of the object inverse nonlinear dynamic model are realized here on the basis of NN. An inclusion of NN in the control loop provides the compensation of the object nonlinearities, facilitating a problem of forming the desired dynamic properties of the control system as a whole.

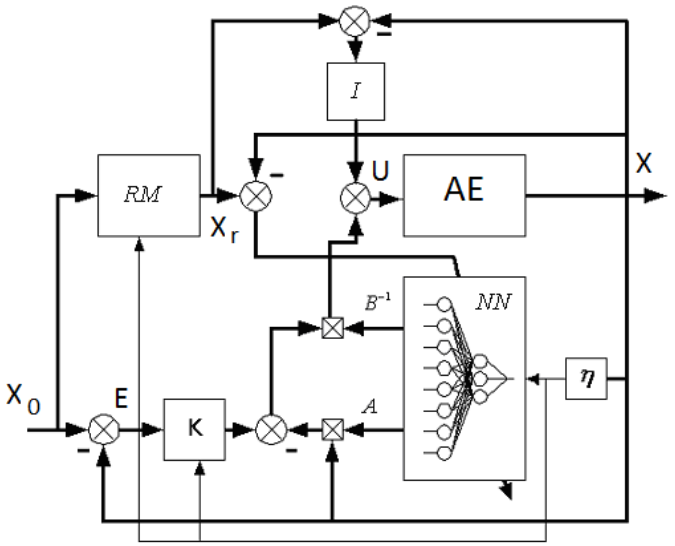


Figure 3
Gas-turbine engine control system with inverse NN-controller

Consider the construction of a multivariable controller for AE as the dynamic object represented by the set of equations (12). On the basis of the approach described above, the NN is designed. The value of AE mode parameter η is the input of this NN, and at its output, the values of matrices $A(\eta)$ and $B^{-1}(\eta)$ coefficients, necessary for the compensation of nonlinearity of AE characteristics, are calculated. The performance of the transient processes in the control system is determined by the values of matrix K coefficients (Fig. 3). If necessary it is possible to train again NN that allows us to compensate for the errors of the inverse NN-controller and the deviations of AE parameters as the control object.

As the inverse model of the gas-turbine engine, the model described above can be used. It is necessary to train it for calculation of the values of matrix $B^{-1}(\eta)$.

In Figure 3 the following designations are used: *RM* is the reference model specifying the desirable performance of the transient processes depending on AE operation mode (η); *I* is the integrator, providing the zero steady-state error of the control system. The transient processes obtained for the variable $\overline{n}_L$ of the gas-turbine engine and the variable $\overline{n}_{LR}$ from the output of the reference model are presented in Figure 4.

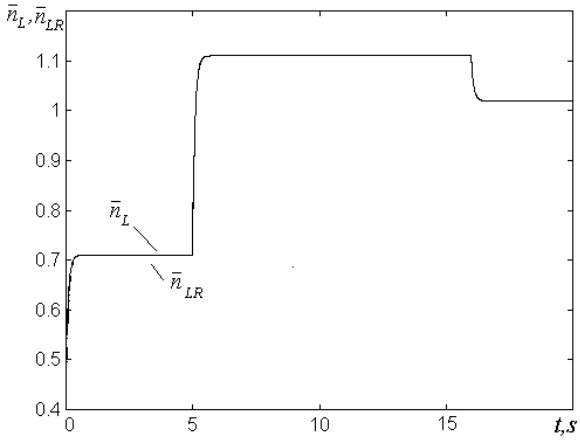


Figure 4
Simulation results of AE processes

As it is seen from Figure 4, the required performance of the transient processes is provided for different operation modes of the control object.

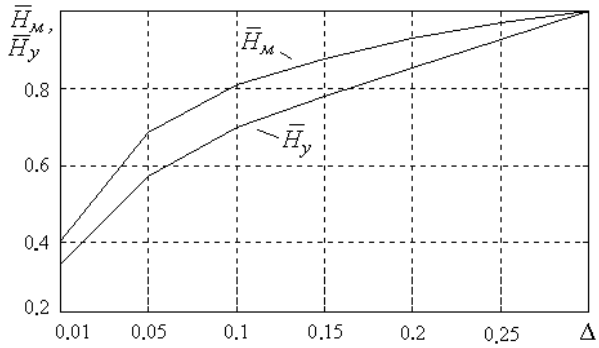


Figure 5

The change of entropy at the outputs of the system and AE inverse model

Consider the influence of the entropy at the output of the inverse NN - model on the entropy of the control system. For this purpose the numerical experiment was carried out during which the parameters of the inverse model were changed increasing a mismatch with parameters of the control object (the loop with integrator was open). Then, the entropy at the output of the inverse model was calculated as

$$H_M = \log(\|\Delta A\| * k1) + \log(\|\Delta B^{-1}\| * k2) \tag{13}$$

where $\|\Delta A\|, \|\Delta B^{-1}\|$ are respectively the errors of the inverse model matrices ($\|\cdot\|$ is a norm of a matrix), $k1$ and $k2$ is the scaling coefficients. In Figure 5 the curves of the entropy change at the output of AE model and the entropy at the output of the control system depending on the value of the error Δ of inverse model are presented (the plots are given for the dimensionless values:

$\overline{H}_M = H_M / (H_M)_{\max}, \overline{H}_y = H_y / (H_y)_{\max}$.

The entropy at the output of the control system was calculated here with the aid of the following expression

$$H_y = \int_{t_0}^{t_k} (k_M \sum_{i=1}^{2} \dot{y}_{M_i}^2 - k_y \sum_{i=1}^{2} \dot{y}_{y_i}^2) dt, \tag{14}$$

obtained on the basis of considered in [7] approach to the entropy determination at the output of the dynamic system. As follows from Figure 5, the entropy of the control system is proportional to the entropy at the output of the inverse model. So, it is possible to set the requirements to the accuracy of gas-turbine engine inverse NN - model depending on requirements to the accuracy of the control system as a whole. For reduction of the entropy at the output of the system in this case the relearning of NN-model is possible while the control system operation.

Conclusions

The paper discusses the general problem of designing multi-level intelligent control systems for complex technical objects on the basis of the theoretical-information approach. At the same time, when designing it is necessary to take into account the rigid constraints imposed on the computational processes implementing the control algorithms. As an example, the problem of constructing the control system for an aviation engine as a complex technical object is considered. The application of the proposed approach in practice will make it possible to obtain well-grounded decisions in the field of intelligent control systems and develop high-effective multi-level control algorithms for advanced technical objects.

References

[1] Saridis G. N. Hierarchically Intelligent Machines, World Scientific Pub., Singapore, 2001

[2] Saridis G. N. Entropy in Control Engineering, World Scientific Pub., Singapore, 2001

[3] Ashby W. R. Introduction to Cybernetics, Stanford, 1955

[4] Solodovnikov V. V., Biryukov V. F., Tumarkin V. I. Principle of Complexity in Control Theory, Moscow, Nauka Pub., 1977 (In Russian)

[5] Problems of Design and Development of Gas-Turbine Engine Control Systems / S. Kusimov, B. Ilyasov, V. Vasilyev, S. Valeyev, et al. – Moscow, Mashinostroyenie Pub., 1999 (In Russian)

[6] Neurocomputers in aviation (aircrafts) / Eds.: V. Vasilyev, B. Ilyasov, S. Kusimov, Moscow, Radiotechnica Pub., 2004 (In Russian)

[7] Yu H. Network Complexity Analysis of Multilayer Feed-forward Artificial Neural Networks. – In: Schumann J., Lin Y. (Eds.), Applications of Neural Networks in High Assurance Systems, SCI 268, Springer – Verlag, Berlin, Heidelberg, 2010, pp. 41-55

[8] Rzevski G. A. Practical Methodology for Managing Complexity // Emergence: Complexity & Organization // International Transdisciplinary Journal of Complex Social Systems. 2011, Vol. 13, No. 1-2, pp. 38-56

[9] Funahashi K. On the approximate realization of continuous mappings by neural networks // Neural Networks. 1989, Vol. 2, pp. 183-192

[10] Hornik K. M., Stinchombe M., White H. Multilayer feedforward networks are universal approximators // Neural Networks. 1989, Vol. 2, pp. 359-366

[11] Pinkus A. Approximation theory of the MLP model in neural networks // Acta Numerrica. 1999, Vol. 8, pp. 143-195

[12] Gomez I., Franco L., Jerez J. Neural Network Architecture Selection: Can Function Complexity Help? / Neural Pro-cess Lett., 2009, 30, pp. 71-87

[13] Kon M. A., Plaskota L. Complexity of Predictive Neural Networks. In: Unifying Themes in Complex Systems. Chapter 18, 2006, Springer – Verlag, pp. 181-191

[14] Silva L., J. M. de Sa, Alexandre L. A. Neural network classification using Shannon's entropy // Proc. of the 13th European Symposium on Artificial Neural Networks, Bruges, Belgium, April 27-29, 2005

[15] Bianchini M., Scarselli F. On the Complexity of Neural Network Classifiers: A Comparison Between Shallow and Deep Architectures // IEEE Transactions on Neural Networks and Learning Systems. August 2014, Vol. 25, No. 8, pp. 1553-1565

[16] Kurkova V. Complexity of Shallow Networks Repre-senting Finite Mapping. In: Artificial Intelligence and Soft Computing // Proceedings of the 14th International Confer-ence ICAISC, 2015, pp. 39-48

[17] Erdogmus D., Principe J. C. Entropy minimization algorithm for multilayer perceptrons // Proceedings of International Joint Conference on Neural Networks (IJCNN'01) 2001, Vol. 4, pp. 3003-3008

[18] Vasilyev V., Valeyev S., Sun Jianguo, Identification of Complex Technical Objects on the Basis of Neural Network Models and Entropy Approach // Proc. of the 9th World Multi-Conference on Systemics, Cybernetics and Informatics, Vol. IX, July 10-13, 2005, Orlando, USA, 2005, pp. 89-93

[19] Advanced Multivariable Control Systems of Aeroengines / Eds.: Sun Jianguo, V. Vasilyev, B. Ilyasov, Beijing, China, 2005

[20] Vasilyev V. I., Valeyev, S. S., Sun J. Identification of Complex Technical Objects on the Basis of Neural Network Models and Entropy Approach, Proc. of the WMSCI 2005 - The 9th World Multi-Conference on Systemics, Cybernetics and Informatics, Vol. 9, pp. 89-93, 2005

Architecture of the Security Access System for Information on the State of the Automatic Control Systems of Aircraft

Arkady Isaakovich Frid, Alexey Mikhailovich Vulfin, Victoria Viktorovna Berholz, Dmitry Yurievich Zakharov, Konstantin Valerievich Mironov

Faculty of Informatics and Robotics, Ufa State Aviation Technical University (USATU), K. Marx Street 12, Ufa, 450008, The Republic of Bashkortostan, Russian Federation
frid.ai@net.ugatu.su, vulfin.am@net.ugatu.su, berholc.viktoriya@net.ugatu.su, zakharov.dju@net.ugatu.su, mironov.kv@net.ugatu.su

Abstract: This article discusses a problem of a secure access provision, via Web-service, to a database. Databases contain critical information about the parameters of the life-cycle of complex technical devices (CTD). Information life-cycle support (LS) of CTD provides access to a projects organization, production facilities, suppliers, service organization and the end users in long-term projects, which complex system projects actually are, at all stages of the device's life-cycle.

Keywords: data base; web application; secure access; telemetry; aircraft

1 Introduction

Prospects for the development of aviation technology are closely related to the information support of flights and the operation of aircraft and their subsystems. Digital control systems, with a distributed structure, with elements of artificial intelligence, are increasingly used. Wireless control systems are being developed for the exchange of information between the components of aviation systems. Such types of exchange are carried out by means of radio signals. This can also be applied to automatic control systems (ACS), with gas turbine engines (GTE) [1]. Thus, information from sensors of the engine operational process (telemetric information, TMI), can be accessed through the radio channel, the control actions generated by the calculator. These fundamental issues should be resolved during development a wireless ACS of GTE [2]:

- Ensure the effective transmission of radio signals and interference immunity of communication channels
- Working of information channels in real time
- Reliability of transmission (passing) of the signal on board
- Autonomy of the power supply of certain elements of the information channels

However, the list of issues that need to be discussed is much wider. For example, the if there is a danger of external intervention, the problem of information security will arise.

2 Problems of Telemetric Information Transmitting

The more that the information exchange infrastructure on board an aircraft is developed, the more questions arise, related to provision of noise immunity and integrity of information. If the issues of protection from interference are solved by well-tried methods (anti-jamming code, etc.), the provision of integrity requires non-trivial solutions, in some cases. Attackers can intervene in the flight control process or in the process of transferring and storing data on the current state of aviation equipment. This type of data is called telemetric information. Cases of such intervention are known. Current telemetry transmission systems demonstrate vulnerabilities that are sufficient, not only for the leakage of passenger and airline data, but also for altering the course of the aircraft [2]. For example, a study of ground-to-board communication systems showed that the ACARS system, despite its versatility and ubiquitous use, is vulnerable. If it is hacked with ADS-B, an attacker can gain access to the flight control system and download flight plans and detailed commands [3].

The traditional technical solutions are used on equipped air routes are completely unsuitable for hard-to-reach areas, sparsely populated areas, with an undeveloped land infrastructure.

At the same time, the real costs of maintaining existing TMI transmission systems are based on satellite communications and are quite large and makes its use, very problematic.

Aeronautical telecommunication systems are also not reliable, but the combination of SATCOM satellite communications and HFDL data transmission mode can provide a higher level of system reliability. However, issues of information security do not allow for the classification of such systems as “reliable” class TMI transmission systems. A successful attack can jeopardize the management of the satellite communication channels. They are used by the Future Air Navigation system, the Pilot Data Link Communications (CPDLC) controller, or the onboard

address and system communication system (ACARS). Despite detailed vulnerability analysis, researchers do not report on the technical means used to break into satellite communication systems [4].

Thus, there is a need to ensure the reliability of data transmission systems on the status of elements and subsystems of aircraft based on the application of modern (including intelligent) technologies for the protection and processing of TMI. Further consideration of issues of ensuring the reliability of data transmission systems on the state of aviation equipment, to the enterprise, will be carried out with reference to the CCD GTE.

At the stage of operation of aviation equipment, the actual task is to collect information on the actual state of operational product and transfer it to the developer enterprise. On the basis of such information support, the enterprise can develop additional recommendations on the technology for further operation of the product, make improvements to the technical documentation, implement a number of measures, that are aimed to improve operational characteristics. To collect and store this information, enterprises create databases. The information on the stage of operation of the products and aggregates, the results of equipment inspections, failures, etc. are collected and stored. Requirements for information management systems of the initial stages of operation for enterprises developing and producing ACS by aviation GTEs should include [5]:

1) Automatic recording of information transferred from service
2) Ensuring the integrity and consistency of data as it accumulates in the database
3) The ability to extract from the database, at any time and for a long time (up to30 years) and in any required query format.

This ensures the integrity and consistency of data that enters and stores in the database requires the development of special security measures.

The current stage in the development of on-board telemetry systems (BTMS) is characterized, firstly, by the large amount of data generated by the sensors of the equipment and other types of information sources [6], and secondly, by the need to eliminate errors associated with data transfer from instruments and all telemetry- and, thirdly, the need to auto-mate the processing of most of the information, including the work on decision-making. Therefore, the problem of transmission and analysis at ground-based service centers of operational parametric information collected by BTMS is urgent for solving the tasks of supporting the life cycle of the gas turbine at the stage of operation and maintenance. The main parameters necessary for the registration and reflection of such information in the database of the GTE status are:

- Where is the operation of the product
- Object on which the product is operated

- The total operating time - the time of operation of the product from the beginning of operation
- Operating time - the time of operation of the product with reference to a specific object
- A fixed event linked to the product (repair, malfunction)

To solve these problems, it is necessary to develop an automated information system (AIS), capable of being distributed in large areas, processing large amounts of information and capable of providing remote access to necessary data. The problem of providing secure access to a database containing critical information about the parameters of the life cycle of the CCD GTE is relevant and in connection with the entry into force of new documents in the field of technical control by FSTEC [6]. The issues of ensuring secure access to such databases were considered in [7, 8, 9, 10], but, without taking into account changes in the regulatory framework.

The article proposes the structure of the organization for receiving and storing data obtained from the operation sites of the SAU GTE units at the enterprise-developer. The proposed scheme allows providing secure access to a database containing critical information about the GTE being operated on the basis of a protected WEB application.

The purpose of this study is to improve the security of access to a database containing critical information about the operated gas turbine engine.

Proceeding from the set goal, the following task is formulated:

1) Development of the architecture of a secure WEB-application for access to the data-base of the AIS system for supporting the life cycle of the SAU GTD.
2) Analysis of the existing structure of the AIS supporting the life cycle of the SAU GTD.

AIS is a set of software and hardware needed to receive, store and process information,

AIS solves the main tasks associated with receiving telemetric information about the state of the SAE GTE. Telemetry data from the BTMS board is planned to be obtained in three different ways:

1) Directly from the side of the aircraft during the whole flight time. This method of data collection is of the greatest interest for the study, since the use of real-time monitoring technologies will allow a full (operational, searching and intelligent) analysis of the operability of the aircraft systems, crew condition and control of its actions directly during the flight. If there

are technical problems with any of the modules, the Engineers will learn about this in advance and will be ready to repair before the aircraft lands.

2) Devices for reading the event log from the sensors of the aircraft modules. When performing technical inspection and maintenance of an aircraft on the ground, devices of this type are read and stored on the status of the modules throughout the previous flight.

3) Entering events into the database manually. The operator processes the information and enters information through the web application.

In the second and third cases, information enters the database through a web application that is an insulating layer between external networks and the internal structure of the AIS, since access from the external network is one of the most vulnerable places of the system.

A typical scenario for using a database to store critical information involves its operation in an environment that is hidden from everyone except network administrators and database administrators.

The widespread use of cloud technologies, mobile devices, virtualization, the Bring Your Own Device (BYOD, personal devices for work) and various technologies for remote work leads to the erosion of traditional perimeter protection of the corporate network. With the expansion of ways to access corporate resources (DB), the network no longer has a single entry point [6]. In this scenario, the risks of information security are increasing, as the company provides more and more access rights to external entities, which may become an indirect reason for the intruder to gain access to the database server located in the internal network of the company. One of the solutions to the problem is the organization of access to the information stored in the database, using a web browser that interacts with the middleware of the company, that accesses the database.

Such access to the database can be limited by means of access control facilities and provided only to a few permitted roles. Thus, if an attacker can overcome the perimeter protection of the network and will be able to perform queries against the database, provided that he does not have an account with which one of these roles is assigned, the database will still be safe. This process simplifies access control and ensures that no user (including an attacker) can access the database directly, but only through an account that has an associated role.

3 Interaction between AIS and Web Application

Web application provides secure access to AIS. It is an intermediate link between the AIS and the client. It is connected to a client via internet and to AIS via

enterprise's local network. When a client refers to a web application service they believe they are working directly with AIS. [7] A general diagram of interaction between AIS and web application is shown in Figure 1:

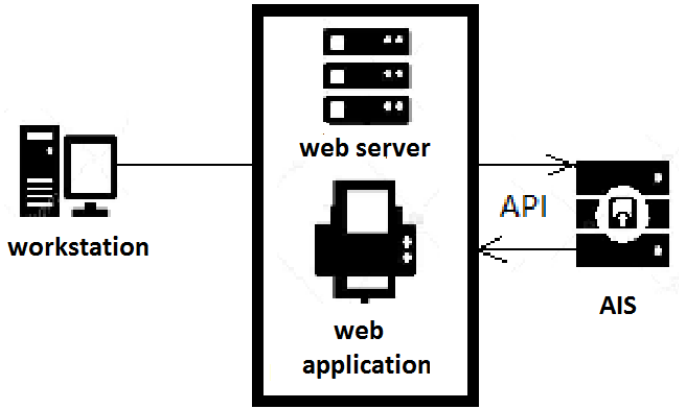


Figure 1
Interactions of the AIS and the web application

The Web Application is developed to make access to AIS more secure. It serves as a middle layer between user and AIS.

Web application provides access through 4 functions:

1. Identification
2. Authentication

   It is required to use two-way authentication, at the Web server level, to develop secure web application. Nowadays, the most secure technology of two-way authentication is authentication using Transport Layer Security (TLS) protocol.

   Using this technology requires a separate subdomain and configuring a Web server to work with TLS. Subdomain verifies a user's certificate. Server also must have a certificate that was certified by the certification agency.

   It is important to note that using a two-way authentication reduces probability of any type of attack. Two-way authentication occurs before user can open first HTML page. Thus, an attacker has no chance to interact with web application without a certificate. The attacker's actions will be stopped during the access to the web server by software which processes the connection. Therefore, web application will not even know about malicious actions.

After authentication at the TLS protocol level the server and the client begin to communicate through the user interface

3. Work providing by means of AIS

   It is necessary to restrict access to the AIS, the AIS access server and to local network by firewall. The firewall should be installed behind the outer. Such installation scheme creates a zone for the enterprise's local network enterprise and a demilitarized zone as shown in Figure 2.

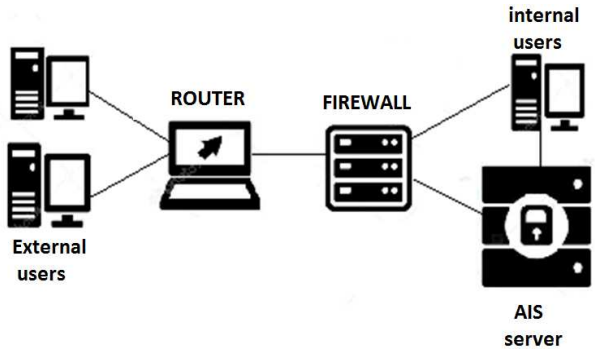


Figure 2
Wiring diagram of the firewall

4. Maintaining the History of Events

   Maintaining the history of events implies the following actions: generating event.

4 User Authentication Scheme at the Web Server Level using the TLS Protocol

It is required to use two-way authentication at the Web server level to build a secure Web application. It is required to build a separate subdomain, that will provide a user's certificate verification. Also, it is required to use a web server configuration that works with TLS. The web server should have a certificate.

The user authentication process at the web server level, according to RFC 5246 (The Transport Layer Security Protocol Version 1.2), is presented as a UML activity diagram in Figure 3 [8].

Two-way authentication occurs before the user can open the first HTML page. Thus, an attacker has no chance to interact with the web application without a certificate [9]. The attacker's actions will be terminated during the access to the

web server by the software that processes the connection. Therefore, the web application will not even be aware of the malicious actions.

After authentication at the TLS protocol level, the server and the client begin to communicate through the user interface.

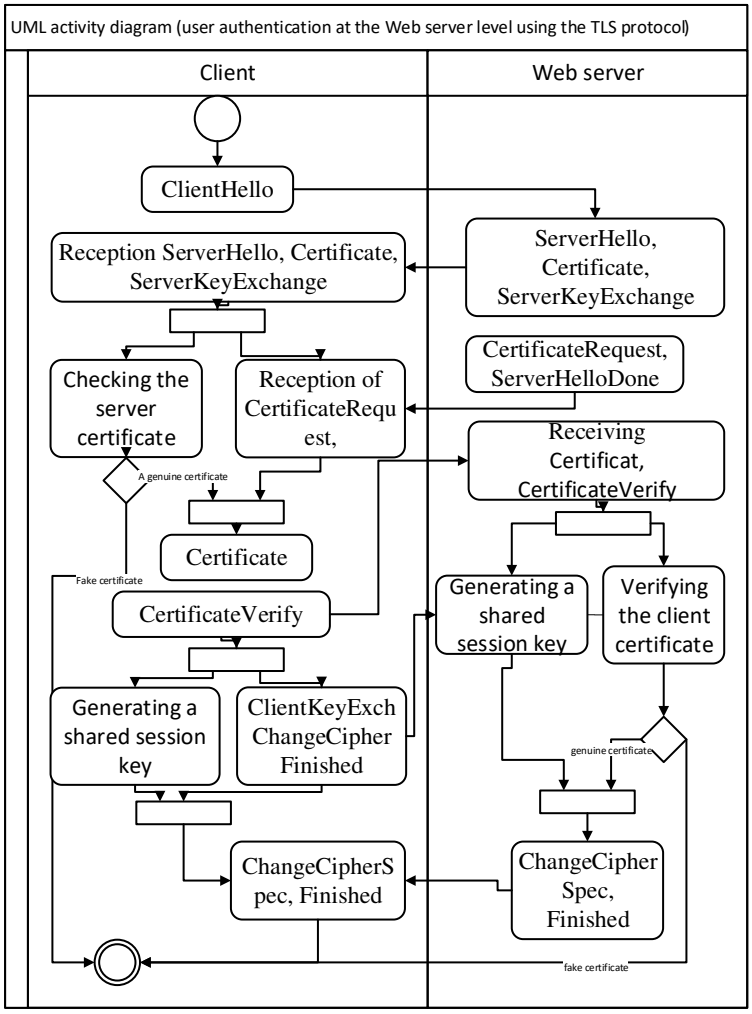


Figure 3
User authentication process at the Web server level

5 User-Level Authentication Scheme at the Application Level

The server and the client begin to communicate through the user interface after authentication at the TLS protocol level. The process of user authentication in the system (at the level of application logic) is presented as a UML activity diagram in following figure:

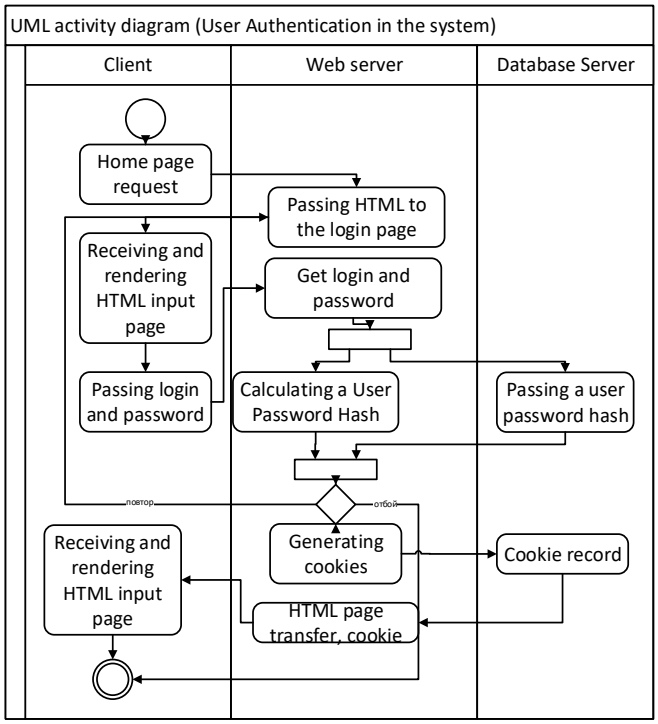


Figure 4
The process of user authentication in the system

6 The Scheme of Processing Requests from Remote Clients

The user can continue to work with the server. Authentication is at the web application level. It becomes possible because the cookie contains the identifier of the valid session. The process of processing user requests is presented in the form of the UML activity diagram, in Figure 5.

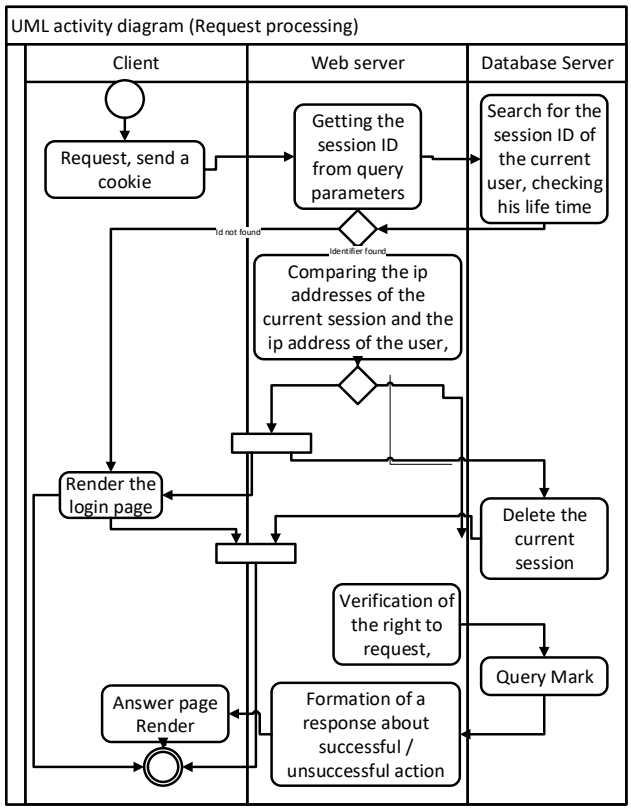


Figure 5
The process of processing user requests

7 Logical Data Model

In the process of developing the architecture of the web application, taking into account the general security requirements and requirements related to countermeasures against attacks on attacks, a scheme for application data was developed.

When analyzing all requirements, 5 entities were identified. Each entity and each data field has a semantic load and is introduced to provide the functionality of the application:

1) Users - an entity that stores user data associated with its identification and authentication. Fields of the entity:

a) user id - unique user identifier

b) email - the user's mail, a unique field on which it is possible to uniquely identify a user in the AIS. This field is also used to recover the password

c) encrypted_password - password hash of the user. The field is necessary for the requirement of secure data storage and the operation of the authentication algorithm

d) reset_password_token - a one-time token of the application for password recovery, which can be sent to the user's email account upon password recovery

e) reset_passwd_sent_at - time of creation of the previous field

f) sign_in_count - the field containing the number of moves of the user to the system

g) current_sign_in_at - the field containing the date and time of the current logon

h) last_sign_in_at - the field containing the date and time of the last logon

i) current_sign_in_ip - the field containing the current ip address

j) last_sign_in_ip - the field containing the ip address at the last login

k) failed_login_count - the field storing the number of failed inputs in a row

l) status - the field for storing the user status (active, blocked, the password reset is expected)

m) name - the field containing the user name

2) Roles - an entity necessary for grouping several actions with the subsequent assignment of action groups to specific users. Fields of the entity:

a) id-unique identifier of the role

b) name - the name of the role (for example, administrator, operator, extras, etc.)

3) User roles (UserRoles) - a table for communication between users and roles. Fields of the table:

a) id - identifier of the user role

b) user_id - user identifier

c) role_id is the role identifier

4) Actions - an entity that describes the actions that can be performed with a specific re-source:

a) id - the identifier of the action

b) api_url - url of the resource

c) methods - methods that can be implemented by an action

d) role_id is the role identifier

5) History - the entity for storing data about user actions:

a) id - the identifier of the history record

b) created_at - the date the record was created

c) user_id - user identifier

d) action_id - identifier of the user action

e) data - data that the user transmitted during the execution of the action

f) status - the status of the operation

For the interplay of the essence there were wings with each other by the following connections:

1) User - role - many to many

2) Role - action - one to many

3) Action - history - one to many

4) User - history - one to many

Conclusions

The anticipation of vulnerabilities in the WEB-application was carried out through the implementation of measures for the development of secure software, established by GOST R ISO / IEC 12207.

Modelling of security threats and detected vectors of possible attacks, as well as, their analysis, allows for the formulation of countermeasures for each of the vectors at the different architectural levels of the web application.

Countermeasures and indications to use protection technologies, with respect to the specific features of the architecture of the projected web application based on the Model-View-Controller pattern [10], made it possible to increase the safety of the use of AIS.

Therefore, increasing the security of access to a database containing critical information about the operated CCD GTE is based on the development of a secure web application architecture, that serves as an isolating layer for external AIS clients, which allows for the provision of transmission and analysis in ground service points of the operational parametric information, from the aircraft board and to provide the possibility of remote access to essential data.

Acknowledgement

This article is supported by RFBR grant № 20-08-00668

References

[1] O. S. Gurevich. Prospectives for the development of the SAU GTE in the work of the CIAM SSC / Abstracts of the International Scientific and Technical Con-ference "Aviadvigateli XXI Century" Moscow, CIAM, 2010, p. 838

[2] O. S. Gurevich, A. S. Trofimov, M. G. Kesselman, V. I. Chernyshov, A. A. Semin. Wireless Demonstration Control System for GTE / Abstracts of the International Scientific and Technical Conference "Aviadvigateli XXI Century" Moscow, CIAM, 2010, p. 812

[3] Mohsen Riahi, Manesh Naima Kaabouch Analysis of vulnerabilities, attacks, countermeasures and overall risk of the Automatic Dependent Surveillance-Broadcast (ADS-B) system International Journal of Critical Infrastructure Protection, 2017, Vol. 19, pp. 16-31

[4] Mohamed Slim Ben Mahmoud, Alain Pirovano Nicolas Larrieu Aeronautical communication transition from analog to digital data: A network security survey, Computer Science Review, Vol. 11-12, pp. 1-12

[5] G. I. Pogorelov, Yu. O. Bagaev. Features of the management of logistics and parame- ter data of the SAU GTE at the initial stages of operation / Abstracts of the international scientific and technical conference "Aircraft engines of the XXI century" Moscow, CIAM, 2010, p. 777

[6] E. Naderi, K. Khorasani Data-driven fault detection, isolation and estimation of aircraft gas turbine engine actuator and sensorsm, Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on, 2017, web [http://ieeexplore.ieee.org/document/7946715/] 10.10.2017

[7] E. Naderi, K. Khorasani Data-driven fault detection, isolation and estimation of aircraft gas turbine engine actuator and sensorsm, Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on, 2017, web [http://ieeexplore.ieee.org/document/7946715/] 10.10.2017

[7] Gerald Beuchelt Securing Web Applications, Services, and Servers Computer and Information Security Handbook (Third Edition) 2017, pp. 183-203

[8] T. Premalatha, S. Duraisamy A certificate based authorization and protected application layer protocol for IoT, Computer Communication and Informatics (ICCCI), 2017 International Conference on, 2017, pp. 1-5

[9] HassinaNacer, NabilDjebari, HachemSlimani, DjamilAissan A distributed authentication model for composite Web services, Computers & Security, 2017, Vol. 70, pp. 144-178

[10] S. S. Hasan, R. K. Isaac An integrated approach of MAS-CommonKADS, Model–View–Controller and web application optimization strategies for web-based expert system development, Expert Systems with Applications, 2011, Vol. 38, pp. 417-428

Routing in CNC Cutting Machines: Engineering Constraints

Alexander A. Petunin[1,2], Pavel A. Chentsov[2,1]

[1]Ural Federal University, 19 Mira Str., 620002 Ekaterinburg, Russia, a.a.petunin@urfu.ru

[2]Krasovskii Institute of Mathematics and Mechanics (IMM UB RAS), 16 S. Kovalevskaya Str., 620990 Ekaterinburg, Russia, chentsov.p@uran.ru

Abstract: The paper is devoted to the problems of cutting path optimization with engineering specifics. There are two severe constraints of thermal cutting: thermal expansion and sheet stiffness constraints. Our solution uses the dependence of the cost functions on the previous part of the route and on the direction of movement of the tool along the contour.

Keywords: general traveling salesman problem; cutting tool motion optimization; CNC cutting machines

1 Introduction

One of the main ways of metal processing is metal sheet cutting. The common problems of metal sheet cutting [1] represent serious practical interest. A detailed overview is given in paper [2]. The main problem is to get parts of the best quality with the minimal cost. The cost includes plant work time and used resources during cutting process (gas, electricity, etc.).

The common problems of metal sheet cutting (see [1]) consist of the nesting problem and searching of accessible piercing points and cutting path optimization (tool path problem). This work is devoted to the tool path problem and its special constraints. There are constraints for enclosure of contours, for thermal expansion and for sheet stiffness.

The tool path problem is a very complex problem. It unites discrete and continuous optimization. There is no mathematical formalization of the problem for the common case. In the paper [3] it was proposed the version of the problem as a generalized travelling salesman problem (GTSP) with the inclusion of contours. Work [4] has a similar problem statement, but another way of heuristic

solution (Tabu search). In the paper [5] the modification of the problem for the Leather cut process is shown. So, the existing mathematical models and algorithms for solving the problem do not consider most of the real technological constraints of the cutting process. In particular, it concerns the constraints connected with thermal deformations of parts at thermal cutting. But only the class of problems which are reduced to the minimization of idling motion of the cutter are considered. Some technological constraints of the cutting process and solving of tool path problem are described in [6].

The tool path optimization problem can be reduced to GTSP by adding new constraints and adding contour visit parameters. The constraints are precedence conditions, thermal and rigidity constraints. The tool motion direction is a parameter.

There are versions of GTSP with fixed permitted motions from point to point and the exact branch-and-cut algorithm and heuristic algorithm are given in [7].

The exact dynamic programming-based algorithm for a routing problem depending on the list of tasks' costs was proposed in [8]. The dinamyc programming fitted for CNC cutting machine tool routing is shown in [9].

The basic theory of tool routing optimization for CNC cutting machines is given in [10], the heuristic algorithm of problem solution is proposed in [11].

In [12] there are proposed heuristic methods of GTSP solving. It is a genetic algorithm with local improvement of solution. The same process was used in [13].

In the paper [14] we can see a hard metric heuristic algorithm and significant comparison with existing algorithms. The comparison shows the high efficiency of the algorithm.

This work is a continuation of [15]. The main difference is cost functions with using of rigidity constraint by penalties.

2 Mathematical Formulation

The steel sheet of fixed size and planned within it contours of the cutting details are given. By p^{st} and p^{fin} denote a start and finish points of the cutting tool motion ($p^{st} \in R \times R$, $p^{fin} \in R \times R$, R – real line). N – number of contours. V_{idling} – idling speed of the tool. V_{cut} – work speed. The piercing points must be realized out of the detail contour curves, because it leads to strong temperature expansion. Moreover, the hole in the point of insert can exceed cut width. Point $p^{on}, p^{on} \in R \times R$, is a piercing point. The cutting tool moves on work speed

V_{cut} to the point of the enter in the contour $p^{cut}, p^{cut} \in R \times R$, and performs cut of the contour with finish in the point p^{cut}. The cut can be made clockwise or counter clockwise. A cutting tool switch-off can damage a detail edge, and cutting tool moves on work speed to the cut-off point $p^{off}, p^{off} \in R \times R$. The point p^{off} placed out of the contour curve. In the point p^{off} the cutting tool realizes switch-off, then starts motion on idling speed to the next contour. These points are shown on Figure 1.

There are thermal restrictions, requiring that near the finishing line of the contour cut there must be enough metal. It's shown in Figure 2.

The rigidity restriction, requiring that around large parts must be enough metal. It's shown in Figure 3.

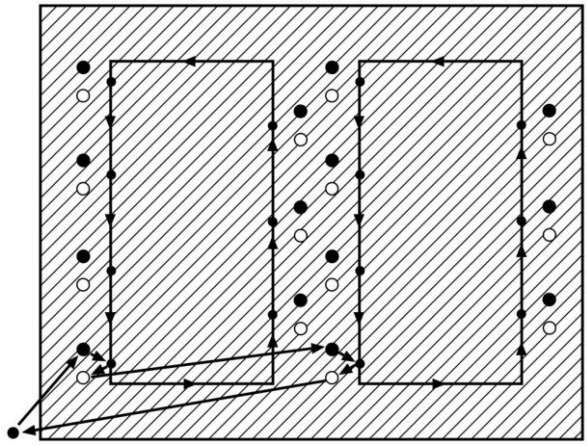


Figure 1
The cut points

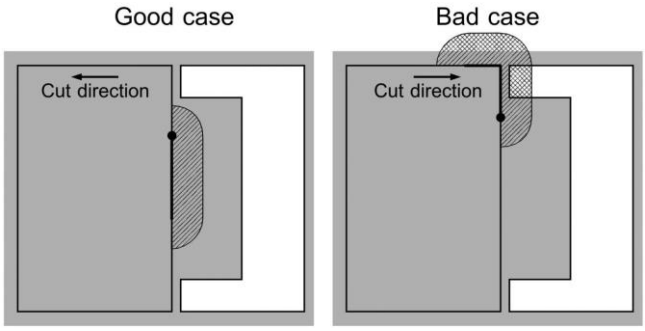


Figure 2
Thermal constraints

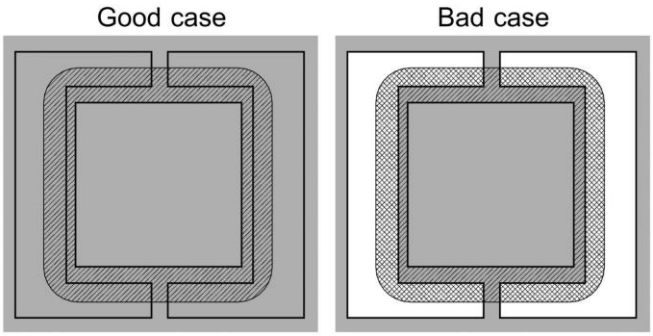


Figure 3
Rigidity constraints

Every contour cut begins and ends from the point triplet – p^{on}, p^{cut}, p^{off}. The number of these triplets is $\tilde{N}$. Set $M, M = \overline{1, \tilde{N}}$, is the set of all triplet indexes. Let $M_1, ..., M_N$, $N \geq 2$, be sets of the contour point triplet indexes for contours. Each contour has its own number of the cut point triplets: $|M_i| = n_i$, $n_i \geq 1$. Note that

$$\bigcup_{i \in \overline{1,N}} M_i = M,\ M_i \cap M_j = \varnothing\ \forall i \in \overline{1,N}, \forall j \in \overline{1,N} \setminus \{i\}.$$

For the definition of the cut points denote three functions:

$$p^{on} : M \to R \times R,$$

$$p^{cut} : M \to R \times R,$$

$$p^{off} : M \to R \times R.$$

By A we denote the family of the ordered sets. If $a \in A$, then

$$a = (a_i)_{i \in \overline{1,N}} : \overline{1,N} \to M, \forall a_1 \in a, \forall a_2 \in a \setminus \{a_1\}, \neg\exists i \in \overline{1,N} :$$

$$(a_1 \in M_i)\ \&\ (a_2 \in M_i)$$

For $a \in A$ the route of cutting tool motion is

$$p^{st} \to p^{on}(a_1) \to p^{cut}(a_1) \to p^{off}(a_1) \to ...$$

$$... \to p^{on}(a_N) \to p^{cut}(a_N) \to p^{off}(a_N) \to p^{fin}.$$

For the definition of the internal cost (thermal restriction penalties), we need the direction sign $d, d \in \{0,1\}$. Value 0 corresponds to counter-clockwise contour cut process, 1 – clockwise.

By D denoting the ordered set of contour tool motion directions for route: if $d \in D$, then $d = (d_i)_{i \in \overline{1,N}} : \overline{1,N} \to \{0;1\}$.

If $d_i = 1$, then cut of the contour cut in route position i direction is clockwise, if $d_i = 0$ – counter clockwise.

The speeds of the tool are significant parameters used in the formulas of cost functions. By V_{idling} denoting idling speed. By V_{cut} denoting work speed. An estimation of the idling motion from point to point is the motion time. If x_1 and x_2 are start and finish points of the switched-off cutting tool local motion (p^{st}, p^{fin}, p^{on} or p^{off}), then this motion can be evaluated by

$$C(x_1, x_2) = \frac{\rho(x_1, x_2)}{V_{idling}}.$$

In this formula ρ is Euclidean distance between points x_1 and x_2.

An estimation of contour $i, i \in \overline{1,N}$, cutting with cut points triplet $m, m \in M_i$ and direction d and cut out at this time contours set $K, K \subset \overline{1,N} / \{i\}$:

$$c(m,d,K) = \frac{\rho(p^{on}(m), p^{cut}(m))}{V_{cut}} + \frac{\rho(p^{cut}(m), p^{off}(m))}{V_{cut}} + \tag{1}$$
$$c_d(m,d,K) + \tilde{c}_d(m,K).$$

Contour cut time is excluded from the estimation because it must be made once, and does not depend on route.

Function c_d is penalty function for cut finishing by location of finish cut area with respect to holes in metal from cut out contours at this time. Note that for final result estimation $c_d \equiv 0$, it is needed only when counting.

Function c_d is defined as follows:

$$c_d(m,d,K) = \frac{S_m^*(m,d,K)}{S_m(m,d,K)} P$$

$S_m(m,d,K)$ - square of checking area near the finishing line of contour $i, i \in \overline{1,N}$, if cut makes with using triplet $m \in M_i$. Checking area near the finishing line of contour further will be referred as cut finish area. $S_m^*(m,d,K)$ – square of intersection area of cut finish area with cut out parts or out of sheet space (for detail see figure 2); P – penalty coefficient. The method of S_m and S_m^* computation is shown in [17].

Function $\tilde{c}_d$ is penalty function for spaces around parts (rigidity constraint). It is meaningful for the large and heavy parts. So, as it shown in Figure 3, large parts must have enough metal around they external contours. If $m \in M_i$, then function $\tilde{c}_d$ is defined this way:

$$\tilde{c}_d(m,K) = \sum_{j \in \overline{1,N} \setminus (K \cup \{i\})} \tilde{P}_1 \left(\frac{\hat{S}(j)}{\tilde{P}_2}\right)^2 \times \left(1 - \frac{\tilde{S}^*(j, K \cup \{j\})}{\tilde{S}(j)}\right) \times T(j) \qquad (2)$$

$\hat{S}(j)$ - square of metal within contour with index $j, j \in \overline{1,N}$ (without square of internal contours). $\tilde{P}_1$ - rigidity constraint weight coefficient. $\tilde{P}_2$ - coefficient of part weight checking. $\tilde{S}(j)$ - square of checking external space near contour with index $j, j \in \overline{1,N}$ (further external contour area). $\tilde{S}^*(j,K)$ - square of external contour area for contour $j, j \in \overline{1,N}$, intersection with cut out contours and out of sheet space.

$T(j)$ for contour $j, j \in \overline{1,N}$, has value 1 if contour j is the external contour of the part, and 0 otherwise.

In every step of route construction, $\tilde{c}_d(m,K)$ gives information on the correctness of the arrangement of the remaining contours with respect to the spaces in the metal and the space outside the sheet.

Note that for final result estimation, $\tilde{c}_d \equiv 0$, it is needed only when count as in case of c_d function.

This problem has precedence conditions. In this paper a complete cutting of each contour is considered. Cut out contours drop down and became inaccessible to any further processing. In some situations cut out parts have support from below and stay accessible to processing. But they lose connection with the sheet and can have unaccounted by plant displacement. This leads to inaccuracies in the cut. So, the reason for the precedence conditions – internal contours must be cut out before external.

It should be said that if we will consider situations with partial contour cut, it allows for making the cut of the internal contours after the partial cut of the external contours.

Let Z be the set of address pairs z for precedence condition:

$$z = (z_1, z_2), z_1 \in \overline{1,N}, z_2 \in \overline{1,N}, z_1 \neq z_2, |Z| \geq 0.$$

For $z \in Z$ contour with index z_1 must be visited before contour with index z_2.

The set Z makes some routes from A inaccessible. We need to define the family of accessible routes $\tilde{A}$, $\tilde{A} \subset A$: $\forall a \in \tilde{A}$, $\forall z \in Z$ if $a_i \in M_{z_1}$ and $a_j \in M_{z_2}$ then $i < j$.

If $a \in \tilde{A}$ and $d \in D$, then it's estimation is

$$T(a,d) = C(p^{st}, p^{on}(a_1)) + c(a_1, d_1, \varnothing) +$$
$$\sum_{i=2}^{N} \left(C(p^{off}(a_{i-1}), p^{on}(a_i)) + c(a_i, d_i, \bigcup_{j \in \overline{1,i-1}} a_j) \right) + C(p^{off}(a_N), p^{fin}).$$

The main task is

$$T(a,d) \to \min, \ a \in \tilde{A}, d \in D. \tag{3}$$

3 Computation of Area Squares for Rigidity Constraint

The way of S_m and S_m^* computation for thermal constraints is shown in [17]. It is made by using special area matrixes in not an exact way but made permissible for this problem's precision.

The same method is used for $\hat{S}(j), \tilde{S}(j)$ and $\tilde{S}^*(j,K)$ computation.

The steel sheet and space around the sides splits in square areas – it is matrix cells. If coordinates of the cell are located out of the sheet, it links with value -2. If the cell locates out of any contours, but within the steel sheet, it links with value -1. If the cell intersects with some contour or locates within it, this cell links with this contour index (see Figure 4).

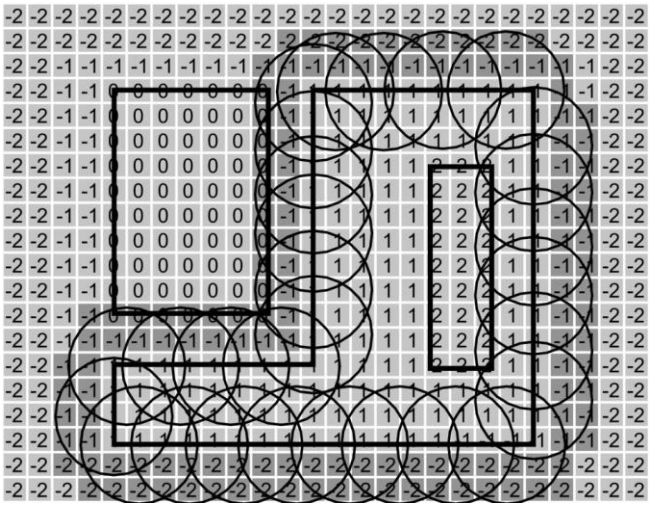


Figure 4
Rigidity area squares count

Count of the $\hat{S}(j)$ makes this way: for the contour with index $j, j \in \overline{1,N}$ we need to count the number of cells with value j, and multiply it by the cell square.

The following describes the method of $\tilde{S}(j)$ computation for the contour with index $j, j \in \overline{1,N}$. Locate N_L points on the contour on an equal range. Make a circular area around every point, and check intersections with area matrix cells. Calculate the number of cells, which value differs from i (in Figure 4 it is 1) and indexes of its nested contours (in Figure 4 it is 2). Every cell must be checked once. Multiply this value by square of the cell. It gives $\tilde{S}(j)$.

The following describes the method of $\tilde{S}^*(j,K)$ computation for the contour with index $j, j \in \overline{1,N}$, and cut out at this time contours set $K, K \subset \overline{1,N}$. Locate N_L points on the contour on equal range. Make a circular area around every point, and check intersections with the area matrix cells. Calculate the number of cells, which value is -2 or exists in the set K and not an internal contour of contour with index j. Every cell must be checked once. Multiply this value by the square of cell. It gives $\tilde{S}^*(j,K)$.

4 Algorithms for Problem Decision

To solve problem (3) use exact dynamic programming (DP) method based algorithm from [9]. The heuristic algorithm of problem solving was proposed in [17], and is usable in this case too. This algorithm can be used with the new rigidity constraint because this constraint is implemented by penalty (2) in cost function (1), and does not make any improvements directly in the algorithms.

The exact method form [9] can be used for small samples with the number of contours less then 27. The heuristic algorithm allows for obtaining route for hundreds of contours. It is usable in most real industrial cuttings.

5 Computing Experiments

Calculations were made on the computer with the Intel i7-2630QM processor, 8 GB memory and operating system Windows 7 (64-bit).

The program is developed with using C++, in the development environment Microsoft Visual Studio 2013.

The samples contain parts of different sizes. It allows for showing the work of algorithms with using new cost functions.

Full information on coordinates of contours, cut points, and cut sequence is not included due to space economy reasons. Even for the first small sample, it requires a lot of text pages;

The external contour area width value is 200 mm. This width seems to be enough for making good rigidity for most large parts, but the search of the most appropriate value of this parameter can be a topic for the new research. Coefficient values: $\tilde{P}_1 = 1,\ \tilde{P}_2 = 1000000$. These parameters were used for all samples.

Example 1. Parameters: $N = 16$, $|Z| = 4$.

The first count was made with DP exact method with a counting duration of 21 seconds. Value of estimation is 64,82. The route is shown in Figure 5.

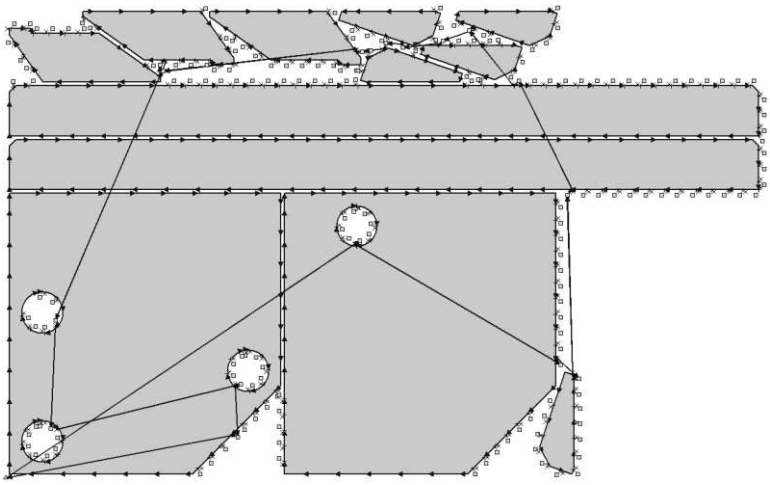


Figure 5
Result for DP count, sample 1

Second count was made with heuristic method. Counting duration 4 seconds. Value of estimation 66,88. The route is shown on Figure 6.

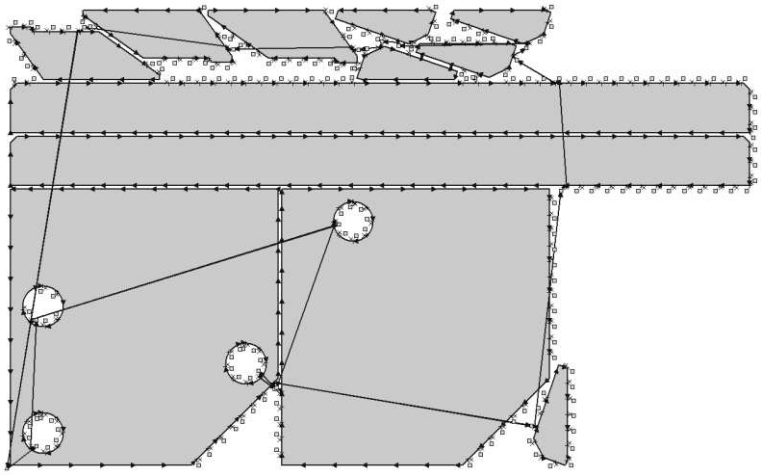


Figure 6
Result for heuristic count, sample 1

Example 2. Parameters: $N = 86$, $|Z| = 29$. To obtain decision the heuristic method (to many contours for DP) was used. The counting duration was 13 seconds. For computation aiterative process was used with a total iterations count of 10. If count time seems to long, it can be reduced to a lower value (less then 10). Value of estimation was 295,33. The route is shown on Figure 7.

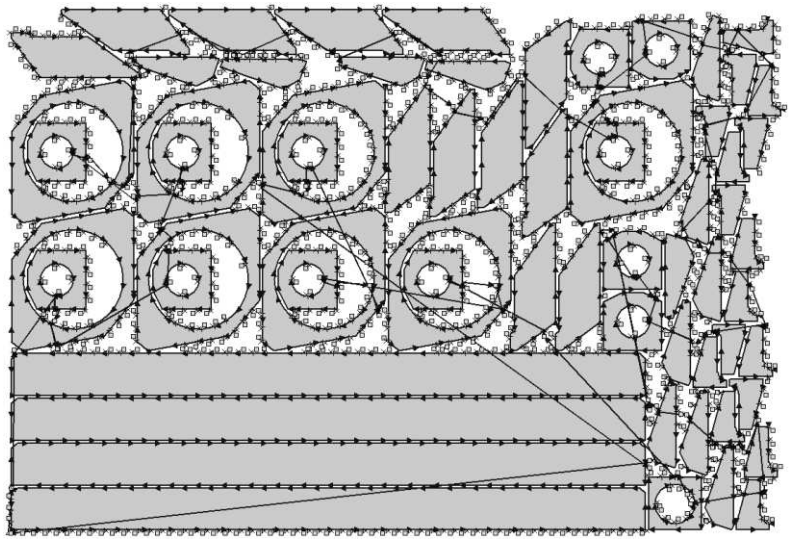


Figure 7
Result for heuristic count, sample 2

So, we can see that the performance of the algorithm allows for calculation of tasks of large dimensions. Is makes the algorithm applicable in manufacturing.

The heuristic algorithm does not gives an exact decision, but the formulation of problems is based on our assumptions about the quality of the cut. We take coefficient and parameter values by our mention on the basis of our ideas about the physical process. This assumptions can differ with real physical process features. In this situation the exact decision of the formulated tasks can be difficult to obtain real cut quality. This circumstance makes the application of exact algorithms not so important, although desirable.

Conclusions

Some problems of tool routing for CNC sheet cutting machines are considered. Complex specific engineering constraints are constructed. The method of constraints computation is proposed.

New constraint is fitted to use with existing exact and heuristic algorithms of tool path optimization problem.

Several computing experiments were performed. The resulting routes look good in respect to constraints.

Acknowledgement

This work was supported by Russian Foundation for Basic Research, grant 20-08-00873.

References

[1] A. A. Petunin: Development of CAM-system for sheet cutting machines as an innovation example, Innovative information technologies: Theory and practice. International scientific edition: materials of the International workshop (Karlsruhe – Ufa – Dresden, April 8-13, 2011), 2011, Ufa, pp. 47-50

[2] Petunin A. A., Stylios C. Optimization models of tool path problem for CNC sheet metal cutting machines. IFAC-PapersOnLine. - 2016 - V. 49, l. 12. - pp. 23-28

[3] K. Castelino, R. D'Souza, P: Wright Tool-path optimization for minimizing airtime during machining, Journal of Manufacturing Systems 22 (3), 2002, pp. 173-180

[4] Dewil, R., Vansteenwegen, P., Cattrysse, D. (2016) A review of cutting path algorithms for laser cutters. Int J Adv. Manuf. Technol. 87:1865-1884

[5] W. B. Yang, Y. W. Zhao, J. Jie, W. L. Wang: An Effective Algorithm for Tool-Path Airtime, Optimization during Leather Cutting. Advanced Materials Research. 2010, 102-104, pp. 373-377

[6] Petunin, A. (2009) About some strategies of the tool path modeling at the control programs generation for the flame cutting machines. Vestnik UGATU, Vol. 13, 2(35), 280-286

[7] M. Fischetti, J. J. Salazar-Gonzalez, P. Toth: A Branch-and-Cut algorithm for the symmetric generalized traveling salesman problem: Operations Research. 1997, 45 (3), pp. 378-394

[8] Chentsov, A. G., Chentsov, A. A. Dynamic programming in the routing problem with constraints and costs depending on a list of tasks. Dokl. Math. 2013, 88, 637-640

[9] A. G. Chentsov, P. A. Chentsov, A. A. Petunin, A. N. Sesekin: Routing problems: constraints and optimality, IFAC-PapersOnLine. 2016, Vol. 49, Issue 12, pp. 640-644

[10] Chentsov, P. A. About one application of greedy algorithm for the solution of the generalized salesman problem. Proceedings of the 16h International Workshop on Computer Science and Information Technologies. Sheffield, England, 2014, Vol. 2, pp.148-151

[11] A. A. Petunin, A. G. Chentsov, P. A. Chentsov: About routing of tool motion in the sheet cutting machines with numerical program control, Scientific and technical sheets SPbGPU. Informatics. Telecommunications. Management, 2013, 2 (169), pp. 103-111

[12] L. V. Snyder, M. S. Daskin: A random-key genetic algorithm for the generalized traveling salesman problem, European Journal of Operational Research, 2006, 174, pp. 38-53

[13] J. Silberholz, B. Golden: The Generalized Traveling Salesman Problem: a new Genetic Algorithm approach. Extending the Horizons: Advances in Computing, Optimization, and Decision Technologies, 2007, 37, pp. 165-181

[14] G. Gutin, D. Karapetyan. A Memetic: Algorithm for the Generalized Traveling Salesman Problem, Natural Computing, 2010, 9(1), Springer, pp. 47-60

[15] P. A. Chentsov, A. A. Petunin: Tool Routing Problem for CNC Plate Cutting Machines, IFAC-PapersOnLine. 2016, Vol. 49, Issue 12, pp. 645-650

An Effective Scheduling Method in the Cloud System of Collective Access, for Virtual Working Environments

Leonid V. Legashev, Irina P. Bolodurina

Orenburg State University, Pobeda avenue 13, 460018 Orenburg, Russia,
ais@mail.osu.ru, prmat@mail.osu.ru

Abstract: This paper describes the cloud system of collective access to virtual working environments as a means of providing remote access to paid and free software, for students of educational institutions in secondary education. The problems of efficient cloud system scheduling and optimizing the usage of cloud virtual machines and paid software licenses has been studied in detail, herein. In addition, the mathematical model of cloud system resources control is presented, as well as, the functional model. Two evolutionary scheduling algorithms are proposed. Also, the UML class diagram of the cloud system simulator is described. The statistical analysis of the fitness function value distribution, to evaluate the proposed algorithms, is also performed.

Keywords: desktop as a service; cloud computing; scheduling; simulated annealing; genetic algorithm, statistical analysis

1 Introduction

Currently, cloud computing is one of the most popular technology solutions for data processing, various calculations on cloud servers, and services providing. Many educational institutions of secondary education are facing problems with regular renewal of the computer park and software purchase, installation and configuration. Taking into account the existing restrictions on the number of software licenses and the duration of user's sessions, the development of efficient methods of using the resources of the cloud system is an urgent task to be solved.

One of the solutions is to create a cloud system of collective access (CSCA) to virtual working environments [1], as shown in Figure 1. This system is based on the DaaS (Desktop-as-a-Service) scheme [2]. Users get access to virtual desktops with installed software via Internet browser. Each user forms one request from the educational organization. The user specifies a required number of virtual machines (VM) instances, choose required software from the list of available software and

sets a working schedule of organization. Data Storage System (DSS) contains the VM images and license servers for paid software. A materials database stores educational-methodical materials information, concerning the subjects of the school courses. Resources administration module is used to find a close-to-optimal schedule.

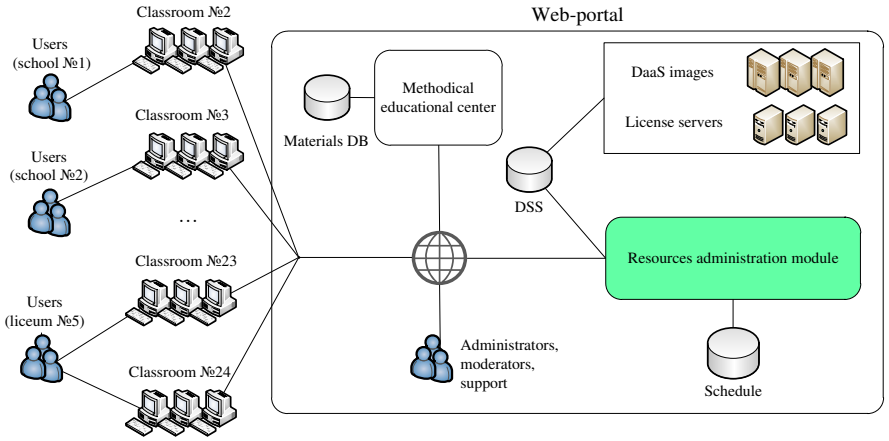


Figure 1
CSCA architecture

Administrators and moderators provide the technical support. To experience CSCA functionality, the end users only needed low-cost computers with Internet access. Collective usage of cloud system leads to reduced costs of the software and computer components purchase and leads to opportunity to use modern software in the educational process.

There is a great deal of research on scheduling problems, virtual machine allocation and cloud technologies using in the educational process. In paper [3], a scheduling algorithm that optimizes virtual machine placement across multiple clouds is described. The purpose of the algorithm is to allocate n virtual machines between m available clouds improving criteria such as performance, cost, or load balancing. The authors of paper [4] considered a problem of increasing energy consumption in a cloud server environment. The paper presented three heuristic scheduling algorithms of virtual machines allocation to minimize energy consumption in the memory system. An algorithm based on genetic programming as a meta-heuristic method for solving static scheduling problem in heterogeneous computing systems proposed in [5]. The authors of paper [6] proposed the usage of cloud computing technologies in the business processes of universities. The significant impact of cloud computing on input-output data quality dimensions of universities collaborative processes is shown. The authors of paper [7] described load balancing of virtual machines resources in cloud environments. Authors proposed scheduling strategy of virtual machines placement based on genetic

algorithm. Software as a Service (SaaS) model in the Chinese education system is described in paper [8]. Students get access to the office suite of applications, libraries and cloud applications (such as GoogleApps and Zoho Office). The disadvantage is in the usage of open-source software only. The authors of paper [9] introduced background and fundamentals about emerging of three technology paradigms – Cloud Computing, Internet of Things, and Connectivism – in the educational sector. Analysis of the existing literature shows the lack of efficient solutions for the CSCA resources administration within the constraints of corresponding time slots and software licenses.

It should be noted, there are many solutions for implementing remote access to information resources. Amazon company offers its own solution Amazon WorkSpaces, which provides Desktop as a Service mechanism and allocates virtual desktops for users based on Microsoft Windows operating systems. Amazon WorkSpaces is not suitable as a turnkey solution of the problem due to the need to set up a virtual desktop for each user individually. Two alternative solutions are offered based on Microsoft Azure Remote Desktop Services – MyCloudIT platform and Workspot solution. Virtual desktops based on MyCloudIT and Workspot come with pre-installed software (Microsoft Office, Outlook, Skype, Adobe Reader, Mozilla Firefox, Internet Explorer). We can also mention Citrix XenDesktop, VMware Horizon, Quest vWorkspace as existing platforms for implementing the Desktop as a Service scheme. They are not suitable too as a turnkey solution of the problem for several reasons: high cost, limited ability to function in a browser, the usage of simplistic administration algorithms, the lack of software license constraints and collective usage in the specifics of educational institutions. TSPlus Server [10] technology is chosen to implement the DaaS mechanism for our CSCA. The main advantages of using the cloud systems are as follow:

- Economic benefits. Organizations do not need to buy expensive paid software and modern computers. They rent only actual services of the cloud system.
- End-user data storage. Various user data, which is necessary for work (documents, files, projects, databases, etc.), are stored in the network storage of the cloud system.
- Broad access to educational software. A larger list of paid and open source software is available for students of educational institutions to perform virtual experiments, mathematical calculations, modelling, machine learning, rendering, etc.

Implementation of the CSCA faces the NP-complete problem of requests placement in the schedule within the constraints of corresponding time slots and limited software licenses.

2 Mathematical Model of Cloud System Resources Control

The cloud system mathematical model is developed to describe the basic stages of operation of the system and to provide DaaS model experience to end-users.

Cloud system of collective access to virtual working environments is formalized as tuple $C_{cloud} = \{Z_{temp}, S_{soft}, R_{req}, U_{users}, F_{flav}, D_{data}, \hat{S}\}$, where its elements are defined as follows: $Z_{temp} = \{Z^1_{temp}, Z^2_{temp}, ..., Z^i_{temp}\}$, $i \in N$ – the set of scheduling templates, $S_{soft} = \{S^1_{soft}, S^2_{soft}, ..., S^j_{soft}\}$, $j \in N$ – the set of accessible software, $R_{req} = \{R^1_{req}, R^2_{req}, ..., R^n_{req}\}, n \in N$ – the set of requests, U_{users} – the number of users, $F_{flav} = \{F^1_{flav}, F^2_{flav}, ..., F^m_{flav}\}, m \in N$ – the set of virtual machines flavors, $D_{data} = \{D^1_{data}, D^2_{data}, ..., D^p_{data}\}$, $p \in N$ – the set of datacenters, $\hat{S}$ – the set of schedules.

Schedule template is defined as tuple $Z^i_{temp} = \{Z^i_{title}, Z^i_{length}, Z^i_{\text{intervals}}\}$, where Z^i_{title} is the template title, $Z^i_{length} \in N$ – the duration of working session in minutes, $Z^i_{intervals} = \{[z_r, z_r + Z^i_{length}] | z_r \in \{0,1,...,1439\}, r \in N\}$ – the set of schedule time slots, z_r – the beginning of the working session. The interval [0,1439] corresponds to a twenty-four-hour time segment (00:00-23:59).

Free and paid software of the CSCA is defined as tuple $S^j_{soft} = \{S^j_{OS}, S^j_{title}, S^j_{license}, S^j_{install}, S^j_{delete}\}$, where its elements are defined as follows: S^j_{title} – the software title, $S_{OS} \in \{OS_1, OS_2, ..., OS_o\}, o \in N$ – the supported operating system, $S^j_{license}$ – the number of licenses, $S^j_{install}$ – the software installation time, S^j_{delete} – the software uninstallation time. Virtual machine flavor is formalized as tuple $F_{flav} = \{C_m, M_m, D_m\}, m \in N$, where C_m – is the number of CPU cores, M_m – the RAM size in Gb, D_m – the disk volume in Gb.

Each user request is defined as tuple $R^k_{req} = \{Z^{i'}_{temp}, R^k_{count}, \hat{S}^k_{soft}, t^k_{arrive}, T_k, \hat{T}_k, w_k, R^k_{status}\}, k = \overline{1,n}$, where its elements are defined as follows: $Z^{i'}_{temp} \in Z_{temp}, i' = 1,..,i$ – the chosen schedule template, R^k_{count} – the number of virtual machines, $\hat{S}^k_{soft} \subseteq S_{soft}$ – the subset of software,

t^k_{arrive} – the request arrival time, $T_k \subseteq Z_{intervals}$ – the required time slots, $\hat{T}_k \subseteq Z_{intervals}$ – the picked time slots, w_k – the weight coefficient, $R^k_{status} = \{-1,0,1\}$ – the request status. R^k_{status} is equal to "1" in case, when user request sets in the schedule at required time, $\hat{T}_k \equiv T_k$. Otherwise, R^k_{status} is equal to "0", and user get access to the virtual machines at different time, $\left(\hat{T}_k \neq T_k\right) \wedge \left(\left|\hat{T}_k\right| = \left|T_k\right|\right)$. If there is no possibility to satisfy user request within current constraints of cloud system, then R^k_{status} is equal to "–1", and $\hat{T}_k = \varnothing$. Cloud system schedule *S* is represented as follow:

$$S = \begin{vmatrix} R^1_{status} & \hat{T}_1 & \hat{S}^1_{soft} & R^1_{count} & T_1 & \hat{Z}^1_{temp} \\ R^2_{status} & \hat{T}_2 & \hat{S}^2_{soft} & R^2_{count} & T_2 & \hat{Z}^2_{temp} \\ R^3_{status} & \hat{T}_3 & \hat{S}^3_{soft} & R^3_{count} & T_3 & \hat{Z}^3_{temp} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ R^n_{status} & \hat{T}_n & \hat{S}^n_{soft} & R^n_{count} & T_n & \hat{Z}^n_{temp} \end{vmatrix}. \tag{1}$$

Each row of schedule *S* corresponds to single request R_{req}. Generated requests added to the queue with FCFS (First-Come, First-Served) service disciplines. Scheduler maximizes value of fitness function, which is described by equation (2):

$$F(S) = \sum_{k=1}^{n} w_k x_k(S) \to \max_{S \in \tilde{S}}, \tag{2}$$

$$\forall S \in \tilde{S},\ \forall k = \overline{1,n}: \quad x_k(S) = \begin{cases} \alpha \cdot R^k_{count}, \text{ if } \hat{T}_k \equiv T_k, \\ \beta \cdot R^k_{count}, \text{ if } \left(\hat{T}_k \neq T_k\right) \wedge \left(\left|\hat{T}_k\right| = \left|T_k\right|\right) \\ -\gamma \cdot R^k_{count}, \text{ if } \hat{T}_k = \varnothing. \end{cases}$$

Constraints on paid software licenses are described by the following inequalities:

$$\forall j = 1..|S_{soft}|, \forall l = [t_{start}, t_{end}]: G_{S,j,l} \leq S^j_{license}, \tag{3}$$

where $G_{S,j,l}$ is the number of virtual machines running the j^{th} software at the l^{th} time interval according to the schedule *S*, t_{start} – the launch time of virtual machines; t_{end} – their shutdown time.

Constraints on the number of placed virtual machines are described by equation (4):

$$\forall k = \overline{1,n}: H(R^k_{req}) = R^k_{count}, \tag{4}$$

where $H(R^k_{req})$ is the number of placed VMs.

Functional model (Figure 2) of the CSCA is based on mathematical model. It describes the main operations within the cloud system from the moment of requests generation to the moment of virtual machines assignment.

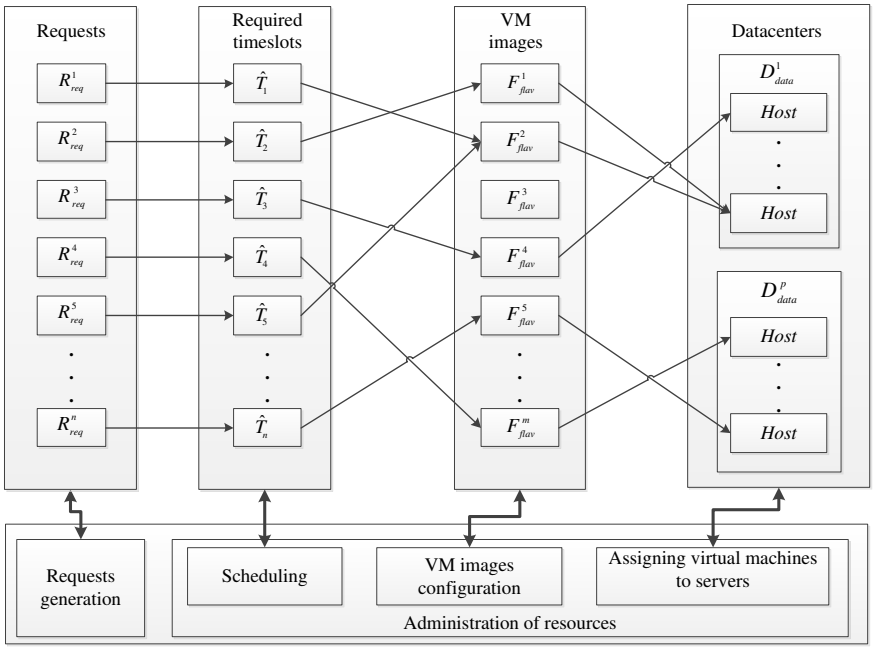


Figure 2
Functional model of the CSCA

Administration of cloud system resources denotes the sequential solution of three tasks:

- Virtual working environments scheduling. End-user requests R_{req} are placed to the schedule *S* to maximize the fitness function value (2) under constraints (3) and (4).
- Dynamic creation of virtual working environments F_{flav} according to the current schedule of CSCA.
- Selection of the physical servers of the cloud provider for launching the required virtual working environments at the specified time periods.

To satisfy constraints (3) and (4) during scheduling operation we define the schedule matrix *SWCheck* and initialize it with zeros (Figure 3). Each row of *SWCheck* corresponds to single software, and each column – to a single minute of time.

$$SWCheck = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{vmatrix}$$

Figure 3
Schedule matrix initialization

For each request in queue in all required time slots, the R^k_{count} virtual machines are adding to corresponding elements of matrix *SWCheck*, as it shown in Figure 4.

$$SWCheck = \begin{vmatrix} 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 10 & 10 & 10 & 10 & 10 & 17 & 17 & 17 & 17 & 17 & 17 & 17 & 17 & \dots & 0 \end{vmatrix}$$

Figure 4
Schedule matrix filling

If the total number of licenses of some software exceeds the maximum number of licenses for a certain period of time (Figure 5), the user virtual machines are subtracted from the schedule matrix.

$$SWCheck = \begin{vmatrix} 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & 15 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & \dots & 0 \\ 12 & 12 & 12 & 12 & 23 & 23 & 23 & 23 & 23 & 11 & 11 & 11 & 11 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 10 & 10 & 10 & 10 & 10 & 17 & 17 & 17 & 17 & 17 & 17 & 17 & 17 & \dots & 0 \end{vmatrix}$$

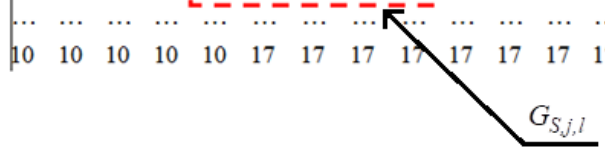


Figure 5
Excess of maximum number of licenses for 5 minute intervals

3 Scheduling Algorithms

3.1 Simulated Annealing Technique

The scheduling problem is NP-complete. We cannot use exhaustive search methods to solve the problem under consideration due to its exponential complexity $O(n^k)$ (where k is the number of requests, n – the number of combination without repetitions from all time slots of the schedule template by the number of required time slots). We have implemented two evolutionary algorithms: Simulated Annealing (SA) and Genetic Algorithm (GA), to solve this problem in polynomial time. The initial schedule is created by using the Round-Robin (RR) method of cyclic load distribution which distributes requests over the initial schedule in a uniform manner, and the primary assignment of request statuses is occurred.

The Simulated Annealing is a stochastic algorithm, which allows finding a close-to-optimal, solution of the scheduling problem and avoids a local maximum of the fitness function value (2). The parameters of the SA algorithm are presented in Table 1.

Table 1

Simulated annealing algorithm parameters

№	Parameter	Value
1	Maximum temperature	t_{max}=100
2	Temperature step	t_{step}=0.9
3	Mutation probability of schedule	0.01
4	Temperature breakpoint	0.05
5	Reward/penalty coefficients	$\alpha,\ \beta,\ \gamma$

In each iteration of the SA algorithm, the current schedule mutates randomly, and the value of fitness function is calculated. The current schedule is accepted as the best solution if the fitness function value is greater than the previous one. In case when the fitness function value is less than the previous one, the current schedule is also accepted as the best solution with acceptance probability function (5):

$$P = e^{-\left|\frac{F(s_{current})-F(S_{best})}{t_{curr}}\right|}. \quad (5)$$

At the first iterations of the SA algorithm, the acceptance probability function (5) will be quite high. As a consequence, the “worst” schedule, will often be chosen as an optimal schedule. This feature prevents the SA algorithm from becoming stuck at a local maximum of the fitness function value (2).

3.2 Genetic Algorithm

Another way to solve the scheduling problem is to use genetic algorithms. Each individual in the genetic algorithm represents one of the schedule options. We can depict it as columns of request statuses (Figure 6).

$$\begin{vmatrix} 1 & -1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & -1 \\ -1 & 1 & -1 & \cdots & -1 \\ 1 & 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & -1 & 1 & \cdots & -1 \end{vmatrix}$$

Figure 6
Initial population of schedules

The parameters of the genetic algorithm are presented in Table 2. The initial population size corresponds to the number of requests. To ensure the diversity of the initial schedule options, we apply the Round-Robin method to do the random permutations of requests in the queue.

Table 2
Genetic algorithm parameters

№	Parameter	Value
1	Initial population size	$\lvert R_{req}\rvert$
2	Temporary population size	$2\cdot\lvert R_{req}\rvert$
3	Crossover ratio	100 %
4	Mutation ratio	50 %
5	Mutation probability of schedule	0.01
6	Selection ratio	50 %
7	Number of epochs	20
8	Reward/penalty coefficients	$\alpha,\ \beta,\ \gamma$

The crossover operator provides the inheritance by a child of optimal VM assignments from both parents within the previous population as it shown in Figure 7.

$$\begin{array}{ccccc} Parent_1 & & Parent_2 & & Child \\ 1 & \times & 1 & \rightarrow & 1 \\ 1 & \times & -1 & \rightarrow & 0 \\ -1 & \times & -1 & \rightarrow & -1 \\ 1 & \times & 1 & \rightarrow & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & \times & 1 & \rightarrow & 1 \end{array}$$

Figure 7
Crossover operation

The mutation operator changes a single schedule, which is randomly chosen from the current population (Figure 8).

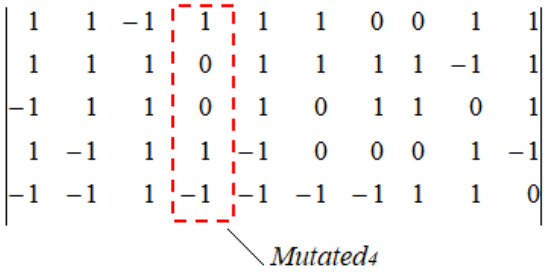


Figure 8
Mutation operation

During the selection operation the individuals with the maximum fitness function value (2) are saved as the base for the next population as shown in Figure 9. When the epochs' number is reached, the first individual will represent the best close-to-optimal schedule *S*.

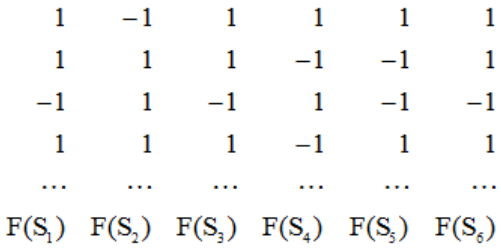


Figure 9
Selection operation

4 Experimental Research

To carry out the experiments, we implemented the CSCA simulator, using C++ language and Visual Studio 2012 IDE. To correctly generate the parameters of the cloud system and the requests we reviewed the needs of virtual resources for educational institutions of the Orenburg region. We also performed a statistical analysis of the distribution of fitness function to evaluate proposed scheduling algorithms.

4.1 CSCA Simulator

Figure 10 shows the UML class diagram, which depicts classes and relationships between them. Classes and methods correspond to the mathematical model of the CSCA.

The **ScheduleTemplate** class has field 'Intervals', which contains a set of time slots for each schedule template. The methods setDate() and getDate() are used for time slot write/read operations. The **Software** class contains information about software available in CSCA such as the software index (id), the supported operating system, the software title, the maximum number of licenses, and the software installation and uninstallation time.

The **Request** class represents the end-user request. Its TemplateIndex attribute refers to the chosen schedule template, SoftwareSet attribute is a subset of required software. The method setSoftware() randomly generates a subset of software according to the user request. The method setTimeSlots() randomly generates the time slot indices according to the user request. CloneT() is used to copy time slots indices, CloneS() – to copy software indices. The method SoftwareCheck() of the **RoundRobin** class is used to create initial schedule for CSCA.

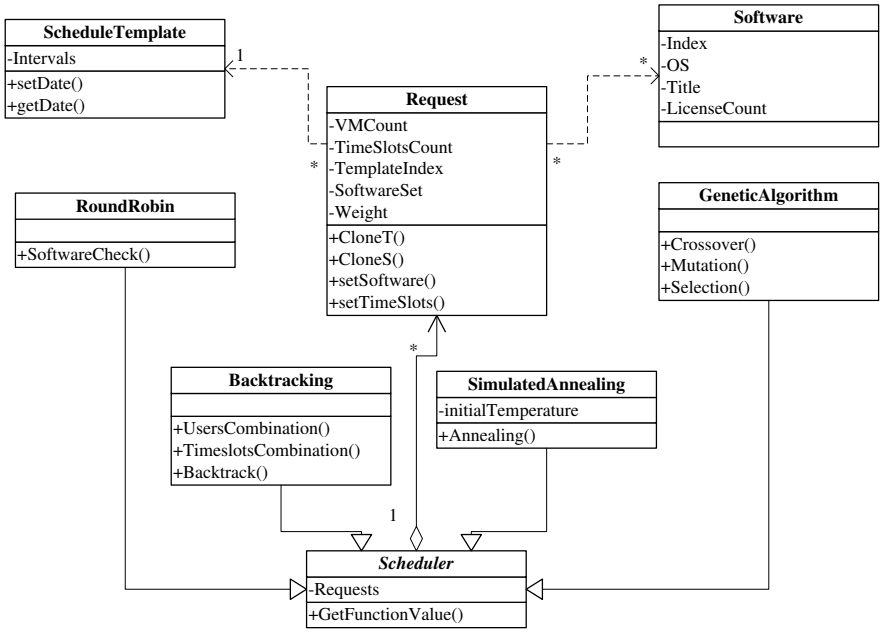


Figure 10
UML class diagram

The **Backtracking** class describes exhaustive search implementation, which is used to estimate the efficiency of proposed simulated annealing and the genetic algorithms. Its methods UsersCombination() and TimeSlotsCombination() are used to build all the possible combinations of requests placement in the CSCA schedule.

The method Annealing() of the **SimulatedAnnealing** class creates a close-to-optimal schedule of the CSCA. The field 'initialTemperature' is used to control the number of iterations of SA algorithm. The methods Crossover(), Mutation() and Selection() are used to implement corresponding operations of the **GeneticAlgorithm** class.

Scheduler is a parent abstract class for other classes, which implements corresponding heuristic scheduling algorithms: Round-Robin, Backtracking, Simulated Annealing and Genetic Algorithm. The GetFunctionValue() method is used to calculate the fitness function value (2).

150 experiments of requests generation and scheduling were performed within CSCA simulator for a large number of users. Examples of generated requests are shown in Figure 11.

```
                    UsersCount = 15

VM count: 5                     VM count: 10
Template index: 0               Template index: 1
0 1 2 3 4 5 6 7 8 9             0 1 2 3 4 5 6 7 8 9
1 2 7 8 4 5 3 0 6 9             9 6 8 5 1 7 3 0 4 2
Software count: 4               Software count: 4
Software indexes: 1 2 8 4       Software indexes: 9 5 7 3
Timeslots indexes:5             Timeslots indexes:7 0 3

VM count: 5                     VM count: 8
Template index: 0               Template index: 0
0 1 2 3 4 5 6 7 8 9             0 1 2 3 4 5 6 7 8 9
6 0 3 1 8 5 9 2 4 7             5 1 4 7 2 9 0 3 8 6
Software count: 1               Software count: 1
Software indexes: 6             Software indexes: 5
Timeslots indexes:4             Timeslots indexes:6

VM count: 13
Template index: 0
0 1 2 3 4 5 6 7 8 9
1 0 7 9 2 3 5 6 8 4                     ...
Software count: 1
Software indexes: 7
Timeslots indexes:0 4
```

Figure 11
Generated requests

The values of reward coefficients α, β and the penalty coefficient γ obtained empirically during the testing of the simulator. These values are interdependent quantities. When choosing the coefficients of the fitness function, we are guided by the following rule: $\alpha + \beta + \gamma = 2$. The worst case is the rejection of the user request for access to information resources, therefore the penalty coefficient has the greatest value of all three coefficients. However, the successful placement of the user virtual machines to the schedule is generally prevalent, and placing the virtual machines at the required time is the most preferable option. Thereby, the reward/penalty coefficients are chosen as follow: $\alpha = 0.7,\ \beta = 0.4,\ \gamma = 0.9$.

The simulated annealing algorithm satisfies 89.7% of generated requests in average. It is 1.88 times more efficient than the Round-Robin method. In addition, the genetic algorithm satisfies 86.5% of generated requests on average. And it is 1.48 times more effective than the Round-Robin method. The average execution time of the Simulated Annealing algorithm is 35 ms, of the genetic algorithm – 210 ms.

4.2 Statistical Analysis

The values of fitness function are varying in the range from -40 to 80. We divide this interval into 24 segments, each of length 5, and calculate the hit frequency of the value of fitness function (2) in each segment for the proposed heuristic algorithms as shown in Figure 12.

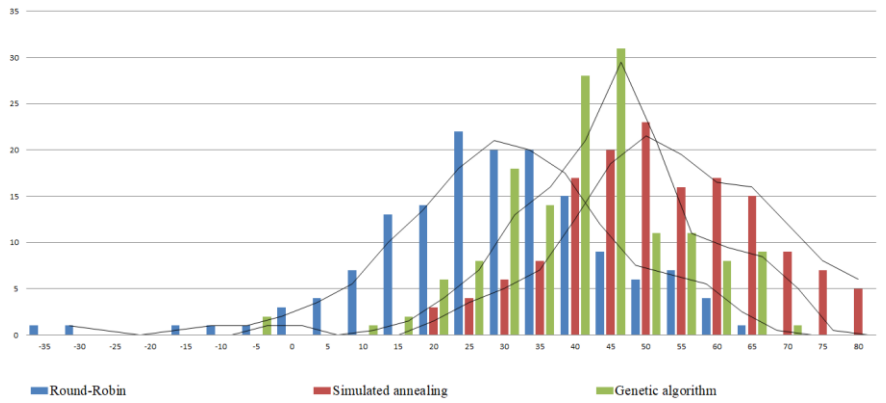


Figure 12
Interval distribution of value of fitness function (2)

After analyzing the graphs, we put forward a hypothesis H_0, that the observed frequency distribution of the value of fitness function (2) is a normal distribution.

Pearson's chi-squared statistical test [11] (χ^2) is used to accept or reject this hypothesis at significance level $\alpha_P = 0.05$ and the number of degrees of freedom K = 24-2-1 = 21. The test values are compared to one-sided critical value, which is equal to K_{crit} = **32.7** for the corresponding α_P level and the K number. Table 3 shows the statistical test results.

Table 3
Pearson's chi-squared statistical test (χ^2)

Algorithm	$\overline{X}$	σ	K_{obs}
Round-Robin	26.26	16.52	124.39
Simulated annealing	49.36	13.91	**6.82**
Genetic algorithm	38.73	13.26	87.33

Hypothesis H_0 is accepted for SA algorithm data set since the test statistic is less than the critical value of χ^2, and it is rejected for both the GA and RR algorithms because the test statistic greatly exceeds the critical value of χ^2.

In both cases, there are outliers in univariate data sets. Grubbs's test [12] is used to detect such outliers and reject them. To test whether the minimum value x_{min_obs} is an outlier, the Grubbs' test statistic is:

$$\tau_1 = \frac{\overline{X} - x_{\min\_obs}}{\sigma} \tag{6}$$

The critical value of the τ-distribution with a significance level of α_P is $\tau_{crit} = 3.81$. For both the GA and RR algorithms the x_{min_obs} and x_{min_obs+1} values were detected as outliers and rejected. Table 4 shows the χ^2-test results for the corrected data sets. Hypothesis H_0 is accepted for the Round-Robin and Genetic Algorithms since the test statistic is less than the critical value of χ^2.

Table 4
Grubbs' test to detect outliers in a univariate data set

Algorithm	$\overline{X}$	σ	K_{obs}
Round-Robin	27.09	15.00	**8.11**
Genetic algorithm	39.35	12.20	**17.96**

The hypothesis H_0 is accepted for all three data samples of fitness function value distribution, therefore, we can use Student's t-test [13] for independent samples with unequal variances (also known as Welch's t-test). Statistic t_e is defined by the following formula:

$$t_e = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}}} \tag{7}$$

Table 5 shows the significant differences between the data samples of proposed evolutionary algorithms. It can be concluded that the simulated annealing algorithm with the highest expected value is the most efficient scheduling solution for implementation.

Table 5
Student's *t*-test for independent samples with unequal variances

Data sample № 1	Data sample № 2	t_e	t_{crit}
Round-Robin	Simulated annealing	13.28	1.96
Round-Robin	Genetic algorithm	7.71	
Simulated annealing	Genetic algorithm	6.60	

Conclusions

The main results of this work are the following:

- The cloud system of collective access as the means of providing an economically profitable remote access to paid and free software was implemented for educational institutions of secondary education of the Orenburg region.
- The approach in solving the problem of efficient scheduling is described in detail.
- A new mathematical model which formalizes the process of cloud system operation, time limits and software license limits is developed.
- Two evolutionary scheduling algorithms based on simulated annealing and genetic programming are proposed.
- The UML class diagram of the CSCA is presented to give a proper look at finding a close-to-optimal schedule of a cloud system.
- The statistical analysis of the distribution of fitness function values is performed. Proposed simulated annealing algorithm shows the best results in experiments and was chosen to be implemented as a part of the Resources administration module of CSCA architecture.
- Machine learning prediction models (such as SARIMAX, Prophet, etc.) can be used to improve end-user request generation accuracy. See our paper [14] for more relevant information.
- The obtained results can be used in future research, to create multi-cloud shared access systems aimed primarily at providing access to virtual working environments with installed software.

Acknowledgement

This work was supported by the Russian Foundation for Basic Research (project 20-07-01065) and the the Russian Federation Presidential grants for support of leading scientific schools (NSh-2502.2020.9).

References

[1] Shukhman A. E., Bolodurina I. P., Polezhaev P. N., Ushakov Yu. A., Legashev L. V.: "Creation of regional center for shared access to educational software based on cloud technology" In: Selected Papers of the XI International Scientific-Practical Conference Modern Information Technologies and IT-Education SITITO-2016, 2016, pp. 180-188

[2] Eaves A., Stockman M.: "Desktop as a service proof of concept". In: Proc. of the 13th annual conference on Information technology education (SIGITE-12), 2012, pp. 85-86

[3] Tordsson J., Montero R. S., Moreno-Vozmediano R., Llorente I. M.: "Cloud brokering mechanisms for optimized placement of virtual machines across multiple providers" In: Future Generation Computer Systems. 2012, pp. 358-367

[4] Jang J. W., Jeon M., Kim H. S., Jo H., Kim J. S., Maeng S.: "Energy Reduction in Consolidated Servers through Memory-Aware Virtual Machine Scheduling" In: IEEE Transactions On Computers, 2011, pp. 552-558

[5] Akbari M., Rashidi H., Alizadeh S. H.: "An enhanced genetic algorithm with new operators for task scheduling in heterogeneous computing systems". Engineering Applications of Artificial Intelligence Journal, 2017, pp. 35-46

[6] Petkovics I., Tumbas P., Matković P., Baracskai Z.: "Cloud computing support to university business processes in external collaboration". Acta Polytechnica Hungarica, 11(3), 2014, pp. 181-200

[7] Hu J., Gu J., Sun G., Zhao T.: "A scheduling strategy on load balancing of virtual machine resources in cloud computing environment" In: Parallel Architectures, Algorithms and Programming (PAAP), 2010 Third International Symposium on, 2010, pp. 89-96

[8] Wang B., Xing H. "The application of cloud computing in education informatization".: In: Proc. of International Conference on Computer Science and Service System (CSSS), 2011, pp. 2673-2676

[9] Sarıtaş M. T.: "The Emergent Technological and Theoretical Paradigms in Education: The Interrelations of Cloud Computing (CC), Connectivism and Internet of Things (IoT)". Acta Polytechnica Hungarica, 12(6), pp. 161-179

[10] All-in-one solution for Remote Access and Web Portal / Terminal Service Plus, 2016, URL: https://www.tsplus.net/

[11] Pearson K.: "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling". Philosophical Magazine Series, 1900, 157-175

[12] Grubbs F.: “Procedures for Detecting Outlying Observations in Samples”. Technometrics, 1969, 11: 1-21

[13] Ruxton G. D.: “The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test”. Behavioral Ecology, 2006, pp. 688-690

[14] Polezhaev P., Legashev L., Shukhman A., Bolodurina I.: “Simulation Modeling of Virtual Machine Based Services in Multi-Cloud Shared Access Centers” In: 7th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2019), 2019, pp. 219-224