

# Distributed computing system for creating digital portraits of complex systems

Igor Starostin<sup>1</sup>, Sergey Khalyutin<sup>2</sup>, Victoria Pavlova<sup>1,2</sup> and Elena Punt<sup>1\*</sup>

<sup>1</sup>Moscow State Technical University of Civil Aviation, Russia, Moscow, 125993, Kronshtadtsky Boulevard, 20

<sup>2</sup>LLC «Experimental laboratory NaukaSoft», Russia, Moscow, 129085, Godovikova street, 9, building 1

**Abstract.** Various methods of operation of complex transport systems imply knowledge of mathematical models of their components. To obtain adequate models of such components, it is necessary to take into account the physical and chemical processes occurring in them. Previously, the authors developed a potential-flow method within the framework of modern nonequilibrium thermodynamics – a unified approach to the analysis and modeling of processes of various physical and chemical nature. In accordance with this approach, as well as with the methods of mechanics, the theory of electric and magnetic circuits, electrodynamics, etc., state functions for the properties of substances and the processes under consideration are set up to the experimentally studied constant coefficients. The system of equations of the considered processes dynamics is obtained from the given state functions. The desired model (digital portrait) of the considered component is constructed by numerical-analytical transformation of the dynamic equations system based on the use of experimental data. The need to automate the proposed method of obtaining digital portraits is due to its complexity and the need to process a large amount of data. An information and computing system is proposed, which implies the construction of a block diagram of the processes in the component under consideration (model-oriented approach). Modeling these processes using a block diagram at different values of unknown parameters allows us to approximate the model (digital portrait) a component based on the resulting set of output characteristic dynamics using machine learning libraries. Process modeling and further approximation of the model is parallelized. This paper is devoted to a distributed information-computing system that implements the creation of various complex systems digital portraits.

## 1 Introduction

Complex technical physical systems (physical and physico-chemical), including components of transport systems, contain simpler subsystems [1 - 9]:

- mechanical, described by the methods of mechanics [1, 2];

---

\* Corresponding author: [elena.punt@mail.ru](mailto:elena.punt@mail.ru)

- electric and magnetic circuits described by the methods of the electric and magnetic circuits theory [3], electrodynamics [3, 4];
- subsystems characterized by physical and chemical processes of different nature (for example, diffusion, chemical transformations, heat transfer, materials science processes), described by the methods of modern nonequilibrium thermodynamics [5 – 8];
- composite subsystems, including mechanical subsystems, electrical and magnetic circuits, physical and chemical subsystems, as well as complex subsystems (in turn, including the mentioned subsystems) [9].

The described methods for constructing systems of equations for the dynamics of processes in complex systems (and the dynamics of subsystems of complex systems) allow us to synthesize a system of equations for the dynamics of processes in the system under consideration, for the numerical implementation of which it is necessary to have experimentally studied parameters of the dynamics of processes in the system [1 - 8, 10]: constant coefficients of this model of processes, the initial state of the system, unknown external influences. The resulting system of equations must be supplemented with equations for the observed and controlled parameters of the system [9, 10].

In order to use this system of equations to solve practical problems of system development (synthesis of optimal structure and selection of optimal parameters [11 – 14]) in accordance with the specified requirements [11, 12], as well as the tasks of operating these systems (synthesis of the control system [15], diagnostics and forecasting of the technical condition [12, 16], analysis of the system reliability [12, 16, 17], development of the system maintenance methodology [17]), it is necessary to transform the resulting system of equations of process dynamics in such a way that only the observed and controlled parameters remain in the resulting system of equations [9, 10]. As a result of this transformation (in general, by numerical and analytical methods [10, 18-20], which consists in calculating the dynamics of the observed and controlled parameters of the system under study for various randomly set parameters of the processes dynamics in this system and further approximating the model of the system under study on a set of the observed and controlled parameters dynamics [10, 18, 19]), the relationship between the controlled parameters of the system and the observed parameters is obtained with accuracy to the control parameters (obtained from the test results of the considered system and laboratory systems) [10]. The controlled parameters of the system also include the observed parameters values predicted for subsequent time moments [10].

Based on the control parameters statistics from the considered system obtained model, the conditional probabilistic characteristics of the controlled parameters are determined [19] (using the methods of probability theory [21]), for which the model [19] is constructed, which is directly used to solve the practical problems described above. The resulting mathematical model, which connects the controlled parameters with the observed ones, is a digital portrait of the considered system [22], which can be used, in particular, for diagnosing and predicting the state of electricity consumers from the intelligent distribution system [22 - 26], as well as electricity sources [22]. The digital portrait is implemented in the mathematical core of the diagnostics and forecasting center [22] in the power distribution system.

As can be seen from the above, the formation of a system model is quite a time-consuming process [9, 10, 18, 19]. This leads to the need for software implementation of the presented methodology for creating a digital portrait, which includes multiprocessor and multicomputer parallelization [9, 19, 27].

## **2 Features of software implementation of the system model formation procedures**

### **2.1 Implementation of the processes dynamics equations system**

The processes dynamics equations system of different physical and chemical nature in an investigated system as a physical and physico-chemical processes block diagram, as well as more complex subsystems (model-oriented approach) [8, 9, 28]. The specified parameters of the dynamics of processes in the system are included in such block diagrams [9, 28]. Then, in accordance with the block diagram of processes in the investigated system, the proposed information and analytical system calculates the dynamics of processes in the investigated system, including the observed and controlled parameters, for the specified parameters of the processes dynamics [9, 28].

In general, for a complex system, the diagram of the processes occurring in it (or its subsystems) can be distributed (and therefore the calculations on this diagram are parallelized) on several computing nodes [9]. For this purpose, the considered block diagram is broken into subdiagrams by weaker connections, and the corresponding intermediate dynamics of the considered quantities are introduced in place of the broken connections [9]. And then each such subdiagram is implemented separately on the corresponding computing node [9]. The dynamics of processes in subdiagrams is calculated in parallel (on different computing nodes) iteratively [9].

Since different complex systems can generally have different subsystems, it is necessary to have a database of these subsystems.

### **2.2 Implementation of the numerical-analytical transformation of the process dynamics system equations**

The proposed information and analytical system accept the following data at the input:

- block diagram of physical and chemical processes in the system under study,
  - containing the observed and controlled parameters;
- ranges of changes in the parameters of dynamic processes in the studied system.

Having received these input data, the proposed information and analytical system generates random values of the parameters of the dynamics of processes in the studied system in the specified ranges, then simulates the corresponding dynamics of the observed and controlled parameters for these generated parameter values [9, 28], using the acceleration of counting due to parallelization. Then, based on the set of these system controlled and observed parameters obtained dynamics, its model is approximated [10], using symbolic regression methods [29] and neural networks [30].

For this purpose, the set of obtained dynamics of the observed parameters (and the corresponding control parameters) of the system is divided into subsets, on each of which a private model of the system is constructed [9, 18], using clustering methods [31]. Moreover, these subsets are different for each set of values of the parameters of the dynamics of processes in the considered system [9, 18]. The private models are then combined into a general model [9, 18]. The construction of private models, as well as their integration into more general ones, is performed in parallel on different computing nodes-executors [9, 18]. For this purpose, analytical expressions of particular and general models are collected in a database, which shows which general models are reduced to which particular models.

Moving from particular models to more general models, we refine the coefficients of general models [9]. To this end, we generally divide the resulting model into components that are weakly related to each other [9], and the coefficients in these components on the

corresponding input data (on those where these coefficients are most pronounced) are adjusted by optimizing in parallel on different computing nodes [9].

Then we generate additional dynamics of the observed and control parameters of the system from the system of equations of process dynamics and use these dynamics to check the final model of the system [9] - the result of a numerical and analytical transformation of the system of equations of process dynamics [9, 10, 20].

### **2.3 Getting a system model (digital portrait)**

Having obtained a model of the considered system by converting the processes dynamics equations system (a block diagram of the system), and having the test statistics of the system various instances (statistics of the control parameters of the system instances), we test the resulting model [19]. If such a model has been tested, then the processes in the considered system, as well as the state functions for the properties of substances and processes [7, 8], were set correctly [9], otherwise, it is necessary to consider other options for the processes and state functions for the properties of substances and processes [7, 8, 9]. Thus, an experimental study of the processes in the system, as well as the properties of substances and processes, is performed [7, 8].

Then, in the general case, we construct a probabilistic model of this system that can be used to solve various practical problems [19]. For this purpose, the resulting model is used to calculate the statistics of the values returned by this model (parallelization is performed for subsets of the values of the available statistics), as well as, if necessary, all the necessary statistical characteristics [19]. Then, for these obtained values, a model (in general, probabilistic) is constructed [19] (parallelization is similar to the one described above) – a digital portrait of the considered system [22]. Previously, from the available statistics (source and calculated), it is necessary to allocate test data for testing the resulting digital portrait.

## **3 Distributed information system for building digital portraits of the user system**

### **3.1 User-defined functionality of a distributed information system for digital portraits building**

As follows from the above, the functionality of the proposed information system is:

- building digital portraits of user-defined systems of various physical and chemical natures;
- database management:
- database of block diagrams of processes (in subsystems);
- database of system models (accurate to the parameters);
- database of state functions for the properties of substances and processes;
- database of ranges of parameters of the dynamics of processes in the system;
- database of statistics of test results of various systems;
- database of analytical expressions (necessary for building system models, setting state
- functions for the properties of substances and processes).

This database management functionality is:

- adding information;
- deleting information;
- reading information;
- updating information.

These databases are generally distributed [32], as well as intelligent [32]. This makes it possible, for example, to obtain processes, as well as the properties of substances and processes, depending on the chemical composition of the system.

### **3.2 Architecture of a distributed system for building digital portraits of user systems**

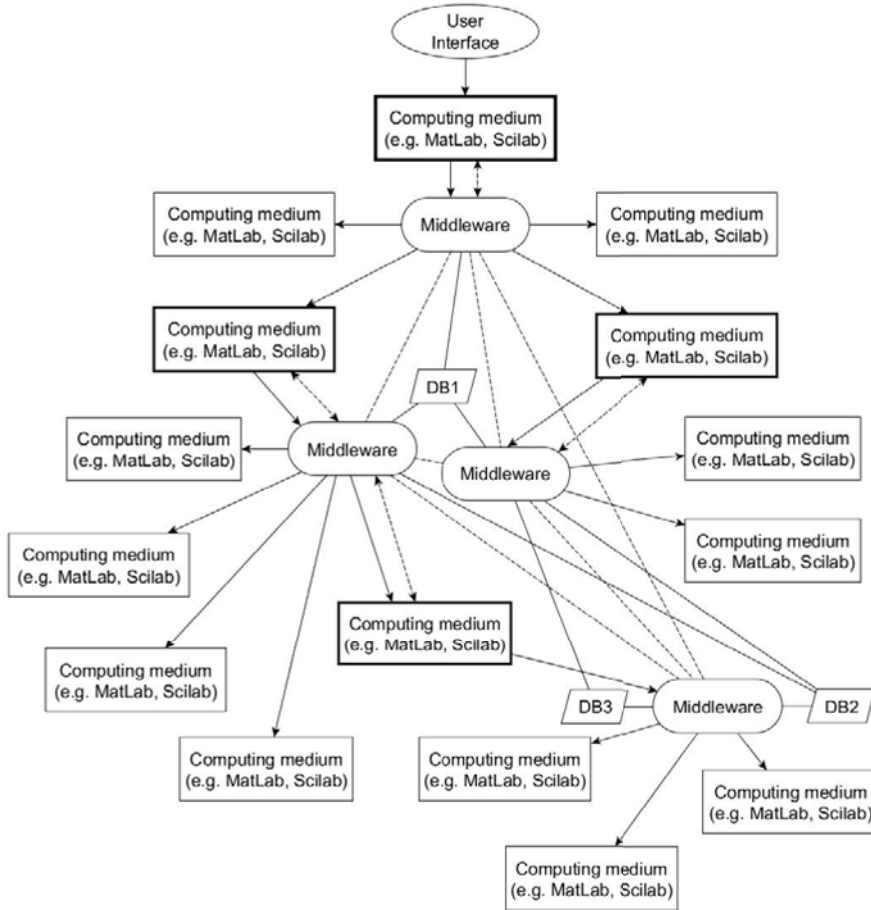
As can be seen, the basis of the formalisms for constructing digital portraits of various user systems is multicomputer parallelization [9, 27], which is based on the MPI-messaging interface, as well as the interface for remote application launch over the network [27]. At each computing node in the network, an environment is launched, which also implements calculations based on block models of specified user systems (model-oriented approach [28]). Examples of such environments are: MatLab, Scilab. This environment should include support for the MPI interface, as well as the remote program launch interface over the network (RDP interface) [27, 33]. These supports should also be implemented in the form of middleware [27, 33, 34], which also includes database management [27, 33, 34]. One example of middleware is Hadoop [34].

On each subordinate compute node, the compute environment is waiting for a job. After receiving a task (encrypted in an MPI message [27, 33]), this environment sends the results of the task (also encrypted in an MPI message [27, 33]) to the node from which this task came and waits for the next task. The computing node-the customer plans sub-tasks and distributes these sub-tasks to the computing nodes-performers (Figure 1). Also, on each computing node-executor, a queue of tasks is implemented (based on middleware [33, 34]), each of which is executed sequentially.

Moreover, in the general case, a computing executor node can also be a customer node for other executors (Figure 1). In this case, a group of tasks is submitted to such a node-performer (having subordinate nodes-performers), which is distributed among the nodes-performers (subordinates). This organization of a multicomputer computing system (tree system [27]) allows each computing node-customer to manage a relatively small number of computing nodes; each computing node-customer is responsible for its own (small) class of tasks. Thanks to this approach, the overhead costs of distributing subtasks across the executing computing nodes are rationalized.

The described subtasks solved by the executive nodes are implemented in the form of corresponding expansion modules (Figure 2). The expansion modules of the computing nodes-customers (Figure 3) implement the planning of the corresponding subtasks for the computing nodes-performers (Figure 2) and further planning based on the results of solving these subtasks on the computing nodes-performers, and the corresponding expansion modules of the nodes-performers – algorithms for solving these subtasks (Table 1). On the computing nodes-executors, which have their own executing nodes in their subordination, two of the above classes of extension modules are implemented (those that plan subtasks and solve subtasks) (Figures 2 and 3).

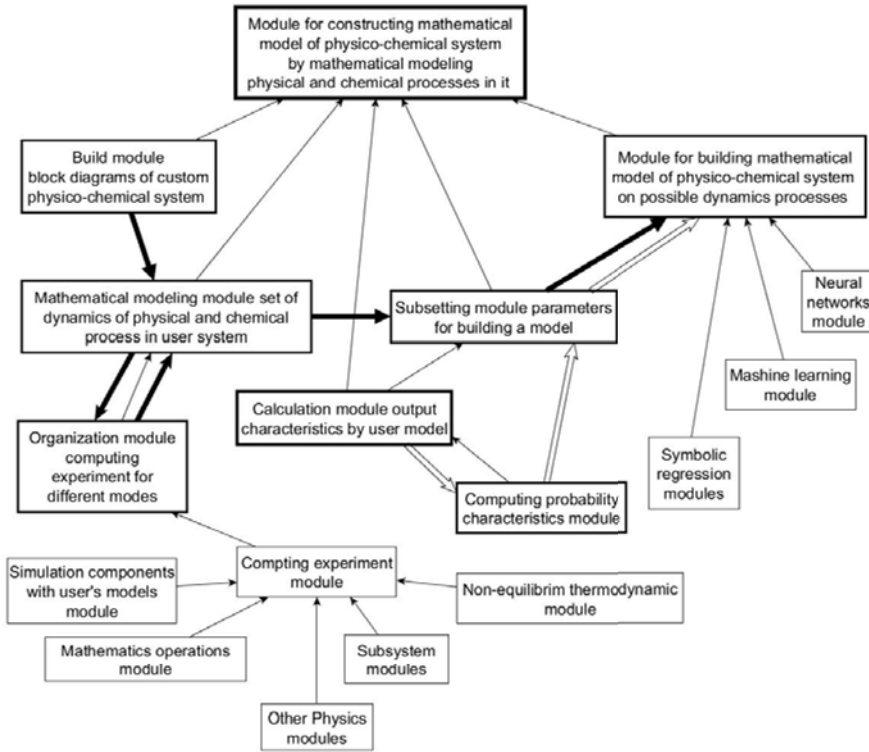
In general, part of the computing nodes in the considered computer network can be connected to each other as via a local network, and the other part – via the Internet. This approach allows you to create a computing system with separate clusters distributed in different locations. Also, access to the database can be either via a local network or via the Internet [27, 32, 33].



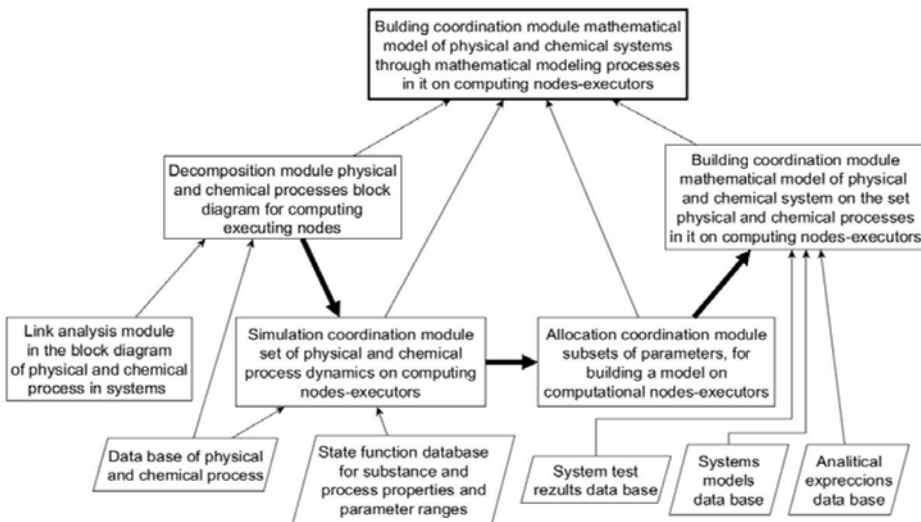
**Fig. 1.** A distributed computing system that implements the formalism of building models based on the analysis of processes in systems. The arrows show the directions for transmitting tasks, and the results of solving problems are transmitted along these arrows in the opposite direction. Dotted lines without arrows – interstitial interaction between the middleware components. Dotted lines with arrows – access from the computing environment to the database via middleware. Databases with numbers – local databases on each node in the network, solid lines without arrows-data exchange of middleware components with databases. Bold computing environments-customers

Auxiliary data stores are also used, which store the above-mentioned frozen dynamics (iteratively adjusted). Additional storage is also implemented on the nodes where the corresponding calculations are performed, and is also connected over the network. In addition to these additional data stores, additional data stores are implemented for the dynamics of the output characteristics of the system.

To increase the reliability of the described information and analytical system, it is necessary to have duplicates of various computing nodes-performers who have their own executing nodes in their subordination [27, 32 – 34].



**Fig. 2.** Modules for building a mathematical model on computing nodes-performers. The thick arrows show the sequence of library outputs during the construction of the system model (the black arrows represent the transformation of the process dynamics model, the white arrows represent the construction of the digital portrait itself). Thin arrows – hierarchy of extension modules



**Fig. 2.** Module for coordinating the formation of a mathematical model of a given physico-chemical system. The thick arrows show the sequence of library calls in the process of building a system model. Thin arrows – hierarchy of extension modules.

The described architecture of the proposed information and analytical system guarantees its universality for the described tasks of constructing digital portraits of various systems.

## References

1. A.G. Gorshkov, L.N. Rabinskiy, D.V. Tarlakovskiy, *Tensor Analysis Basics and Continuous Environment Mechanics* (Moscow, 2000)
2. I.M. Belenkiy, *Introduction to Analytical Mechanics* (Moscow: Higher School, 1964)
3. L.A. Bessonov, *Theoretical basis of electrical engineering* (Moscow: Higher School, 1996)
4. V.P. Yakubov, *Electrodynamics* (Tomsk: NTL Publishing House, 2006)
5. D. Jou, J. Casas-Vázquez, G. Lebon, *Extended irreversible thermodynamics* (New York, USA: Springer, 2006)
6. V.A. Etkin, *Energodynamics (synthesis of theories of energy transfer and conversion)* (St.Petersburg: Science, 2008)
7. I.E. Starostin, V.I. Bykov, *Kinetic theorem of modern non-equilibrium thermodynamic* (Raleigh, North Carolina, USA: Open Science Publishing, 2017)
8. I.E. Starostin, A.G. Stepankin, *Program implementation of methods of modern non-equilibrium thermodynamics. And the system of simulation of physicochemical processes SimulationNonEqProcSS v.0.1.0* (Bo Bassen, Mauritius: Lambert academic publishing, 2019)
9. I.E. Starostin, S.P. Khalyutin, A.V. Altukhov, A.O. Davidov, *Parallelization Applied to the Synthesis Methodology and Operation of Complex Systems Based on the Analysis and Modelling of their Physical and Chemical Processes*, 2020 1st International Conference Problems of Informatics, Electronics, and Radio Engineering (PIERE), p 287 (2020)
10. I.E. Starostin, S.P. Khalyutin, *Identification of system models from potential-stream equation on the basis of deep learning on experimental data*, Civil aviation, **v.23(2)**, p 47 (2020)
11. M.A. Gaides, *General systems theory (systems and system analysis)* (Moscow: Globe Press, 2005)
12. A.V. Antonov, *Systems analysis* (Moscow: Higher School, 2004)
13. I.P. Norenkov, *Fundamentals of Computer Aided Design* (Moscow: MSTU, 2002)
14. V.A. Ovchinnikov, *Graphs in the problems of analysis and synthesis of structures of complex systems* (Moscow: MSTU, 2014)
15. N.D. Egupov, *Methods of classical and modern theory of automatic control. T. 3. Methods of modern theory of automatic control* (Moscow: MSTU., 2000)
16. L.P. Kolodezhnyi, A.V. Chernodarov, *Reliability and technical diagnostics* (Moscow, 2010)
17. E.Y. Barzilovich, *Complex Systems Maintenance Models* (Moscow: Higher School, 1982)
18. I.E. Starostin, *Software implementation on the basis of SCM Scilab methodology for constructing a mathematical model of the system from the equations of physicochemical processes in it*, Proceedings of the International Symposium "Reliability and Quality", T.1, pp.41 (2020)



19. I.E. Starostin, *Obtaining reliability models of technical objects from potentially stream equations of physicochemical processes in these objects*. Reliability and quality of complex systems, **v.1(29)**, pp. 6 (2020)
20. V.N. Lantsov, *Methods of lowering the order of models of complex systems* (Vladimir: Vladimir State University named after A.G. and N.G. Stoletovs, 2017)
21. A.V. Pechenin, A.V. Teskin, G.M. Tsvetkova, *Probability theory* (Moscow: MSTU, 2004)
22. S.P. Khalyutin, *Aircraft power distribution system - aircraft equipment diagnostics and forecasting center*, Power supply, v. 2, pp. 4 (2020)
23. E.A. Punt, S.P. Khalyutin, A.O. Davidov, *Analysis of tasks of forming thermal imaging of electrical devices*, 1st International Conference Problems of Informatics, Electronics, and Radio Engineering, PIERE 2020, pp 104–107 (2020)
24. N.A. Simankov, A.A. Savelov, S.P. Khalyutin, A.O. Davidov, *Possibility of aircraft electrical equipment diagnostics by the local load control units*, International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices, EDM, 2020-June, pp. 318–323 (2020)
25. V.I. Pavlova, S.P. Khalyutin, A.A. Savelov, A.O. Davidov, *Monitoring and Diagnostics of the Technical Condition of Built-in Power Sources of Aviation Equipment*, International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices, EDM, 2019-June, pp. 464–468 (2019)
26. V.M. Dmitriev, *Application of impedance analysis method for assessment of technical condition of power supply system component*, Power supply, **v. 3**, p 4 (2020)
27. V.V. Voevodin, V.I. Voevodin, *Parallel computings* (St. Petersburg: BHV-Petersburg, 2002)
28. E.V. Petrunina, *Functional graphs in the simulation of complex systems*, Works of the All Russian Scientific and Practical Conference "Simulation Modeling: Theory and Practice", p. 280 (2015)
29. Dang Thi Fuk, A.I. Diveev, E.A. Sofronova, *Solving Problems of Identification of Mathematical Models of Objects and Processes by Symbolic Regression*, Cloud of Science, **v. 5(1)**, p.147 (2018)
30. S. Haykin, *Neural Networks (A Comprehensive Foundation)*, (Upper Saddle River, USA: Prentice hall, 2006)
31. K.V. Vorontsov, *Clustering and multidimensional scaling algorithms. Lecture course* (Moscow:MSU, 2007)
32. S. Davy, D.B. Arno Meysman, Ali Mohamed, *Introducing data Science. Big data. Mashine learning and more. Using Python Tools* (Shelter Island: Manning, 2017)
33. S.P. Kopysov, A.K. Novikov, *Parallel Computing Middleware* (Izhevsk: "Udmurt University", 2012)
34. White T, *Hadoop. The definitive guid* (Beijing, Cambridge, Farnham, Koln, Sebastopol, Tokio: O'Relly, 2013)