

# Architecture of data warehouse of the intelligent oil wells control system

**K.F. Tagirova<sup>1</sup>, A.R. Ramazanov<sup>1</sup>, A.M. Vulfin<sup>1</sup>, A.V. Nikonov<sup>1</sup>**

<sup>1</sup>Ufa State Aviation Technical University, K. Marks st. 12, Ufa, Russia, 450077

**Abstract.** The purpose of developing data warehouse architecture is improving the storage subsystem as part of an intelligent oil well control system. To achieve this goal, it is necessary to develop the structure and information model of the storage subsystem. It is proposed to expand the exploration and exploitation data model of the Epicentre oil fields by adding new entities and a hierarchy of links. Highly effective architectural templates, taking into account the specifics of the research object and the accumulated experience of industry experts, are used. This will reduce the costs for development of the data warehouse architecture and will provide an opportunity for integration with existing information systems of a higher level of control and management – MES, ERP – and with other software and hardware complexes of the current level.

**Keywords:** software system architecture, intelligent control system, information model.

## 1. Introduction

Nowadays, the oil industry in Russia has entered to the stage of oil field development and oil reservoirs and wells exploitation on a scientific basis with fitting out oil fields and wells with modern equipment. Various “smart” production facilities are being actively implemented – intelligent wells and intelligent fields that allow reducing capital costs and reducing production costs by saving energy and improving control efficiency [4]. At the moment, the second generation of smart fields is being implemented, the main difference of which is realization of the field development remote control [5]. However, the implementation of such systems is impossible without the use of modern measuring instrumentation as well as control of technological equipment.

On the Russian Federation territory about 50 thousand wells are operated by installations of sucker rod pumps (SRP), which provide to 20% of all produced oil [1]. However, most of the wells from the SRP fund – low-yield, so it is economically inefficient to equip them with modern expensive controllers with complex intelligent control algorithms. It is necessary to use an intelligent control algorithms realized as additional modules in the control station of the SRP. Therefore, intelligent monitoring and control systems are proposed to solve the problem of control the SRP equipped wells. This system allows to solve the following tasks [6]:

1. Control of the technical condition of the well equipment;
2. Calculation of the well production rate to select control mode;
3. Control of the balance of pumping rocking machine.

During operation of such system the problem of processing and analyzing large volumes of technological data arises. Therefore it becomes an actual problem to organize a data warehouse for efficient access to the accumulated data and to apply intelligent algorithms for well control. Appropriate requirements for such storage are [7]:

1. Integration with existing information systems of higher level of control and monitoring – MES, ERP [8];
2. The ability to integrate with other software and hardware complexes of the current control level, for example, with a variety of equipment monitoring systems;
3. Ensuring the logical integrity and consistency of the stored data as well as the coherence of data in distributed storage systems.

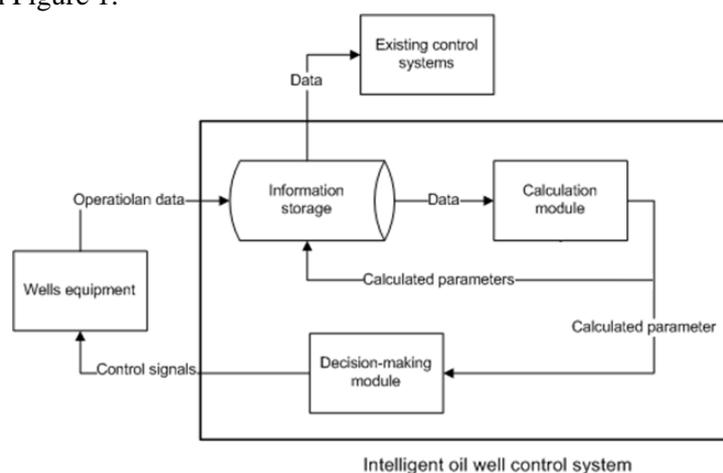
The purpose of data warehouse development is improving storage subsystem as a part of the intellectual oil well control system.

To achieve this goal it is necessary to solve the following tasks:

- Develop a structure of the storage subsystem,
- Develop an information model of the storage subsystem.

## 2. Structure development of a data warehouse subsystem as part of an intelligent oil well control system

Proposed information storage is a part of the intelligent oil wells control system which structural diagram is shown in Figure 1.



**Figure 1.** Structural diagram of the intelligent oil wells control system.

Equipment receives wattmetrograms and dynamograms of SRP operating, which are stored in the information storage. From this storage, the data enters the calculation module, which determines the state of equipment, calculates the flow rate and determines the balance of the rocking machine. These data are returned to the information storage and are also transferred to the decision module for determining control actions.

## 3. Information model development of a storage subsystem based on object-oriented decomposition

In order to use existing high-performance architectural templates for a proposed storage subsystem that takes into account the specifics of the research object and the experience gained by industry specialists embodied in existing information systems for processing and storing data of oil producing enterprises, the exploration and exploitation data model of the Epicentre deposits is taken as the basis. It is developed by the corporation POSC (Petrotechnical Open Software Corporation). Expanding the existing POSC model to implement intelligent management and diagnostic functions, it is possible to build an effective information model of the data warehouse taking into account design features [9, 22, 23, 24].

Epicentre model is created independently of the applications and at a logical level, with no matter of the type of final data warehouse, what allows integration of existing solutions in case they are built with respect account the proposed model. This simplifies the task of communication of modules and systems from different manufacturers and following to the basic principles of open systems.

More than 1,000 existing technical and business objects associated with the oil exploration and production are defined in the Epicentre data model. In the terminology of POSC data modelling, these objects are named entities. Characteristics than may be contained in entities, defines in the model as entity attributes. The most important are the attributes that define the relationship between the entities [2].

When designing data model for operating the technological parameters of intelligent control system of oil wells stations, it is necessary to enter new entity to expand Epicentre model, namely the entity of wattmetrogramm (*wattmetrogramm*), dynamogramm (*dynamogramm*), pretreated dynamogramms (*prepared\_dynamogramm*) and classified dynamometer (*classified\_dynamogramm*), which are necessary for solving the problem of diagnostics of equipment condition according to dynamometer data and for control tasks [10, 11, 13-21].

Wattmetrograms and dynamograms can be considered as part of technological time series (TTS) [26, 27]. These entities must be integrated with existing entities by building a hierarchy of links. The Epicentre model for providing communication instead of repeating the identifier, an attribute (attribute), pointing to another entity, is added in the logically integral element (entity). Thus, to organize relations with the rest of data, the entities of *wattmetrogramm* and *dynamogramm* contain a reference to the well (well), and the entities *prepared\_dynamogramm*, *classified\_dynamogramm* contain a reference to the dynamometer. For example, the entity of *wattmetrogramm* will be as shown in figure 2:

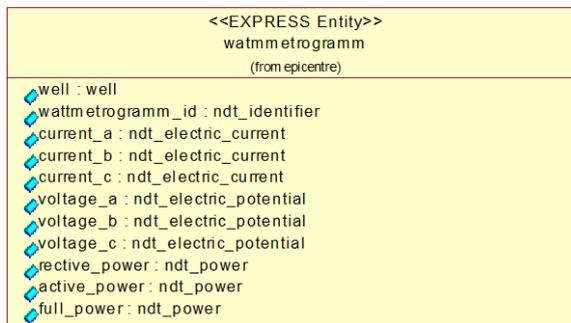


Figure 2. Wattmetrogramm entity.

In the Epicentre data model attributes presented by named types, having precise definition (*named defined types*), which consist of one or several basic data types. In the Epicentre data model documentation, named types with a precise definition have their names indicated by the prefix “Ndt”. Therefore, while creating the data warehouse model, already existing types were used, but for complete description of the model two own types to storage dynamograms were implemented: plunger stroke counts array – *Ndt\_position\_array* and counts of loads per rod array – *Ndt\_load\_array*. Figure 3 shows the basic types used during development of warehouse:



Figure 3. Main named types used in the development of warehouse.

Thus, the overall structure of the information model is as follows (Figure 4) [22-25]:

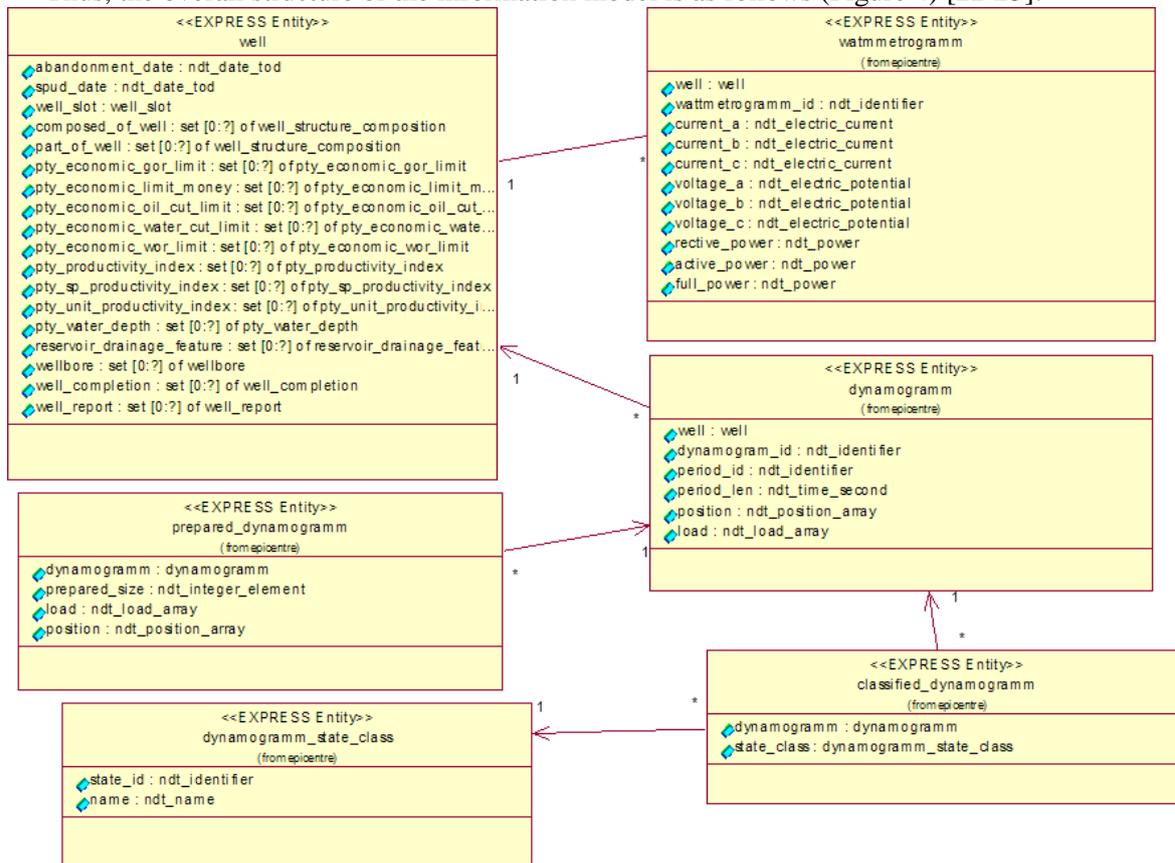


Figure 4. General structure of warehouse model.

Epicentre data model is described in EXPRESS language, but EXPRESS is not the equivalent of the data definition language (DDL), such as the Standard Query Language (SQL) DDL [3]. Therefore, for development of data warehouse it is necessary to convert expressions in EXPRESS language into a set of expressions on DDL language. This conversion POSC corporation called projection. So the creation of the physical warehouse is presented in Figure 5.

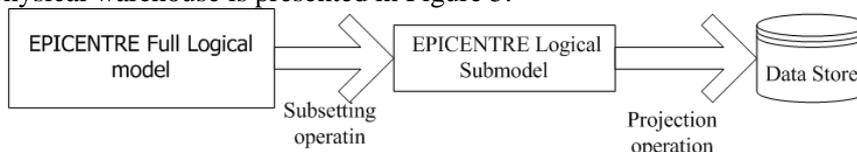


Figure 5. Creating a data warehouse model of EPICENTRE.

In order to project Epicentre model into a set of expressions on DDL language, POSC Corporation proposes to use the tool to project Epicentre data model to the four types of Relational Database Management System (RDBMS) DDL: ANSI, Informix, Oracle, and Sybase.

Based on the Epicentre 3.0 model and the subsets of the Epicentre “AC3 data model: Process Data”, “DM1: Document Specification”, “DM3: Document Schema”, “DM4: Document Content”, it is forming an information model of data processed in the intelligent oil well control system – time series – and in the subsystem for storing the parameters of the intelligent analysis module (Figure 6).

Detailed disclosure of attributes of key entities and basic links is given in the section devoted to the development of the Intelligent Control System (ICS) PC (Figure 7).

From Epicentre data model it was selected three groups of entities characterizing the different stages of processing the technological time series as the main data source for implementing the methods of intellectual analysis, additional elements that store and process the accumulated data.

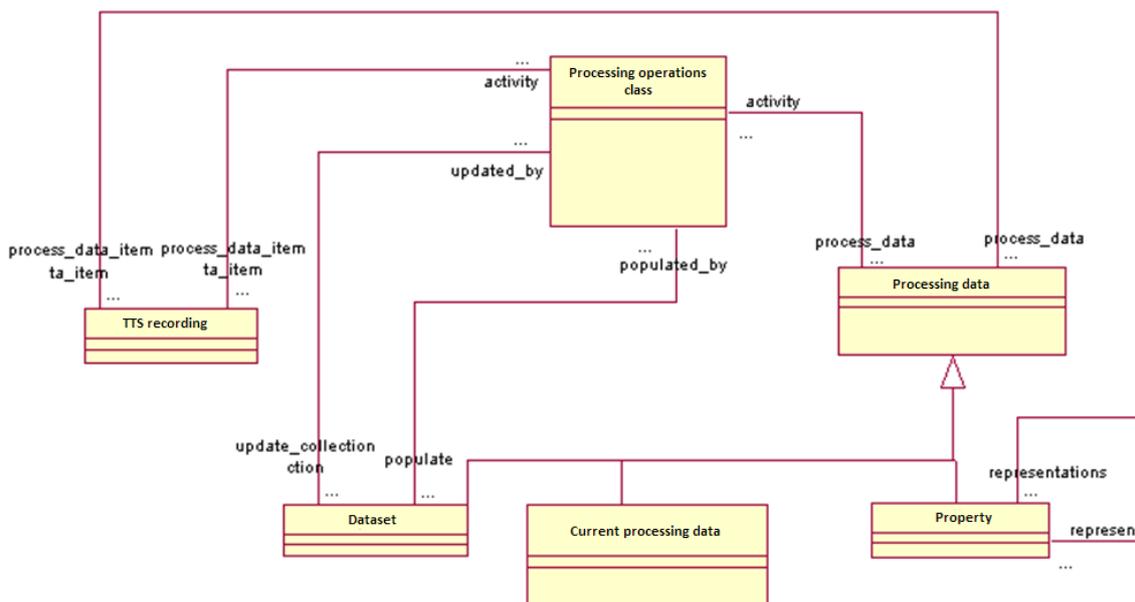


Figure 6. Truncated and augmented subset AC3: Process Data.

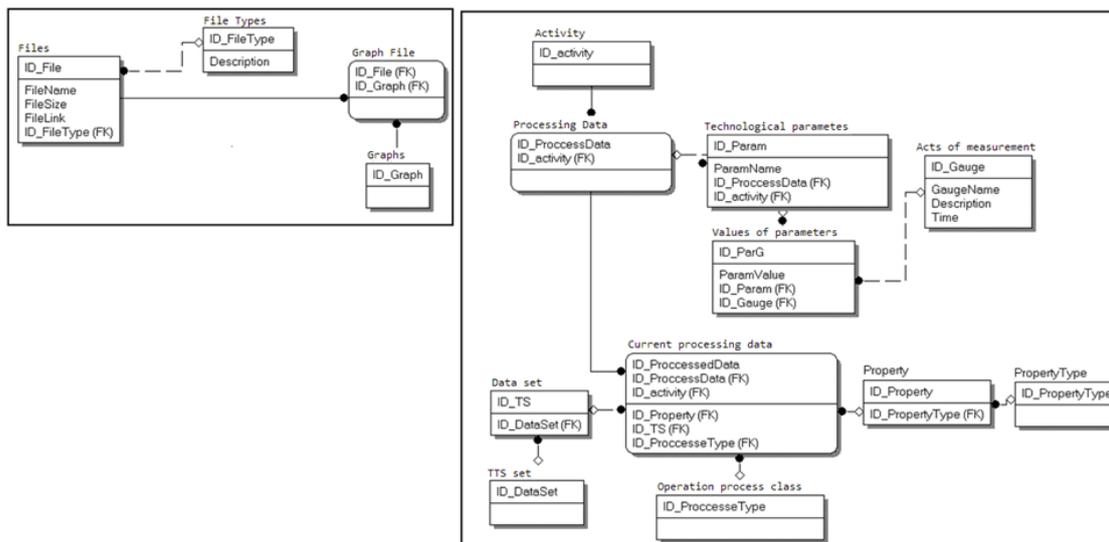


Figure 7. Prototype of the logical model of technological time series processing in the ICS.

In [28] the application of intelligent analysis technologies in diagnostics and control tasks is shown: committees of neural network models and preprocessing algorithms for technological time series based on the application of wavelet analysis.

The process of life cycle (LC) management of the developed software complex, which implements the modules for intellectual analysis and management of oil producing wells, corresponds to the PDCA cycle and the basic requirements of the process approach, formulated in MS ISO 9000: 2000, as a functional model. ISO / IEC 12207: 1995 “Information Technology - Software Life Cycle Processes” is the main normative document regulating the composition of the software life cycle processes (Figure 8).

Therefore, an information model of TTS processing as part of an intelligent control system for a group of oil producing wells was developed as a subset of the POSC Epicentre model, which allows the use standard interfaces between application programs, database management systems, workstations and users.

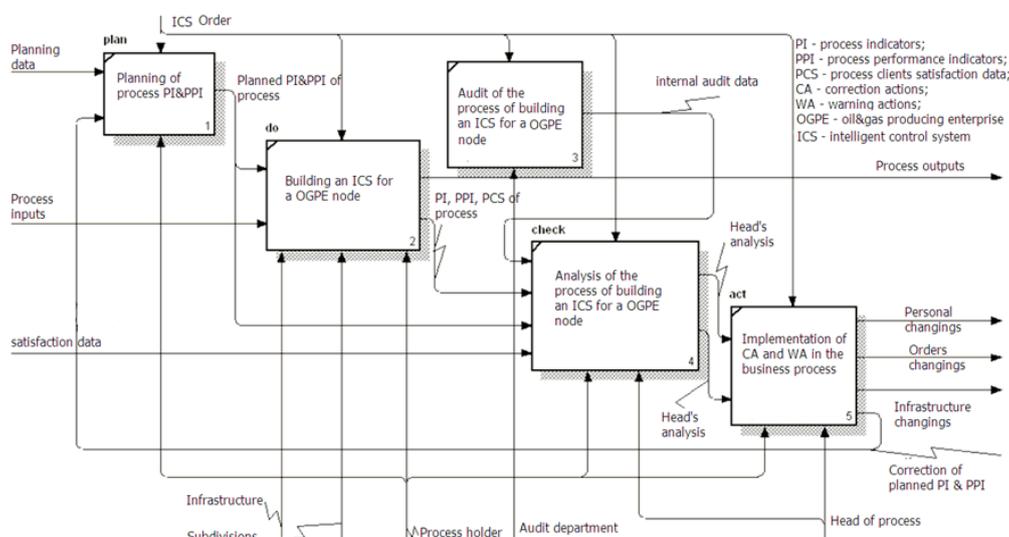


Figure 8. Management of the software development process.

#### 4. Conclusions

Thus, it is proposed to expand the Epicenter model by adding new entities and hierarchy of links in the task of the information model construction to improve the subsystem of data storage as part of an intelligent oil well control system, which will allow:

1. Reduce the labour costs for the development of the data warehouse by using existing high-performance architectural templates that take into account the specifics of the research object and the accumulated experience of industry experts.
2. Provide an opportunity for integration with existing information systems of a higher level of control and management – MES, ERP – and with other software and hardware complexes of the current control level through the use of the principle of open architecture.

#### 5. References

- [1] Market research facilities sucker rod pumps (SRP). The analytical report [Electronic resource]. — URL: <http://www.techart.ru/files/research/walking-beam-pumping-unit.pdf> (20.11.2017).
- [2] Almuhametov, A.A. Learning objects and relationships in Epicentre information-logical model of exploration and exploitation oilfields / A.A. Almuhametov, V.M. Giniatullin, U.I. Zozula, V.M. Filippov. – Ufa, 2003.
- [3] Epicentre Sample Relational Implementation. [Electronic resource]. – URL: [http://w3.energistics.org/archive/Epicentre/Relational\\_Implementation\\_Methodology\\_v2.3/main](http://w3.energistics.org/archive/Epicentre/Relational_Implementation_Methodology_v2.3/main). (accessed: 20.11.2017)
- [4] Gubarev, M.I. Opportunities and perspectives for the construction of “smart” wells / M.I. Gubarev // Materials of XX international scientific symposium of students, postgraduates, and young scientists devoted to the 120 Anniversary of the founding of Tomsk Polytechnic University. – Tomsk, 2016. – Vol. 2. – P. 723-726.
- [5] Dmitrievskii, A.N. Resource-innovation model and the solution of actual problems of development of oil and gas / A.N. Dmitrievskii, N.A. Eremin. [Electronic resource]. – URL: <http://oilconference.ru/d/304647/d/dmitrievskiy> (accessed: 20.11.2017).
- [6] Mazakov, E.B. Representation and processing of knowledge in information automated systems of intelligent field / E.B. Mazakov // Zapiski Gornogo Instituta. – 2014. – Vol. 208.
- [7] Yurev, A.M. Unification of control systems by the information exchange on mes production managements / A.M. Yurev // Automation, telemechanization and communication in oil industry. – 2011. – Vol. 4. – P. 17-20.
- [8] Kizina, I.D. Methodology of development and implementation of oil and gas industry companies energy menegement support information system / I.D. Kizina, K.A Sanarova // Automation, telemechanization and communication in oil industry. – 2011. – Vol 4. – P. 6-1.

- [9] Chengfang, L. Construction and Application of the Data Model in Petroleum Exploration and Production Fields / L. Chengfang // Beijing 2014 International Geophysical Conference & Exposition, Beijing, China. Society of Exploration Geophysicists and Chinese Petroleum Society, 2014. – P. 1319-1322.
- [10] Tagirova, K.F. Neural network algorithms of information processing in the tasks of diagnosing downhole pumping equipment of oil company / K.F. Tagirova, A.M. Vulfin // Automation, telemechanization and communication in oil industry. – 2013. – Vol. 12. – P. 28-32.
- [11] Vulfin, A.M. Enhancement of accuracy of deep-pumping equipment based on data mining / A.M. Vulfin, K.F. Tagirova // Optical Memory and Neural Networks. – 2015. – Vol. 24. – P. 28-35.
- [12] Ayvazyan, S.A. Applied Statistics: Classification and reduction of dimension / S.A. Ayvazyan, V.M. Buhtshtaber, I.S. Enyukov // Moscow: Finance and Statistics, 1989.
- [13] Vasilyev, V.I. Intelligent control systems. Theory and Practice / V.I. Vasilyev, B.G. Ilyasov // Training material, 2009. – P. 33-62.
- [14] Vorontsov, K.V. Machine learning. Lecture courses / K.V. Vorontsov. – [Electronic resource]. – URL: <http://www.machinelearning.ru/wiki/index.php> (accessed: 20.11.2017).
- [15] Belov, I.G. Investigation works well pumps dynamography / I.G. Belov // M.: Gostoptekhizdat, 1960.
- [16] Mansafov, R.Y. A new approach to the diagnosis of sucker rod pumps work on the dynamometer / R.Y. Mansafov // Engineering Practice. – 2010. – Vol. 9. – P. 92-99.
- [17] Vulfin, A.M. Diagnosing the state oil company equipment based on data mining techniques / A.M. Vulfin, A.I. Frid // Materials of II international conference “Information technology intellectual support of decision-making” (ITIDS’2014), 2014.
- [18] Vulfin, A.M. Safety Increasing of Oil Companies Engineering Networks Operation with Use of Artificial Intelligence Systems / A.M. Vulfin, A.I. Frid // Materials of 16 international science conference. – Sheffield. – 2014. – Vol. 3. – P. 53-58.
- [19] Tagirova, K.F. Modified algorithm determination of current parameters of SRP by dynamometric card data / K.F. Tagirova, A.M. Vulfin, A.R. Ramazanov, A.A. Fathulov // Automation, telemechanization and communication in oil industry. – 2015. – Vol. 12. – P. 37-41.
- [20] Tagirova, K.F. Coordinated control system for group of production wells on the basis of a hierarchical system of dynamic models / K.F. Tagirova, A.M. Vulfin, G.T. Bulgakova // Materials of IX scientific-practical conference “Mathematical modeling and computer technologies in field development processes”. – Ufa, 2016.
- [21] Vulfin, A.M. The information processing system for diagnosing oil engineering networks based on data mining techniques / A.M. Vulfin, A.I. Frid, V.M. Giniyatullin, S.A. Shuvalov, M.A. Salikhova // Bashkirskii khimicheskii zhurnal. – 2012. – Vol. 4. – P. 72-78.
- [22] Fowler, M. UML distilled: a brief guide to the standard object modeling language / M. Fowler // Addison-Wesley Professional, 2004.
- [23] Alexandrescu, A. Modern C++ design: generic programming and design patterns applied / A. Alexandrescu // Addison-Wesley, 2001.
- [24] Gamma, E. Design Patterns. Elements of reusable object-oriented software / E. Gamma, R. Helm, R. Johnson, J. Vlissides. – St. Petersburg: “Piter”.
- [25] Levykin, V.M. Design Patterns requirements for information systems: modeling and application / V.M. Levykin, M.V. Evlanov, M.A. Kernosov. – 2014.
- [26] Vulfin, A.M. Neural network model of time series analysis within the framework of the Data Mining methodology / A.M. Vulfin, A.I. Frid // Information-control systems. – 2011. – Vol. 5.
- [27] Vulfin, A.M. Intelligent automated decision support system for the technological complex of oil acceptance and delivery / A.M. Vulfin, A.I. Frid // Mechatronics, automation, control. – 2011. – Vol. 5. – P. 29-34.
- [28] Tagirova, K.F. Improvement of the USSN diagnosis based on data mining dynamometer / K.F. Tagirova // Automation, telemechanization and communication in the oil industry. – 2014. – Vol. 11. – P. 23-28.