Method of designing the modular structure of the information system

Y I Rogozov, S A Kucherov, J Y Lipko, A N Belikov, A Q Maakot and S A Belikova

Institute of Computer Technologies and Information Security, Southern Federal University, Taganrog, Russian Federation

corresponding author's e-mail: skucherov@sfedu.ru

Abstract. The overwhelming majority of modern information systems are built on the basis of a modular principle. This principle involves the creation of independent software subsystems that perform separate groups of tasks. The success of building an information system depends on the quality of the division of tasks into groups. Known approaches to modular partitioning are based on the organizational structure of the enterprise and job descriptions of employees. This leads to unmanaged intermodular connections and loss of the advantages of the modular approach (flexibility, interchangeability of modules, etc.). To solve this problem, the article proposes a new method for designing a modular structure of information systems based on the analysis of information and information flows. This allows minimizing inter-module communications and building systems that are independent of the organizational structure of the enterprise.

1. Introduction

Active changes that occur in organizations and enterprises require rapid changes in information systems and processes that are implemented in them [1,2]. The task of ensuring the flexibility of information systems is most effectively solved by modularly constructing the structure of a system of typed blocks — functional modules [3-5].

The principles of partitioning systems into modules defined by D. Parnas [6] are the only conditionally formalized element of the approach.

Currently, there are three main approaches to the modular partitioning of information systems:

- Individualistic approach based on the consideration of one private system and a particular subject area;
- The approach of the line of software products based on the consideration of a typical task and a set of similar subject areas.
- Configurable information systems based on one of the above approaches with the ability to make changes to the set and functional content of the modules

The main problem in creating a modular structure of an information system is the formation of functional content and the specification of modules.

Depending on the functional content of the module, there may be problems of critical dependencies between modules (replacing or updating a module is impossible without disrupting the system's performance due to the presence of multiple connections and dependencies with other modules), and

the problem of integrating modules into a single system (the almost complete absence of inter-module links leads to the complexity of the processes of continuous processing of the same data by employees at different workstations).
Depending on the level of granularity of the modules, both access control problems can arise (large-scale modules can be used by employees with different levels of access to data), as well as problems of optimizing user workspace (many different modules need to be combined into a single workstation, while still being able to perform automated business processes). and business transaction chains).
Thus, the approaches used to create a modular structure of information systems should offer criteria and methods for optimizing both the functional content and the granularity of the modular structure.
Despite the fact that each of the existing approaches allows us to obtain an information system based on the modules, the problem of effective partitioning into modules remains unsolved. Due to the fact that the module is formed empirically, it is not possible to assess its quality and compliance with the conditions of use. Also certain difficulties are the process of determining the balance between the static part and the extension point of the module.
To ensure the flexibility of the modular structure, an automated procedure is needed to isolate typical elements and structures of the system based on the analysis of business processes and modeling their interaction in the information system [7]. Baseline data for such a technique can be obtained from a survey and analysis of the enterprise [8]. The models constructed as a result of the survey, in contrast to the models of computational algorithms [1], are poorly formalized, graphical schemes [9]. This article proposes a method for constructing a modular structure of an information system based on the analysis of business processes.
To achieve this goal it is necessary first of all to solve the problem of developing a method for formally constructing a modular structure of an information system that takes into account the quantitative assessment of the degree of typification in its structural organization. It is assumed that the source data for the design of the model is a structured database obtained as a result of a pre-project survey of an enterprise using the procedure described in [10, 11].

2. Related work

Most of the known approaches to building models of business processes to obtain quantitative solutions use the idea of using optimization models. Most often, graph models are used in problems of this type [12]. A systematic description of graphs and examples of their real use were made in classical monographs [1,12].
The most formalized approach to the construction of models of processes and systems is a criterion approach based on the introduction of some formal criteria describing the purpose of the modeling process [13-15]. Usually they try to formulate the goal of the system in the form of an optimization problem in the form of $f(x) \Rightarrow \max(\min)$, where f is a certain scalar function, for example, design reliability, diagnosability of the resulting product, etc. x is a vector defining controlled (changeable) parameters, for example, the number of typical modules in a modular system, with $X = \{x^i\}, 0 \le x^i \le x$. Problems of this type are solved by finding the extremum of the function $f(x)$ on the set X in the form (1):

$$f(x) \underset{x \in X}{\Longrightarrow} \max(\min) \quad (1)$$

So, in design automation models [16], the designer is faced with the task of choosing a vector that gives the maximum (minimum) value to several technical functionals: $f_1(x), f_2(x), \ldots, f_v(x)$. For solving problems of this type, methods of linear convolution, check figures, Pareto compromises, etc. are usually used [13, 17, 18].
This approach is implemented when considering well-formalized tasks of computer-aided design, in particular, we will use them to solve the problem of building a modular structure of an information system

3. Method of designing the modular structure of the information system

When analyzing the activities and management of the enterprise, formalization of the organizational structure (OS) plays a large role. The organizational structure is, as a rule, modular, and includes four main aspects: a description of the structure of departments and divisions; communication between them and the external environment; information circulating on these connections (i.e. document circulation), as well as functions performed by structural subdivisions and departments. In accordance with this formalization, the standard methodology for constructing the organizational structure of an enterprise includes the following steps: identifying the main set of structural divisions of an enterprise (accounting, personnel department, sales department, etc.); depending on the goals of the enterprise, delegation of certain functions by this structural subdivision and the definition of links between the structural subdivisions and the workflow.

The distribution of production functions between departments (modules), based on the traditional approach, is usually applied [19]. To increase the efficiency of information exchange, it is often necessary to informally distribute functions between modules (divisions), delegate new functions and organize new connections to them [7, 20]. With this approach, the construction of the organizational structure of the "top" of the structural units (modules), which are traditionally performed functions, as well as the standard workflow [21, 22].

In the approach proposed in the article, everything should be the other way around - the decisive component of the organizational structure should be information and, in accordance with the goals of the enterprise, should determine the functions of its processing, their distribution between modules (departments) and, ultimately, the information flows of the enterprise's information management system [23, 24].

Now you can set the task of building a modular information system model that would allow reflecting not only the links between the structural units of the enterprise and their weight, but also evaluate the essence of the processes occurring in the organization. In this case, what functions of processing and generating information (documents) are performed within the enterprise, taking into account the distribution of these functions between departments [25, 26]. In the following, we will call this information model a functional-modular model. The development of such a model will make it possible to solve the problem of optimizing the organizational structure of an enterprise according to new criteria, for example, the degree of typification of the modules included in the system, etc. This article sets the task of describing a method for constructing a modular information system model and its information filling [26].

The proposed method for constructing a modular information system model includes the following steps:

1. Primary processing of information obtained during the survey of the enterprise;
2. The selection of functions performed by units;
3. Isolation of the simplest (elementary) functions that should be implemented in the AIMC;
4. Definition of connections between functions and their weights;
5. Building a global structure (scenario) of the system;
6. Decomposition of the global structure into a number of substructures (typical modules) taking into account the criterion of maximum typification

The model of an information system from the point of view of systemology [27], based on the general theory of algebraic systems, can be represented as follows:

$$S = (A, R) \tag{2}$$

where A is the set of system elements, and R is the set of relations between them.

A system property that can be used to determine differences in observations of the same model parameter is referred to in systemology as the model base [26]. A typical base, suitable for almost any property of the model, is time. The bases of the three main types - time, space, and group - can be combined.

An object model is a set of properties with which a set of their manifestations is connected, and a set of bases with which a set of its elements is connected. The object model in accordance with (2) can be formally defined as

$$O = \left(\{(a_i, A_i)|i \in N_n\}, \{\left(b_j, B_j\right)|j \in N_m\}\right) \quad (3)$$

where $N_n = \{1,2, \dots, n\}$, a $N_m = \{1,2, \dots, m\}$,; through a_i and A_i respectively, the property and its many manifestations are indicated.; b_j and B_j – base and many of its elements [27].

For information models of production and business processes, the set of features A_i and bases B_j in (3) are determined in studies on the information components of production systems [7, 10, 11]. These databases will be further used in the present work.

From the point of view of the introduced model (3), the variable will be called the operational representation of the property, i.e., the image of the property A_i, determined by the specific system examination procedure introduced above [27]. Each variable is associated with a certain set of values $\{a_i\}$. Through which it manifests itself. These values are states (or values) of a variable, and the whole set is a set of states.

The parameter is called the operational representation of the base B_j. Associated with each parameter is the set $\{b_j\}$. called the parametric set, and its elements - the values of the parameter [27]. If two or more parameters are used, then their common parametric set is the Cartesian product of individual parametric sets. Each specific parameter value (from the general parametric set) should identify a single observation of the corresponding variables [7].

When developing information systems, time, as a rule, enters the model as a parameter and not as a variable; therefore, they are often called planning systems [28]. Variable declaration of input (output) is usually done using the function

$$u: N_n \rightarrow \{0,1\} \quad (4)$$

such that if $u(i) = 0$, then the variable v_i is input and if $u(j) = 1$, then the variable $v_j$is output. In graph models, such defining functions are defined using the adjacency matrices of oriented graphs [1], in ER ++ models using entity interfaces [23, 29], etc. Any ordered set given by (4)

$$\boldsymbol{u} = (u(1), u(2), \dots, u(n)) \quad (5)$$

is called the determinant of the input-output [27].

The purpose of the functioning of an information system can be given by the description of its behavior [27]. The concept of behavior in systemology is used to obtain the characteristics of a general parametrically invariant restriction on model variables. For a given model [26], the range of possible types of parametrically invariant constraints depends on the properties described by the parametric set (3). If no properties are specified, then the possible states of the variables are mutually limited. If the parametric set is ordered (for example, it includes time), then the states of the variables can be limited not only by other states, but also by the states of the selected neighborhood for each specific parameter value. The neighborhood on an ordered parametric set is usually called a mask [27]. It is determined by a parametric set, a set of variables and a set of shift functions defined on the set $\boldsymbol{W}$ [27, 30]. Shift functions are single-valued functions.

$$r_j: \boldsymbol{W} \rightarrow \boldsymbol{W}, j = 1,2, \dots, n \quad (6)$$

The shift function puts each element $\boldsymbol{W}$ in correspondence with a single element $\boldsymbol{W}$, different from the original one. We will consider only fully ordered parametric sets (including time). If the object being modeled does not give direct ordering, we will introduce it artificially in the process of inspecting the object [7, 21, 10, 11, 30]. Following the methodology of parameterization of models [27], we introduce a completely ordered set of variables $\boldsymbol{V}$ and a similarly ordered set of vector shift rules $\boldsymbol{R}$ according to (6). The set of all variable values is then described by the $\boldsymbol{V} \times \boldsymbol{R}$ Cartesian product. The models consider the values given by some relation (a subset of the Cartesian product)

$$\boldsymbol{M} \subseteq \boldsymbol{V} \times \boldsymbol{R} \quad (7)$$

such that each pair $(v_i, r_j) \in \boldsymbol{M}$ corresponds to one shift equation. The relation $\boldsymbol{M}$ represents a neighborhood scheme on a parametric set, in terms of which the sample variables of the model are defined. This scheme is usually called a mask [27].
Thus, in studies of information systems and the design of modular devices [29], the method of artificially introducing the order (numbering) of the functions under study is used, while in the planning of works, their order is specified by binding events to relative time points.
A complex system can contain several subsystem masks, the order of which can also be set quite arbitrarily when the model is formalized. An example of layer-by-layer representation of the subsystem masks on Cartesian products (7) is shown in Fig. 1. for the case of analyzing an information system.

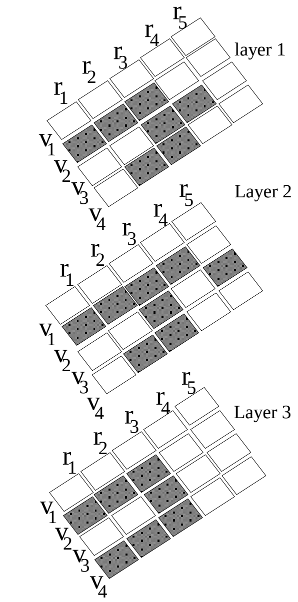

Figure 1. Three-dimensional mask representation for a model of a complex functional system based on the relation (7)

The model in fig. 1 is built on the basis of a single database obtained from a survey of the information structure of the enterprise, in which the enterprise's business processes are formed for each structural or functional information element of the organization (business process of the analytical department or business process of contract approval). In turn, all business processes are described in the form of binary square matrices, where the functions (6) act as rows and columns. In addition, each relationship (ie, "1" in a binary square matrix) is a nested matrix containing a certain set of elements that characterize this relationship, and contains a descriptive part of this relationship (type, membership in a hierarchical level, time constraints, transferable documents, etc.).
The three-dimensional model, in contrast to the previously known one-dimensional matrix models [26, 3] and others, promotes an understanding of how the whole process functions and what changes can be made to make the process more efficient. In a simulated information system, issues related to the organization of the process execution are highlighted. issues of establishing the sequence of tasks and issues of accounting information interdependence between tasks. Note that according to the model of fig. 1 it is easy to single out the typical part of the set of model functions, which can be formalized into a module, understood here as the largest immutable element of the structure of the model functions.
In addition, it is proposed to supplement the characteristics of the connections of the resulting mask with the characteristics "point of view" and "focus of attention" according to the ideology of the

Zachman scheme. This step will allow to determine the usefulness of the existing model, as well as to analyze the visualized model at various levels of abstraction from different points of view. From the point of view of the terminology introduced above, the formalization of the notion of a point of view is carried out by introducing an appropriate mask $\boldsymbol{M}$. Any mask represents a certain point of view by representing constraints on the basic variables [26]. The easiest way to specify a specific mask is to list all the full states of the corresponding variables in the subset (7), which determines the multidimensional relation on $\boldsymbol{V} \times \boldsymbol{R}$. This ratio is most often determined by the function.

$$f_B: (\boldsymbol{V} \times \boldsymbol{R}) \to \{\boldsymbol{0}, \boldsymbol{1}\} \quad (8)$$

such that $f_B(C_{ij}) = 1$ if the state of C_{ij}enters the mask, and $f_B(C_{ij}) = 0$ otherwise. Thus, the f_B function is a typical selection function [19]. Note that the f_B function determines the actually occurring states from $\boldsymbol{V} \times \boldsymbol{R}$, but does not determine the value of the parameter at which they occur. Thus, this function is parametrically invariant

4. Conclusions

The presented model of a modular information system allows you to select the most frequently repeated scenarios (functions) in the form of columns of matching mask cells, since all the scenarios are ordered in a single state space of a common production process scenario. After that, we can proceed to the modeling of individual functions in the form of some finite automata, the action of which is initiated by the input of some document (template) and generates other document (s) transmitted to other modules. The application of the proposed method allows you to create a modular structure of the information system based on the identification of recurring elements of the organization’s business processes.

5. Appendices

The reported study was funded by RFBR according to the research project № 18-07-00908 A.

6. References

[1] Melikhov A.N., Bershtein L.S., Kureychik V.M. The use of graphs for the design of discrete devices. - M.: Nauka, 1974.

[2] Mikita R.M., Rogozov Yu.I., Sviridov A.S., Stukotiy L.N. The Concept of Building an Enterprise Information Model. // Telecommunications. 2004, №8.

[3] Kuznetsov N.A., Kulba V.V., Kovalevsky S.S., Kosyachenko S.A. Methods of analysis and synthesis of modular information management systems.- M.: Fizmatlit, 2002, 800 p.

[4] Rogozov Yu.I., Butenkov S.A., Sviridov A.S. Systematization of models of the life cycle of information systems in the framework of the model J. Zakhman In Coll. Proceedings of the International Scientific and Technical Conference "System Analysis and Information Technologies" (SAIT-2007), Obninsk, September 10-14, 2007, v.2, pp.195-199.

[5] Jacobson A., Buch G., Rambo J. Unified software development process. - SPb .: Peter, 2002. - 496 p.

[6] D. L. Parnas. On the criteria to be used in decomposing systems into modules // Communications of the ACM. — 1972. — Vol. 15, no. 12. — DOI:10.1145/361598.361623.

[7] Kobern A. Modern methods of describing functional requirements for systems, M: Lori, 20 02. - 288c.

[8] Druppov S.A., Mikita R.M. Methods of phased formation of technical specifications for the development of an information system based on the results of an enterprise survey. News TSURE, 2006, №11.

[9] G. Vernikov. Basic Methodologies for Survey of Organizations. IDEF0 Standard // Information Service Director, No. 5, 2001.

[10] Rogozov Yu.I., Sviridov A.S. Application of software at pre-project examination, Taganrog “News of TSURE” No. 1, 2004, p.82
[11] Sviridov A.S. Evaluation of the effectiveness of the survey of the enterprise according to the combined methodology in the design of an automatic information system of the enterprise. Materials of the 59th scientific session of them. Popov. M.: 2004, issue LIX-1, pp. 93-94.
[12] Zykov A.A. Fundamentals of graph theory. M., N., 1987
[13] Vdovenko L.L. A system-information approach to assessing the economic activity of industrial enterprises. - M .: Economic education, 1996.
[14] Vendorov A.M. Designing software for economic information systems. Textbook M .: Finance and Statistics 2002. 325s.
[15] Mamikonov A.G., Kulba V.V. Synthesis of optimal modular SOD.- M.: Nauka, 1986, 276 p.
[16] Norenkov I.P. Introduction to computer-aided design of technical devices and systems.- M: Higher School, 1980, 308 p.
[17] Pole B.T. Introduction to optimization. - M .: Nauka, 1983.- 384 p.
[18] Emelichev A.I. and others. Polyhedra, graphs, optimization. M., N., 1981.
[19] Aliyev T.M. and others. Automation of information processes in integrated industrial control systems of industrial enterprises. - M .: Energoizdat, 1981.
[20] Bobnev S.V., Rogozov Yu.I., Methodology for constructing a functional oriented model of the organizational structure of an enterprise and its information filling. M .: Telecommunications. 2006, Issue 7, pp.30 - 34.
[21] Introduction to the information business. / Ed. V.P. Tikhomirova, A. Khoroshilova. - M .: Finance and Statistics, 1996.
[22] Minyaev MF, Information Technology Management: In 3 books. Book 2. Information Resources, M: Omega, 2003, 432 p.
[23] Bobnev S.V., Rogozov Yu.I., Butenkov D.S. The use of optimization models of information flows for the construction of Case-tools. News TSURE, Taganrog, 2005, no. 3, s. 54-60.
[24] Mikita R.M., Rogozov Yu.I., Sviridov A.S., Stukotiy L.N. The Concept of Building an Enterprise Information Model. // Telecommunications. 2004, №8.
[25] Mikita R.M., Stukotiy L.N. Method of construction and visualization of the information model of the enterprise // Coll. Proceedings of the Scientific and Technical Forum with international participation “High Technologies-2004”, Izhevsk, November 23-26, 2004, p.
[26] Druppov S., Mikita R. Methodology for the phased formation of technical specifications for the development of an information system based on the results of an enterprise survey // Zhurn. “Izvestiya TIT SFU. Technical Sciences, 2006, № p.
[27] Clear J. Systemology. Automation of solving system problems. M .: Radio and communication, 1990, 536 p.
[28] Ljung L. Identification of systems. - M .: Mir, 1975.
[29] Norenkov I.P. Introduction to computer-aided design of technical devices and systems.- M: Higher School, 1980, 308 p.
[30] Mironov S.V., Pishchukhin A.M. Metasystem approach to management: Monograph. - Orenburg: GOU OGU, 2004.- 338 p.