Resource-aware Approach in the Design of Complex Information Systems as a Problem of Multicriteria Optimization

Mikhail Tatur [1], Natalia Novoselova [2] and Viachaslav Prorovski [1]

[1] *Belarusian State University of Informatics and Radioelectronics, Pietrusia Brouki, 6, Minsk 220013 Belarus*
[2] *United Institute of Informatics Problems, Surganova 6, Minsk, 220012, Belarus*

Abstract

One of the modern information society tasks is the design of complex multiprocessor information systems, taking into account the rational use of system resources. It includes the development of a new resource-aware approach for the design of next generation information systems, such as smart home systems, autonomous transport, cyber-physical systems, etc.

The paper has an educational character and is intended for novice researchers with the aim of introducing them to the subject area. It considers a demonstrative explanation of the problem of resource allocation in the design of complex information systems. It describes the basics of the problem of complex assessments of the technical solutions and the formal approach for solving this problem, including the methods from the theory of Data Mining and Operation Research. It also considers the problem of finding a trade-off between system parameters when looking for new solutions.

Approaches to the comparative assessment of multiprocessor systems are demonstrated using simple examples. The ranking of the multiprocessor systems is performed in accordance with a complex criterion. Due to the specifics of technical solutions the formation of such a complex criterion is a rather difficult task, therefore many criteria (parameters) are used to make the assessment. The problem of finding the optimal variant of a multiprocessor system is presented as a problem of multicriteria optimization. The use of multicriteria optimization makes it possible to search for a trade-off between several alternatives or to find the non-dominant solutions, where none of them is better than the other in all the considered parameters and therefore have the same importance. In this case, many solutions are allowed, each of which is acceptable in the absence of preliminary information about the importance of the criteria.

From the set of several non-dominated solutions obtained as a result of multicriteria optimization, it is possible to choose one that is most preferable for a specific applied problem.

Keywords
resource-aware approach, multi-core information systems, data mining, multicriteria optimization

1. Main problems in the design of complex information systems

Novice researchers quickly master modern design technologies [1-2], work on the creation of specific technical solutions and, as a rule, do not pay attention to the complex nature of the work results. (The main thing is that the systems work, perform the prescribed functions). However, with a deeper analysis of the design results, the following problems inevitably arise. Let's name just a few of the most common ones.

The problem of complex assessments includes the following constituents.

1. Comparison of the obtained technical solution with the already known ones (reduced for solving a formal ranking problem).

2. Assessment of how effective the solution is, are there alternative solutions with a better combination of technical and economic parameters? (Reduced to a formal optimization problem).

3. Correctness assessment of technical task requirements for information system development, i.e. compliance of requirements with real restrictions due to the level of scientific, technological, economic, operational and other capabilities.

There are two approaches for solving these problems. The informal approach is represented by the well-known methods of expert assessments [3]. It is believed that experts intuitively "know" the hidden dependencies between

parameters, which allows them to better cope with making evaluative decisions, compared to developers who do not have such experience.

The formal approach uses methods from the theory of Data Mining and Operation Research [4]. This mathematical apparatus is aimed at analyzing data in a multidimensional feature space and allows you to evaluate solutions according to a given criterion or several criteria. The basic provisions are as follows:

−   each technical solution is characterized by a set of qualitative and quantitative parameters and can be represented by a point in a multidimensional space, and a set of alternative technical solutions to the same applied problem, respectively, by a set of points;

−   the technical requirements for the system act as constraints that can be represented by hypersurfaces in space, then the conformity of the technical parameters of the system and constraints is represented as the relative position of a point and a surface.

The key aspect of the problem of assessments in complex analysis is the establishment of relationships between the parameters based on the already obtained technical solutions. Note that this problem is universal and does not depend on a specific applied task.

The problem of finding a compromise is inevitable during design.

It is well known that most of technical solution parameters are interdependent, i.e. in an effort to improve the parameter of interest, other parameters may be degraded. Experienced developers understand these dependencies, and this helps them to find a trade-off between system parameters when looking for (developing) new solutions. However, in practice, it is very difficult (or impossible) to obtain these dependencies in an analytical form. Note that this problem (trade-off) is specific to each design domain (i.e., it is associated with the method of obtaining a technical solution).

The problem of visualizing the design progress and analyzing the results is connected with the above problems, and is primarily associated with work in a multidimensional (more than 3-x) feature space. The known visualization methods do not completely solve the problem, but only certain aspects. For educational purposes, the classical way of visualization is to reduce the number of features to two, with the subsequent generalization of the result to a multidimensional space.

We do not pretend to solve these problems in general terms or in a specific application area. In this article, we will provide a number of related demos that will be useful for novice researchers to familiarize themselves with the problem of resource "assessment and trade-off" in the design of information systems.

2. Demonstration example of solving the problem of comparative assessment of multiprocessor systems

Let there be several multiprocessor systems. Table 1 presents four multiprocessor systems, which are characterized by several parameters: the number of processors, performance, dimensions, power consumption. We will assume that these systems can programmatically solve the same problem, but are implemented in a different design.

Table 1
Parameters of multiprocessor systems

	Number of processors (items)	Performance (Gflops)	Dimensions (Volume, dm3)	Power consumption (W)
System 1	4/0,04	2,0/0,2	4,0/0,4	20,0/0,2
System 2	8/0,08	3,0/0,3	1,0/0,1	12,0/0,12
System 3	16/0,16	5,0/0,5	3,0/0,3	10,0/0,1
System 4	64/0,64	7,0/0,7	6,0/0,6	40,0/0,4

It is necessary to:
1. Give a comparative assessment (perform ranking) of the systems. Choose the most efficient system.
2. Make the comparative analysis of systems in terms of multiobjective optimization.

2.1. Performing ranking

It is not difficult to rank systems according to any of the four parameters. However, if we want to introduce a complex criterion of comparative assessment, taking into account several or all of the parameters, then we will encounter a number of difficulties and uncertainties. At first, we have parameters of a different nature, so we cannot operate with them in absolute terms. To obtain relative values, one of the normalization methods is used, suitable for the problem being solved.

For clarity and simplicity of reasoning, we use the simplest normalization method, when the parameters are reduced to some fixed maximum value (in our case, 100, 10, 10 and 100, respectively). The normalized values of the parameters of multiprocessor systems are shown in Table 1 in the corresponding columns after the sign (/).

If the number of processors and performance are known, statistically confirmed parametric dependencies (Figure 1a), then the dependences of dimensions and power consumption in this example are not visible (Figure 1b). (Although, in theory, we know that with an increase in the number of processors and performance, the power consumption and size should increase). But this example deliberately does not provide any information about the technology, element base, design, purpose, generation, etc., which directly affects both the dimensions and power consumption of the implemented systems.
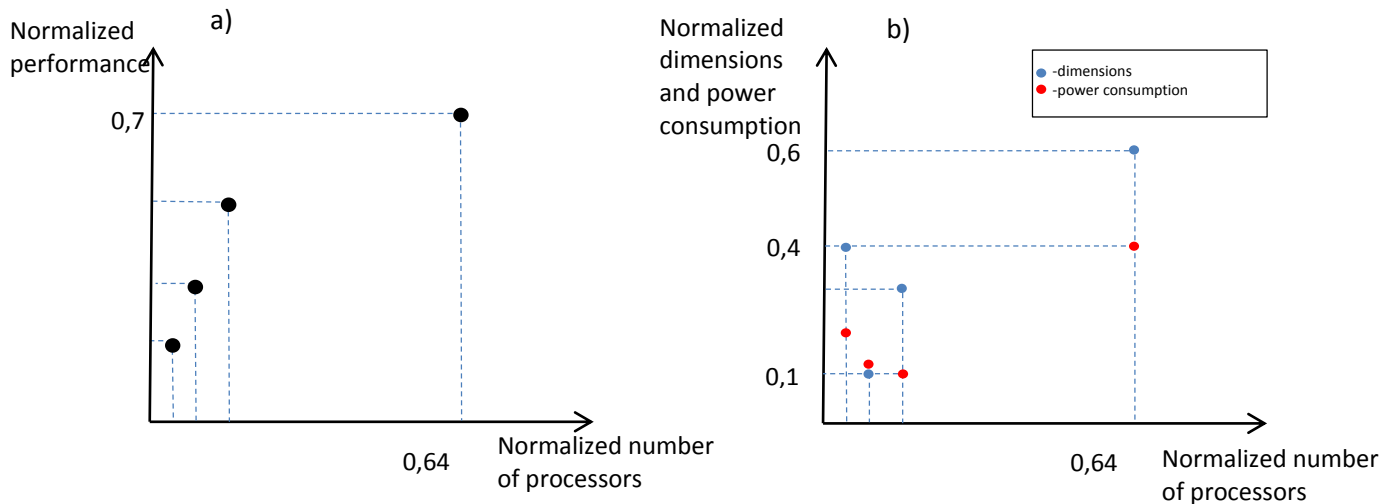

Figure 1: Distribution of multiprocessor systems in various two-dimensional normalized spaces of features

Next, we must offer (justify the choice) of a comprehensive criterion for comparative assessment. And not everything is obvious here! To bring together the parameters of performance, dimensions and power consumption into a single indicator, it is necessary to enter weighting factors. We will assume that it makes no sense to include the number of processors in the complex criterion, since it is indirectly in the performance of systems. (The points on the graphs on Figure 1 do not yet reflect the ratios of the significance of the parameters in making evaluative decisions).

Let us consider the formation of a complex criterion for the comparative assessment of systems as a weighted sum of three criteria. In this case, it is necessary to perform the following steps:
1. Bringing the criteria to a single range of values (normalization or scaling of the criteria):

$$\bar{x}_i = \frac{x_i - x_i^{min}}{x_i^{max} - x_i^{min}},$$

where x_i^{min} is the minimal value of ith criterion, x_i^{max} – is the maximal value of ith criterion. For our example $x_i^{min} = 0$ for $i = 1,2,3$ and $x_1^{max} = 10, x_2^{max} = 10, x_3^{max} = 100$.

2. Selection of weight coefficients w_1, w_2, w_3 for each criterion: performance, dimensions and power consumption.
3. Ranking solutions according to a complex criterion $F(x) = -w_1 x_1 + w_2 x_2 + w_3 x_3$.
The minimum value of the criterion corresponds to the most optimal solution.

The normalized values of the parameters of the systems are presented in Table 1. Let us choose the following values of the weight coefficients $w_1 = w_2 = w_3 = 1$.

Let's calculate the values of the complex criterion for each of the systems $F_1(x) = 0,4, F_2(x) = -0,08, F_3(x) = -0,1, F_4(x) = 0,3$. Thus, the systems presented in Table 1 according to the complex criterion can be ranked as follows System 3, System 2, System 4 and System 1 in ascending order of the criterion value. The most optimal system is System 3.

The disadvantage of this ranking method is the dependence of the optimal solution on the choice of the values of the weight coefficients w_1, w_2, w_3. For example, if the value of the weighting factor for the parameter dimensions is set equal to $w_1 = 2$, then system 2 will be selected as the most optimal.

Due to the fact that the formation of such a complex criterion is a rather difficult task, due to the specifics of technical solutions, it is of interest to use many criteria (parameters), namely, finding the Pareto front of optimal solutions − non-dominated solutions.

2.2. Comparative evaluation of multiprocessor systems as a problem of multicriteria optimization

The problem of finding the optimal variant of a multiprocessor system can be presented as a problem of multicriteria optimization. All parameters of a technical solution presented in Table 1 can be considered as points in a multidimensional feature space, which allows finding optimal solutions according to some complex criterion or simultaneously according to several criteria.

As a rule, the criteria are interdependent, i.e. an increase in one of them can lead to a decrease in the value for the other; therefore, the use of multicriteria optimization makes it possible to single out compromise or non-dominant solutions, where none of them is better than the other in all the parameters under consideration and, therefore, are of equal importance. In this case, many solutions are allowed, each of which is acceptable in the absence of preliminary information about the importance of the criteria.

In general, the problem of multicriteria optimization is formulated as follows:

Find the vector $\bar{x}^* = \left(x_1^*, x_2^*, \ldots, x_n^* \right)^T$ of the values of parameters, which satisfy m inequalities:

$$g_i(\bar{x}) \geq 0, \quad i = 1, 2, \ldots, m, \tag{1}$$

and p equalities

$$h_i(\bar{x}) = 0, \quad i = 1, 2, \ldots, p, \tag{2}$$

and optimize the vector function

$$\bar{f}(\bar{x}) = \left[f_1(\bar{x}), f_2(\bar{x}), \ldots, f_k(\bar{x}) \right]^T. \tag{3}$$

Constraints (1), (2) define the domain G, which contains all feasible solutions to the problem. The vector $\bar{x}^*$ corresponds to the optimal solution in the domain G. In this case, the optimality is meant according to the Pareto concept, the formal definition of which from the point of view of the maximization problem is as follows [5]:

The vector of solution $\bar{x}^*$ is called Pareto-optimal if and only if there is no other vector $\bar{x}$, which dominate $\bar{x}^*$, i.e. if

$$\forall i \in 1, 2, \ldots, k, \quad f_i(\bar{x}) \leq f_i(\bar{x}^*) \text{ и } \exists i \in 1, 2, \ldots, k, \text{ где } f_i(\bar{x}) < f_i(\bar{x}^*).$$

In other words, $\bar{x}^*$ is Pareto optimal if there are no acceptable vectors $\bar{x}$, which allow you to increase the value of one of the criteria, while not decreasing the value of at least one of the remaining criteria. The solution $\bar{x}^*$ is strictly dominates the solution $\bar{x}$, if $\bar{x}^*$ is betters than $\bar{x}$ by all criteria. In general, the solution to the optimization problem is a set of non-dominated solutions.

Having a set of several non-dominated solutions obtained as a result of multicriteria optimization, it is possible to choose a solution that is most preferable for a specific applied problem.

Thus, the assessment of the systems presented in Table 1 is based on the concept of non-dominated decisions and considers the value of each criterion separately. In our case we will perform a manual assessment of the systems. To solve more complex problems with a large number of alternatives, characterized by a large number of parameters, special methods of multicriteria optimization are used [5-7].

In the presented example, the criteria to be optimized correspond to the parameters of the system, and the problem of multicriteria optimization in this case can be written in the following form:

Optimize vector function $F(x) = [x_1, x_2, x_3]$, such as
$$x_1 \rightarrow max, x_2 \rightarrow min, x_3 \rightarrow min$$
and $a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, a_3 \leq x_3 \leq b_3$ – allowed parameter range,
where x_1, x_2, x_3 - performance, dimensions and power consumption of systems.

Consider systems 1 and 2, where the value of the performance criterion of System 2 are higher than the corresponding value of System 1, and the value of the dimensions and capacity criteria for System 2 are lower. Consequently, System 2 has more optimal parameter values and thus dominates System 1. If we compare Systems 2 and 3, then we see that System 3 has higher performance and less power consumption, while its dimensions are larger than the dimensions of System 2. Therefore, Systems 2 and 3 are non-dominated solutions. Likewise, the non-dominated solutions are Systems 2 and 4, Systems 3 and 4.

In Figure 2 the systems under consideration are shown as points in the space of two criteria. Non-dominated solutions are connected with a line.
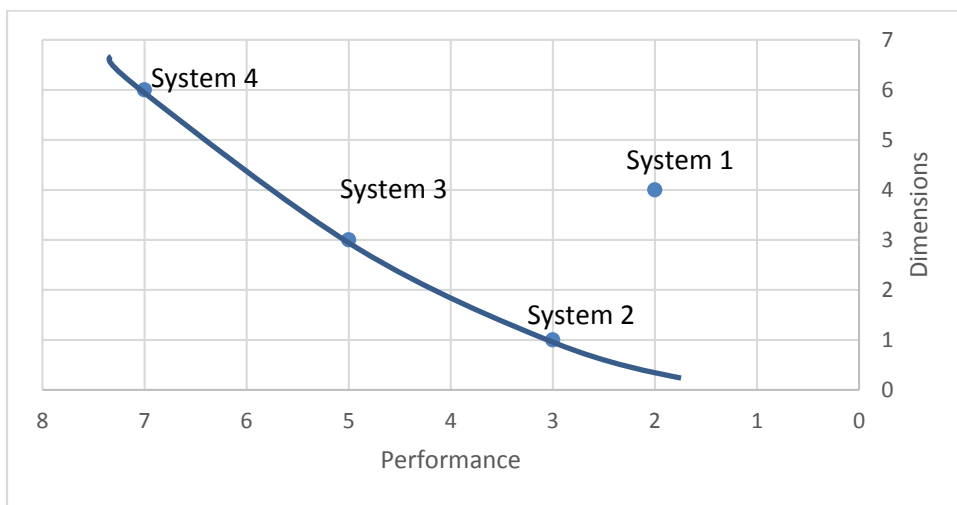

Figure 2: Graphic presentation of technical solutions in a two-criterion space

It should be noted that systems 2,3,4 are not dominant in the specific problem under consideration, i.e. they are locally non-dominated solutions. In the case of adding additional alternative solutions for multiprocessor systems, the number and composition of non-dominated solutions may change, although in this case the solutions will be locally non-dominated. According to the analysis, System 1 can be excluded from further consideration due to the fact that it is the dominant solution for all the considered optimization criteria.

3. Conclusion

The problem of "trade-off" in the design of complex information systems continues to be relevant. Problems of complex assessment and visualization of design results coexist closely with it. The origins of the problems lie in the fact that the course and results of the design are influenced by many factors, and it is not possible to correctly assess the degree of influence of some parameters. The known methods of data analysis and optimization only partially allow formalizing the "trade-off" in the design process. However, their value cannot be underestimated, since when applied correctly, they show the developer the direction to search for new technical solutions and the resources that can be used for this.

4. Acknowledgements

5. References

[1] R. F. de Magalhães, A. D. Danilevicz, J. Palazzo. Managing trade-offs in complex scenarios: A decision-making tool for sustainability projects, Journal of cleaner production 212 (2019) 447-60.

[2] K. N. Otto, E. K. Antonsson, Trade-off strategies in engineering design, Research in Engineering Design 3(2) (1991) 87-103.

[3] C. W. Brownley, Multi-objective decision analysis: Managing trade-offs and uncertainty, Business Expert Press, 2013.

[4] C. Dudas, A. H. Ng, L. Pehrsson, H. Boström, Integration of data mining and multi-objective optimisation for decision support in production systems development, International Journal of Computer Integrated Manufacturing 27(9) (2014) 824-839.

[5] K. Deb, Multi-objective Optimization Using Evolutionary Algorithms, John Wiley and Sons, Ltd, England, 2001.

[6] C. M. Fonseca, P. J. Fleming, Genetic algorithms for multi-objective optimization: Formulation, discussion and generalization, in: Proceedings of the Fifth International Conference on Genetic Algorithms, Morgan Kaufmann, San Mateo, CA, 1993, pp. 416-423.

[7] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach, IEEE Transactions on Evolutionary Computation 3(4) (1999) 257-271.