On the optimal scheme of the possible three predictors in discrete optimization problems using decision-making algorithms

Yulia Terentyeva

Abstract - **The computational complexity of algorithms used to solve discrete optimization problems dictates the need for optimal use of resources, the main of which are time and memory. The article will consider situations when the main algorithm uses predictors, the result of the analysis /solution of which is iteratively taken into account. The situations of one, two or three predictors are analyzed in detail, as well as various schemes of decisive rules, including the so-called majority vote or unanimous vote. The result of the study is evidently presented. According to the obtained result, the most effective scheme is one of the best among the available three predictors. This result has a double effect, since it eliminates the need for the main algorithm to allocate resources for the work of other predictors. Thus, we propose a proven optimal scheme for using predictors in the process of algorithms for solving discrete optimization problems, which are often NP-complex. In addition, the proven optimal scheme of using predictors can be extended to the functionality of decision support systems, when the role of predictors can be played by experts, in particular people or programs for analyzing situations and making decisions.**

Keywords – **heuristic algorithms, discrete optimization problems, graph theory models, greedy algorithms.**

I. INTRODUCTION

The development of modern communication networks is accompanied by the emergence of a whole range of problems arising due to the high dimensions of both the graphs of communication network models and the parameters characterizing communication network objects. At the same time, many problems are NP-hard, and this circumstance, along with the mentioned high dimensionality, determines the need for new research and the development of methods for solving them [1-3]. There are also current problems of polynomial complexity, which, also due to their high dimensionality, stimulate research in the direction of reducing algorithmic complexity. Similar studies in development [4], in particular, led to a result that can be used not only in communication networks.

The article will focus on the concept of decision making during the operation of algorithms. A similar concept of decision making can be generalized to a certain set of communication network problems associated with

Manuscript received November 27, 2023

high-dimensional optimization problems that involve obtaining a pseudo-optimal solution.

Decision-making concepts within the framework of using predictor functions are a frequent prerogative of discrete optimization problems, since the algorithm solving the problem involves selecting a branch of further search for a pseudo-optimal solution, where each branch must be evaluated using predictor functions. As for the types of discrete optimization problems mentioned, where this analysis can be applied, these are, for example, the traveling salesman problem, the scheduling problem.

II.GENERAL DESCRIPTION OF THE CONSIDERED ALGORITHMS

The subject of this article is derived from previous work related to several different components. Firstly, the immediate subject area, the problem of reflectometry, initially considered by the author in [5–8]. The second prerequisite for the material of this article is auxiliary decision-making algorithms for the iterative solution of a discrete optimization problem: such iterative algorithms, as a rule, are either the simple application of some greedy heuristics at each iteration step or, in a more complex case, the branch and bound method; the latter also includes the use of greedy auxiliary heuristics - however, due to natural constraints on execution time, the auxiliary heuristics here tend to be much simpler. However, in both cases there may be several such greedy heuristics - and such situations were briefly considered in [8], in more detail in [9]. At the same time, one of the problems studied in these works was to search for situations when the greedy heuristic, a priori assumed to be the best, should be replaced as a result of the use of additional algorithms - however, in the discrete optimization problems considered in [8,9], no such situations were found was.

The third prerequisite is related to the use - as one of the possible algorithms for multi-criteria optimization - the so-called. dynamic risk functions [10]. If we talk directly about discrete optimization problems (which is also the reflectometry problem), then in them the possible use of dynamic risk functions is associated with auxiliary decision-making algorithms, which we above conventionally called "the second prerequisite for the material of the article."

The fourth prerequisite relates to the creation of an optimal scheme for organizing expert opinions, see [11], which can be effectively used when considering and solving dissertation optimization problems.

In this case, options for organizing an iterative solution, as well as methods for studying the quality of the resulting

Yulia Yu. Terentyeva, Head of the Department of Analysis and Methodology for Improving Information and Telecommunications Systems, Center for Information Technologies and Systems of Executive Authorities (19 Presnensky Val St., building 1, Moscow 123557, Russian Federation), Cand. Sci. (Eng.), terjul@mail.ru, http://orcid.org/0000-0002-2418-003X

algorithms, can be considered in the future. Among similar options for organizing an iterative solution, we mention the so-called. tournament self-learning, [10], etc. It can often be used for several variants of the predictor function - however, in the end everything can be reduced to a comparison between pairs of algorithms; if necessary, the corresponding tournament between a group of algorithms can be organized a posteriori, or more precisely, it can be implemented after the comparison of two algorithms of a certain pair has been implemented.

So, this article examines the effectiveness of using predictors in the operation of discrete optimization algorithms (recall that there must be some iterative process).

The question immediately arises: how to describe several such predictor functions, given that in the cited publications (including the classic monograph [12], as well as in most of the works cited below) only one such function is considered? However, this is possible (apparently, always): for example, in reflectometry problems, each predictor can be made on the basis of some vertex that has not yet been included, and he (the predictor) evaluates the choice of some (other) vertex for inclusion "from his point of view ", that is, it returns a result depending on how much (and whether) the distance from the nearest illuminated vertex has changed

Note that this article develops the approach briefly described in [8,9]. For the problem under study, all this will need to be considered for some completely different predictors [13], as well as for using the branch and bound method [14]; and, of course, it will be necessary to apply a similar technique to other discrete optimization problems.

So, in this work the following simplifications will be used. We will consider several predictors, but since the process is iterative, at each step we must select exactly one solution despite the fact that predictors can give similar "real" results, which, apparently, will correspond to the same "Boolean" results. We can, simplifying it a little, say that the subject of the article is how to make a decision about choosing one element in the case when there are no more than 3 predictors, and we work with their "Boolean" results.

III. ON INCREASING THE EFFICIENCY OF PREDICTOR ALGORITHMS

In this section we will consider questions about possible ways to increase the efficiency of algorithms for selecting a resolving element.

Let us consider the probability of making the right decision if algorithms use predictor functions. To do this, we will simulate situations where the so-called expert group includes from 1 to 3 predictors. At the same time, we will consider variants when they make correct decisions either with the same probabilities or with different probabilities.

Variant 1. The expert group consists of 1 predictor - the case is trivial. The final probability coincides with the initial one.

Variant 2. Expert group – 2 predictors.

Variant 2.1. Consider the case when both predictors make

correct decisions with the same probability, that is $p_1 = p_2$. Then, according to the concept of the final block making the correct decision, namely, if an element is considered good when both predictors evaluate it as good, the probability of

making the correct decision will be equal to $p = p_1^2$ (product probability formula). Let us repeat that the independence of expert predictors is assumed.

Variant 2.2. Case $p_1 \neq p_2$. We get $p = p_1 p_2$ (according to the product probability formula).

Variant 3. Expert group -3 predictors. There are two possible approaches to constructing a final decision-making scheme. The first approach implements the concept of making a decision "by majority vote", and the second - "by unanimous vote". Let's consider each of these approaches, and it is advisable to consider cases of identical and different probabilities for predictor programs to make the right decision. The natural assumption is made that predictor programs make decisions independently.

Variant 3.1. The same probabilities of making the right decision for all three predictors, that is $p_1 = p_2 = p_3$. Then, within the framework of the approaches under consideration, the following probabilities of making the correct decision with the entire set of predictors occur.

Variant 3.1.1. Unanimous decision. In this case $p = p_1^3$ (product probability formula).

Variant 3.1.2. By majority vote, we get $p = -2p_1^3 + 3p_1^2$ (formula for addition and multiplication of probabilities or Bernoulli's formula).

Variant 3.2. Different probabilities of making the right decision for different predictors. That is $p_1 \neq p_2 \neq p_3$.

Variant 3.2.1. Unanimous decision. In this case $p = p_1 p_2 p_3$ (product of probabilities formula).

Variant 3.2.2. By majority vote. Let us denote A_i the event at which the *i* - predictor makes the correct decision. Then the desired probability of making the right decision using several predictors - according to the "majority voting" approach - can be found as follows:

$$
p = P(A_1A_2A_3 + \overline{A}_1A_2A_3 + A_1\overline{A}_2A_3 + A_1A_2\overline{A}_3)
$$
 (1)

Further, according to the formula for adding and multiplying probabilities,

$$
p = P(A_1A_2A_3) + P(\overline{A}_1A_2A_3) + P(A_1\overline{A}_2A_3) + P(A_1A_2\overline{A}_3)
$$
 (2)

Since the predictor functions work independently, then

 $(p = p_1 p_2 p_3 + (1 - p_1) p_2 p_3 + p_1 (1 - p_2) p_3 + p_1 p_2 (1 - p_3)$ (3)

IV. BRIEF ANALYSIS OF THE OBTAINED RESULTS (PROBABILITIES)

A natural question arises about the best option for a group of predictors - from the point of view of their quantitative composition. For clarity, the previously obtained probabilities are presented in the following Table 1.

Let us carry out a pairwise comparison of the probabilities of making a correct decision by a group of predictor functions for different ways of organizing this group. Here we will omit

a detailed analysis of inequalities with parameters. The results of the analysis are shown in Table 2.

Table 2. **Analysis of conditions for preference of options for organizing groups of predictors.**

Variant 7 deserves special attention, in which three expert predictors vote according to the majority vote principle. Based on the fact that this option a priori turned out to be no worse than most of the other options considered, it is legitimate to ask the question about the maximum possible efficiency of this predictor organization scheme. This problem is obviously equivalent to the following problem (in the general case, let's consider the problem on an extremum under restrictions):

$$
xy + yz + xz - 2xyz \rightarrow extr
$$

\n
$$
0 \le x \le 1
$$

\n
$$
0 \le y \le 1
$$

\n
$$
0 \le z \le 1
$$

\n(4)

This function is continuous and defined on a compact set. According to Weierstrass's theorem, it reaches its maximum and minimum. Either at local extremum points, or at the border. To find the critical points, we use Fermat's theorem, which gives the necessary condition for the existence of an extremum. We have

$$
\begin{cases}\ny + z - 2yz = 0 \\
x + z - 2xz = 0 \\
y + x - 2xy = 0\n\end{cases}
$$
\n(5)

The Hessian matrix looks like this:

$$
\begin{bmatrix} 0 & -2z+1 & -2y+1 \ -2z+1 & 0 & -2x+1 \ -2y+1 & -2x+1 & 0 \end{bmatrix}
$$
 (6)

Solving the system gives two points $(0, 0, 0)$ and $(1, 1, 1)$. The Hessian matrix at these points is of indefinite sign. Therefore, these are the saddle points of the function.

When searching for extrema on the boundary, we obtain the following. The function reaches a minimum of 0 on the continuum of points when any two coordinates are zero, the third from the set range. And also the function reaches a maximum of 1 on the continuum of points when any two coordinates are one, the third from the established range. This fact can be interpreted this way. When organizing an expert group of three predictors with voting based on the majority vote principle, if two expert predictors always make the right decisions, then even if there is one poor-quality predictor or simply a predictor with an unstable probability of choosing the right decision, maximum efficiency is still achieved.

Thus, in order to say which combination of predictors will give the best result on average, it is necessary to estimate the probability of fulfilling the preference conditions for each pair of options under consideration. Given the specific probabilities of specific predictors to make the right decisions, you can use the compiled comparative table 2 to decide on the most effective scheme for organizing the work of predictors under given conditions.

V.INTERPRETATION OF THE RESULTS OF THE COMPARATIVE ANALYSIS

The conducted comparative analysis of the schemes for organizing a group of predictors allows us to draw the following conclusions. Options 6 and 4, in which three predictors in the expert group vote with different or equal probabilities according to the principle of unanimous voting, lose to almost all other options (option 6 may give better results than option 2 under certain conditions, but most likely fulfillment of conditions under which option 2 is preferable to option 6). Options 2 and 3 also turned out to be untenable, when the expert group consists of two predictors. It makes sense to consider options 1, 5 or 7. Their preference over each other is analyzed and reflected in Table 2. Since option 5 is a special unlikely case of option 7, option 5 gives a gain only with a low probability of making the right decision at least at least one of the expert predictors, it makes sense to focus on options 1 or 7 of the organization of the expert group.

VI. COMPARISON OF TWO "BEST" OPTIONS FOR ORGANIZING AN EXPERT GROUP

Without a doubt, with fixed probabilities of potential predictors of an expert group, a decision can be made on the most effective scheme for its organization. However, the above analysis will allow us to give an a priori assessment of the preferability of the organization scheme of the expert group under consideration. To do this, it is necessary to evaluate the measures of the areas corresponding to the preference of one or another option. It is advisable, as already noted, to consider option 1, when there is only one expert, and option 7, when the group consists of three predictors voting according to the majority vote principle. In this case, the assumption is made that in a group of three predictors one is precisely a representative of the expert group of option 1.

Since, as indicated in Table 2, option 7 is preferable if the condition is met

$$
p_3 > \frac{1}{1 + (\frac{1}{p_1} - 1)(\frac{1}{1 - p_2} - 1)}
$$
\n(7)

let's take a closer look at this expression.

In order to evaluate how a given surface breaks up the compact under consideration, defined by the probabilities of making the correct decision by the predictors of an expert group of three predictors (note: voting according to the majority vote principle), it is necessary to calculate the corresponding double integral.

$$
f = \int_{0}^{1} \int_{0}^{1} \frac{1}{1 + (\frac{1}{x} - 1)(\frac{1}{1 - y} - 1)} dx dy = \int_{0}^{1} \int_{0}^{1} \frac{x(1 - y)}{x(1 - y) + y(1 - x)} dx dy = \int_{0}^{1} \int_{0}^{1} 1 - \frac{y(1 - x)}{x(1 - y) + y(1 - x)} dx dy
$$
\n(8)

$$
f = 1 - \int_0^1 \int_0^1 \frac{y(1-x)}{x(1-y) + y(1-x)} dx dy = 1 - f \implies f = \frac{1}{2}
$$
 (9)

It follows that the surface under consideration splits the compact into two equal parts. Generally speaking, this means that if groups of experts of options 1 and 7 are formed in an arbitrary order, then none of the options will provide a significant gain. It is also worth noting here that the use of the concept of geometric probability here is not entirely appropriate, since the experiment does not satisfy the conditions of the geometric definition of probability, when its outcomes can be represented by points of a certain $\Omega \in \mathbb{R}^3$ in such a way that the probability of a point falling into any part $A \subset \Omega$ would not depend on the shape or location A inside Ω , but would depend only on the measure of the region *A* and, therefore, would be proportional to this measure. That is, the point (p_1, p_2, p_3) does not have a uniform distribution in the area Ω . At the same time, it should be noted that if a certain order is given to the formation of expert groups, this will lead to a narrowing of the area of outcomes, which in turn will allow us to make the assumption of a uniform distribution of the point (p_1, p_2, p_3)

in the area Ω . This assumption will also be valid taking into account the fact that, due to a number of circumstances, for a specific predictor-expert, the probability of making the right decision may change during the day, if this case with predictors is extended to real experts. Then, by assessing the measures of areas of preference for options for organizing an expert group, one can draw a conclusion about the quality of one or another option from the point of view of the resulting probability of making the right decision by the entire group.

Without limiting the generality of the argument, it is legitimate to assume that $p_1 \geq p_2 \geq p_3$. This assumption occurs because among a group of three, the selected predictor-expert for the group of option 1, consisting of one predictor, should obviously have the maximum probability among the three of making the right decision. Note that if $p_1 = p_2$, then option 7 will give a win only in the case $p_3 > 0.5$ (according to the above reasoning, the first and second experts have qualifications characterized by a higher 0.5probability of making the right decision). To assess the degree of preference of option 7, subject to $p_1 \geq p_2$ and $p_1 \geq p_3$, it is necessary to calculate the following double integrals.

$$
mes\ D_{1} = \int_{0}^{1/21} \int_{y} x dx dy + \int_{1/2}^{1} \int_{1/2}^{x} \frac{x - xy}{x + y - 2xy} dy dx \quad (10)
$$

The area D_1 under consideration, according to paragraph 6 and the above reasoning, is characterized by a system of

inequalities: $x + y - 2xy$ $x > y, x > z, z < \frac{x - xy}{x + y - 2}$ $> y, x > z, z < \frac{x - z}{x}$. The

.

boundary surfaces are shown in Fig. 2.

Here the boundary $y = 0.5$ is obtained as a result of the

intersection of surfaces $z = x$ and $z = \frac{x - xy}{2}$ $+y-2$ $=\frac{x-$ Omitting intermediate calculations, we get

$$
mes\ D_1 = \frac{2}{3} - \frac{1}{2} \ln 2 \tag{11}
$$

Since the volume of the simplex $\{(0, \frac{1}{2})\}$ 1 $(0, 0)$ $(1, \frac{1}{2})$ $\frac{1}{1}$,0) (1,1,0)

(2 $\frac{1}{2}$, 2 $\frac{1}{2}$, 2 1 $(1,\frac{1}{2})$ 1 ,1) (1,1,1)} is equal to $\frac{6}{48}$ $\frac{5}{10}$, then under the conditions $x + y - 2xy$ $x > y, x > z, z > \frac{x - xy}{x + y - 2}$ $> y, x > z, z > \frac{x - xy}{1}$ defining the

area D_2 preferences of option 7, the measure of the area D_2 is equal to

mes
$$
D_2 = \frac{5}{48} - \left(\frac{2}{3} - \frac{1}{2}\ln 2 - \frac{11}{48}\right) = \frac{1}{2}\ln 2 - \frac{1}{3}(12)
$$

Let us estimate the ratio of area measures. We have

Let us estimate the ratio of area measures. We have

$$
\frac{mes\ D_1}{mes\ D_2} = \left(\frac{2}{3} - \frac{1}{2}\ln 2\right) \left(\frac{1}{2}\ln 2 - \frac{1}{3}\right) = \frac{4 - 3\ln 2}{3\ln 2 - 2} \approx 24.2\tag{13}
$$

Let us note again that this D_1 corresponds to the preference area of option 1, and $D₂$ corresponds to the area of preference of option 7.

Fig. 1. Surface delimiting the preference areas of the two best options

Thus, option 1 of organizing an expert group from one predictor has a quality indicator that is more than 24 times better than the quality indicator of option 7, when the expert group consists of three predictors voting according to the majority vote principle. Moreover, this one predictor (option 1) must be one of three in the group of option 7 and have the maximum probability of making the right decision among the

three. Note that the area characterizing the advantage of option 7 over option 1, as already noted, is 24.2 times smaller than the area of inverse preference and, remarkably, corresponds to a situation where all three expert predictors have the probability of making the right decision , close to unity.

According to the comparison, the possible potential damage when organizing an expert group of three predictors voting according to the majority vote principle will result in approximately 24 times greater damage than when organizing an expert group from one predictor, provided that this predictor has the highest probability among the three making the right decision.

Thus, the recommended strategy is one expert predictor with the highest probability of making the right decision.

VII. CONCLUSION

Thus, a demonstrative comparison of the main combinations of options for using predictor functions in discrete optimization algorithms has been carried out. Combined methods of organizing groups of predictors were considered and studied using probabilistic methods. Probability estimates were obtained and the conditions for the effectiveness of various schemes for constructing groups of predictors were determined, taking into account the quantitative composition (from one to three), as well as strategies for organizing voting (by majority vote or unanimously). A detailed comparison table ("each with each") of various schemes has been compiled, showing the ratio of parameters in which one method can be a priority over another. An integral comparison was carried out for all combinations of circuit construction. It has been proven that in the vast majority of cases, the best quality indicators in terms of the probability of making the right decision will be possessed by an expert group of predictors, consisting of one expert predictor with the highest probability of making the right decision among all the others. As a special case, it is shown that increasing the number of predictors in most cases does not lead to a noticeable increase in the efficiency of selection, but can significantly increase the complexity of calculations. As a result, a strategy of using one predictor with the highest probability of making the right decision is recommended.

The results of the study can also be used in organizing schemes for taking into account expert opinions in various systems that require binary filtering of objects using filtering tools, including humans.

REFERENCES

- [1] Yu. Yu. Terentyeva, Determination of the Maximum Set Independent Simple Paths between the Vertices of the Graph. *Modern Information Technologies and IT-Education*, vol. 17, no. 2, p. 308-314, 2021. (In Russ., abstract in Eng.) doi: https://doi.org/10.25559/SITI-TO.17.202102.308-314
- [2] Yu. Yu Terentyeva, Algorithms for calculation of reliability of large-scale communication networks. *Informatization and communication*, no. 6, p.

171-175, 2021. (In Russ., abstract in Eng.) doi: https://doi.org/10.34219/2078-8320-2021-12-6-171-175

- [3] B. Melnikov, Discrete optimization problems some new heuristic approaches. In: Eighth International Conference on High-Performance Computing in Asia-Pacific Region (HPCASIA'05), Beijing, China, 2005, pp. 8 pp.-82. doi: https://doi.org/10.1109/HPCASIA.2005.34
- [4] Yu. Yu. Terentyeva, Some theoretical issues related to the description of practical algorithms for constructing spanning trees. *International Journal of Open Information Technologies*, vol. 9, no. 3, p. 23-27, 2021. (In Russ., abstract in Eng.) EDN: KRVLTR
- [5] B. F. Melnikov, Yu. Yu. Terenteva, On a graph model for reflectometry issues and some algorithms for their solution. Part 1. Issue statement and approaches to algorithmics. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Fiziko-matematicheskie nauki* = University proceedings. Volga region. Physical and mathematical sciences, no. 2, p. 28-39, 2022. (In Russ., abstract in Eng.) doi: https://doi.org/10.21685/2072-3040-2022-2-3
- [6] B. F. Melnikov, Yu. Yu. Terenteva, On a graph model for reflectometry issues and some algorithms for their solution. Part 2. An approach to software implementation. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Fiziko-matematicheskie nauki* = University proceedings. Volga region. Physical and mathematical sciences, no. 3, p. 32-42, 2022. (In Russ., abstract in Eng.) doi: https://doi.org/10.21685/2072-3040-2022-3-4
- [7] B. F. Melnikov, Yu. Yu. Terentyeva, On a graph model for reflectometry problems and some algorithms for their solution. Part III. An approach to test data generation and some results of computational experiments. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Fiziko-matematicheskie nauki* = University proceedings. Volga region. Physical and mathematical sciences, no. 4, p. 31-41, 2022. (In Russ., abstract in Eng.) doi: https://doi.org/10.21685/2072-3040- 2022-4-3
- [8] S. V. Baumgertner, B. F. Melnikov, Generalized indeterministic finite automata. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Fiziko-matematicheskie nauki* = University proceedings. Volga region. Physical and mathematical sciences, no. 2, p. 64-74, p. 2013. (In Russ., abstract in Eng.) EDN: RLEENF
- [9] S. V. Baumgertner, *Mul'tijevristicheskij podhod k zvjozdno-vysotnoj minimizacii nedeterminirovannyh konechnyh avtomatov*: PhD dissertation. Tol'jatti, 2012. 123 p. (In Russ.)
- [10] B. F. Melnikov, Heuristics in Programming of Nondeterministic Games. *Programming and Computer Software*, vol. 27, no. 5, p. 277-288, 2001. (In Russ., abstract in Eng.) doi: https://doi.org/10.1023/A:1012345111076
- [11] Yu. Yu. Terentyeva, On the optimal scheme for organizing an expert group, taking into account the voting strategy for making a collegial decision. *Informatization and communication*, no. 4, p. 138-143, 2014. (In Russ., abstract in Eng.) EDN: SZQYSD
- [12] Gudman S., Khidetniemi S. Vvedenie v razrabotku i analiz algoritmov = Introduction to the design and analysis of algorithms. Moscow: Mir, 1981:368. (In Russ.).
- [13] I. K. Sigal, A. P. Ivanova, *Vvedenie v prikladnoe diskretnoe programmirovanie: modeli i vychislitel'nye algoritmy* [Introduction to Applied Discrete Programming: Models and Computational Algorithms]. Moscow: Fizmatlit; 2007. 304 p. (In Russ.)
- [14] Melnikov B.F., Eyrih S.N. On the approach to combining truncated branch-and-bound method and simulated annealing. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Serija: Sistemnyj analiz i informacionnye tehnologii* = Proceedings of Voronezh State University. Series: Systems Analysis and Information Technologies, no. 1, p. 35-38, 2010. (In Russ., abstract in Eng.) EDN: MUPYVH