UDC 519.179.1:510.51

Vectorial Problem on Combinations on Hypergraphs

Elvira R. Zaripova[*], **Soltan I. Salpagarov**[†], **Leonid M. Goncharov**[†], **Alexander S. Pankratov**[†]

[*] *Department of Applied Probability and Informatics*
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya st., Moscow, 117198, Russian Federation

[†] *Department of Information Technologies*
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation

Email: zaripova_er@rudn.university, salpagarov_si@rudn.university, magicmight2008@gmail.com, pankratov_as@rudn.university

Nowadays there is a very limited list of tasks using the theory of hypergraphs as in the single-criterion and multi-criteria performances. However, the apparatus of the theory of hypergraphs allows to reflect in the model the simultaneous interaction of several factors.This is possible because the edge of the hypergraph is represented as a nonempty set of vertices that contains more than two vertices.

In this paper a mathematical theory of the vector problem of combinations on l-homogeneous l-partite hypergraphs is formulated. An upper bound is obtained for the cardinality of the set of admissible solutions both for problems with equipotential parts and for problems with parts of different cardinality. An algorithm is constructed to solve this problem. The first part of the presented algorithm for finding combinations on l-partite l-homogeneous hypergraphs is based on the classical algorithm for finding matchings in a bipartite graph. The determination of admissible solutions of the problem is performed without taking into account the values of the weights of the edges of a given hypergraph. In the second part of the algorithm, from the admissible solutions, a set of Pareto optimal solutions of the vectorial problem is constructed. So there is no need to specify in the model building process pairwise interaction modeling. This greatly simplifies the modeling process.

It should be noted, that the technology for solving such problems are also poorly understood. Technology solution of the problem is understood clearly describes the system actions that are performed when dealing with it. The main part of this technology is algorithmic. This is due to the fact that even the optimization problem on graphs related to the problems of the class NP. For mass tasks on graphs in an optimization (single-criterion) formulation known to a limited number of both exact and approximate algorithms.

Homogeneity refers to the edges of a hypergraph, where each edge has exactly l vertices. When building the model this means that there is a need for description l of objects that interact simultaneously. In this case the binary relationship is not necessary to describe. We must specify a common property that these l objects have. The concept of l-partite should be attributed to the set of vertices that are divided into l of different classes. When constructing the model, the partite means that l sets are considered that l sets that have different meaning. Thus, each edge contains one vertex from each of these parts.

We obtained an upper bound for the cardinality of the set of admissible solutions both for problems with equal powers of parts, and for problems with parts of different power of parts. Admissible solutions are formed after the construction of admissible edges on the hypergraph. An edge is valid if the pairs of vertices have taken the meaning the purpose for which you built the model. For a problem with equal powers of parts, we have the problem of perfect combinations. For a problem with the same powers of parts, we have the problem of combinations. Theoretically, there are problems of covering a multi-part hypergraph with various connectivity components, such as stars or trees. In these problems, tasks with different powers of parts are used.

Key words and phrases: hypergraph, multicriteria, vector incomparability, Pareto optimal, recursive algorithm.

In: K. E. Samouylov, L. A. Sevastianov, D. S. Kulyabov (eds.): Selected Papers of the VIII Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems", Moscow, Russia, 20-Apr-2018, published at http://ceur-ws.org

1. Introduction

Currently, an increasing need for the development and study of methods of modeling of technical, economic and social processes associated with direct human involvement. These research methods often refer to discrete integer optimization. Discrete optimization finds at present a wide application in various studies. A distinctive feature of processes involving human beings is their weak structuredness and indistinctness in the values of the initial data. Topical issues of modern discrete modeling are the development of methods of structuring meaningful descriptions of discrete problems and formalization of their parameters in conditions of uncertainty [1]. Thus, topical issues of modern discrete modeling are the development of methods for structuring meaningful descriptions of discrete problems and formalizing their parameters under uncertainty. In describing these methods on the basis of graph theory, we understand that the model in its essence will reflect the binary relations between the two factors of the problem. However, in more complex systems, there are factors that are included in the binary relations and qualitatively affect them. In such cases, problems arise that form relationships, where factors reflecting a single interaction, more than two. Unlike the graph [2, 3], we consider the performances of the vector (multicriteria) problems greatly simplify the modeling process of real objects and processes [3]. For example, the classical problem of allocation in which is the search for a minimum (maximum) sum of weights of edges in a weighted bipartite graph. This task can be generalized and represent the task of searching for combinations on a multi-partite hypergraph. In the classical setting the problem is formulated in the optimization or single-criterion formulation. The problem of combinations on hypergraphs is formulated in a multicriteria or vector formulation. The advantage of using the theory of hypergraphs is that the edges of the hypergraph allow one to describe the relation over an arbitrary set of objects, thereby simplifying the process of modeling those problems in which it is impossible to mathematically define an unambiguous dependence of objects on the basis of binary relations, as in graph theory [4, 5]. It should be noted that this type of problem belongs to the class of NP-difficult problems [6, 7]. For such problems, even in single-criterion formulation has not been investigated, and sometimes do not exist both exact and approximate algorithms for its solution [8, 9].

To solve the formulated problem we propose an exact algorithm for finding the Pareto optimal set of admissible solutions. We obtained an upper bound for the maximum power set of admissible solutions [10].

2. The formulation of the vectorial problem

The missing definitions on the theory of hypergraphs can be found in [2, 3]. A hypergraph $G = (V, E)$ is a pair of sets (V, E). Where V is represented by a finite non-empty set, and E is a family of subsets consisting of the set V. $V = \{v\}$ is the set of vertices of the hypergraph, and $E = \{e\}$ is the set of its edges. Pair of vertices v_1 and v_2 are adjacent if they belong to one edge e, $v_1 \in e$, $v_2 \in e$. The number of vertices in an edge is called the degree of this edge. If the degree of each edge is equal to l, then such a hypergraph is called l-homogeneous. If two edges have common vertices, they are called adjacent. A combination in the graph G is a subset of the set of edges E in which each pair of edges is non-adjacent [4, 5].

A hypergraph $G = (V, E)$ is said to be l-partite if the set of its vertices V is divided into parts (subsets) V_s, $s = \overline{1, l}$ so that two conditions are satisfied:

1. Every pair of vertices of one part is not adjacent.
2. For every edge $e \in E$ every pair of vertices $v_1, v_2 \in e$ belongs to different parts.

If every edge $e \in E$ is incident to one vertex from each parts V_s, $s = \overline{1, l}$, such a hypergraph $G = (V_1, V_2, \ldots, V_l, E)$ is called complete l-partite hypergraph. If to each edge $e \in E$ of a hypergraph G a sequence of numbers $w_v(e) \geqslant 0, v = 1, 2, \ldots, m$ is associated, then it is called m-weighted. A hypergraph is called m-weighted if each of its edges is m-weighted.

An admissible solutions to this problem is an edge subhypergraph $x = (V, E)$, $E_x \subseteq E$ the hypergraph $G = (V, E)$ in which each connected component is a combinations.

Figure 1 shows 3-partite 3-homogeneous hypergraph, where $|V| = 12$, $|E| = 10$, $V_1 = \{1, 2, 3, 4\}$, $V_2 = \{5, 6, 7, 8\}$, $V_3 = \{9, 10, 11, 12\}$, $E = \{e_1, \ldots, e_{10}\}$, $e_1 = \{1, 5, 9\}$, $e_2 = \{1, 6, 10\}$, $e_3 = \{1, 7, 11\}$, $e_4 = \{2, 5, 9\}$, $e_5 = \{2, 6, 10\}$, $e_6 = \{2, 7, 11\}$, $e_7 = \{3, 7, 11\}$, $e_8 = \{3, 8, 12\}$, $e_9 = \{4, 7, 11\}$, $e_{10} = \{4, 8, 12\}$.

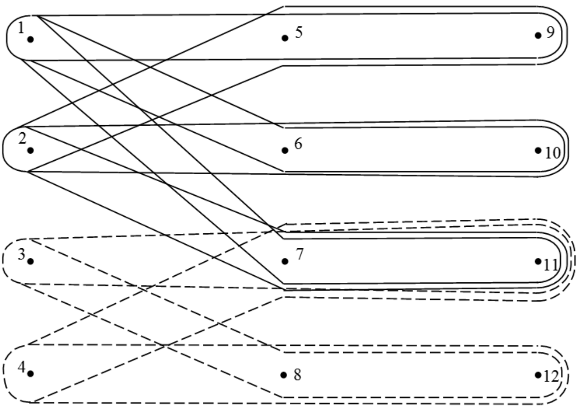

Figure 1. 12-vertebral 3-partite 3-homogeneous hypergraph

We denote by $X = X(G) = x$ the (SAS) set of admissible solutions of this problem. on the hypergraph $G = (V, E)$ in Figure 1 $|X| = 4$, $X = \{x_1, x_2, x_3, x_4\}$, $x_1 = \{e_1, e_5, e_7, e_{10}\}$, $x_2 = \{e_2, e_4, e_7, e_{10}\}$, $x_3 = \{e_1, e_5, e_8, r_9\}$, $x_4 = \{e_2, e_4, e_8, r_9\}$.

The solution of multicriteria problems leads to the fact that there is no single optimal solution. In this case, it becomes necessary to search for a set of alternatives instead of the optimum [9, 10].

The quality of the admissible solutions $x \in X$ is estimated by (VOF) a vector objective function [12, 13].

$$F(x) = (F_1(x), F_2(x), \ldots, F_m(x)), \tag{1}$$

the criteria of which have the form MAXSUM:

$$F_s(x) = \sum_{e \in E_x} w_s(e) \to \max, \quad s = 1, 2, \ldots, m. \tag{2}$$

The definition of these criteria uses weights $w_s(e), s = 1, 2, \ldots, m$ assigned to the edge $e \in E$. VOF (1) whose criteria have the form (2) determines the problem on the SAS, the Pareto set X', consisting of Pareto optima x'. In the case if the equal in value VOF solutions $x', x" \in X$ are considered equivalent (indistinguishable), then from X' a (CSA) complete set of alternatives is introduced X^0. Subset $X^0 \subset X'$ is called CSA if it's have minimal power $|X^0|$ and the equality is true $F(X^0) = F(X')$, where

$F(X^0) = \{F(x) : x \in X^0\}$, $\forall X^0 \subseteq X$ and $F(X') = \{F(x) : x \in X'\}$, $\forall X' \subseteq X$. In this way CSA X^0 is the maximum system of vector-incomparable Pareto optima from X', $X^0 \subseteq X'$. The most expedient solution is selected from the CSA using the procedures of the theory of choice and decision making [14, 15]. As the desired solution of the m-criteria problem under consideration, we take a complete set of alternatives [16, 17].

3. Description of the recursive algorithm

To solve the formulated problem, we describe the algorithm. We consider the problem when the hypergraph $G = (V_1, V_2, \ldots, V_l, E)$ is l-partite l-homogeneous. In the hypergraph G powers of the parts $|V_s|$, $s = 1, \ldots, l$ are different. We introduce the following notation:

- $X(G)$ is a required set of Pareto optimal solutions;
- $W(X)$ is a set of weights (in the form of vectors $(w_1, \ldots, w_m)$ uniquely matched to the set of solutions $X(G)$;
- x_k is a admissible solution of k;
- R is a set of non-adjacent edges complemented by the search for combinations;
- T is a subset edges of E hypergraph $G = (V, E)$, each of which is non-adjacent with any edge in R;
- T' is a subset of edges from T, $T' \subseteq T \setminus \{t_j\} :\ \forall (t_j, t'_k)$, $t'_k \in T'$, $t'_k \bigcap t_j = \oslash$;
- $w_x(x_k)$ is a vector of values of criteria (weights) k combination;
- $w_e(e_k)$ is a vector of values of criteria (weights) k edge;
- $V_{\min}$ is the smallest part of the hypergraph.

The initial data of the algorithm are: l-partite l-homogeneous m-weighted hypergraph $G = (V, E) = (V_1, \ldots, V_l, E)$, $|V_{\min}|$ is a power of the smallest part of the hypergraph G, $|V_{\min}| = \min\{|V_1|, |V_2|, \ldots, |V_l|\}$.

The result of the algorithm is represented by a pair of sets $(X', W(X))$. These sets are uniquely matched. X' is a set of Pareto optimal solutions. The admissible solution is a combination that contains all the vertices of part $V_{\min}$.

The algorithm is divided into two parts. The first part of the algorithm finds SAS without considering the weights of edges. The second part, based on weights and evaluation criteria, allocates a subset of the Pareto optimal solutions from SAS. The algorithm uses intermediary sets T and R. The construction of each admissible solution is based on a recursive selection from a given set of edges E of the subset R. The subset R is complemented by an edge from $T \subseteq E$ at each level of recursion such that it is non-adjacent with any edge in R. The recursion terminates if R becomes $|V_{\min}|$. When $R < |V_{\min}|$ recursive calls continue. In this case, instead of T the nest step of recursion is transmitted to T', the subset of edges in T such that $T' \subseteq T \setminus \{t_j\} :\ \forall (t_j, t'_k)$, $t'_k \in T'$, $t'_k \bigcap t_j = \oslash$. If at the considered step $T = \oslash$, then the construction of the current admissible solution is completed.

An estimate of the maximum cardinality of the SAS for the general case (when the powers of the hypergraph parts $|V_k| > 1$, $k = 1, \ldots, l$).

$$|X(G)| \leqslant \prod_{j=0}^{p} \prod_{k=1}^{l} (|V_k| - j), \quad p = |V_{\min}| - 1.$$

For the special case, when given l-partite l-homogeneous hypergraph $G = (V, E) = (V_1, \ldots, V_l, E)$ is complete and $|V_1| = |V_2| = \ldots = |V_l|$ the estimate represents:

$$|X(G)| \leqslant \left(\left(\frac{|V|}{l}\right)!\right)^{l-1}.$$

When constructing a model, the completeness condition means that there were no restrictions when constructing the edges.

Finding the Pareto optimal solutions among the admissible solutions is constructed according to the following principle. The quality of the admissible solutions $x \in X$ is estimated by the VOF function (1) whose criteria have the form (2). It is assumed that the qualitative characteristics of the solutions found (obtained according to a given law from the set of edges included in them) are represented by numbers and normalized. On the weight vector $(w_1, \ldots, w_m)$, each ordered pair of solutions can be in four states relative to each other. To uniquely determine this relation, it is enough to know the number of characteristics of the first solution, qualitatively superior to the corresponding characteristics of the second, and the number of characteristics equivalent to the corresponding characteristics of the second. In this case, the algorithm for determining the relationship between solutions will have the following form:

1. If the number of equivalent characteristics coincides with the common set of characteristics (m), then these solutions are equivalent.
2. Otherwise, if the number of equivalent and superior characteristics of the first solution coincides with the common set of characteristics (m), then the first solution qualitatively exceeds the second one.
3. Otherwise, if the number of superior characteristics of the first solution is zero, then the second solution qualitatively exceeds the first.
4. Otherwise the pair of solutions is considered vectorally incomparable.

Figure 2 shows the function that performs comparisons solutions x.

```
function Comparison( x ):
   W_x(x_z) := Σ_{e_i∈x_z} w_e(e_i)
   if X(G) = {∅} and W(X) = {∅} then
      X(G) := X(G) ∪ {x_z}, W(X) := W(X) ∪ {w_x(x_z)} return true
   end if f_c := true
   for all x_i ∈ X(G) do:
      a, b := 0
      for k := 1,m do:
         if w_x,k(x_z) = w_x,k(x_i) then b := b+1 continue
         end if
         if w_x,k(x_z) > w_x,k(x_i) then a := a+1 continue
         end if
      end do
      if b = m then
         f_c := true, X(G) := X(G) ∪ {x_z}, W(X) := W(X) ∪ {w_x(x_z)}
         return true
      elif a+b = m then
         f_c := true, X(G) := X(G) \ {x_z}, W(X) := W(X) \ {w_x(x_z)}
         continue
      elif a = 0 then f_c := false, return false
      else f_c := true, continue
      end if
   end do
   if f_c = true then
      X(G) := X(G) ∪ {x_z}, W(X) := W(X) ∪ {w_x(x_z)} return true
   else return false
   end if
end function
```

Figure 2. Function that performs comparisons solutions x

From the point of view of software implementation, the finding of admissible solutions, their evaluation and subsequent allocation, a lot of Pareto optimal solutions occurs as follows. At the beginning of the algorithm, a dynamic list of solution instances is initialized, which will serve as the resultant Pareto optimal set. Alternately, each edge from the given source list of edges is selected by a set of edges that do not have common vertices and are listed in the list after this. For each list that is generated, the algorithm is repeated until there remain lists of vertices that satisfy the conditions of an admissible solution. When finding each new feasible solution, it is passed to the method of assessing the quality of decisions. The solution is compared with each element of the given Pareto optimal set by the algorithm given above.

Variants of the behavior of the method of formation of the set of Pareto optimal solutions:

1. The first solution found is entered in the (PS) Pareto set without checking, respectively.
2. If the solution found qualitatively inferior to all solutions from the PS, it is ignored.
3. If the solution is equivalent to one or more solutions from the PS, it is added to the PS as an alternative.
4. If the solution exceeds one or more solutions from the PS, it is added to the PS, and all qualitatively inferior ones are excluded from it.
5. If the solution is vectorally incomparable to none of the solutions in the PS, it is added to it.

As a result of the verification of all found admissible solutions, we obtain a set of Pareto optimal admissible solutions of the problem that interests us. Figure 1 shows recursive algorithm for finding solutions to the vector combination problem [18, 19].

4. Conclusions

As a result of the work the following results were obtained: the mathematical vectorial problem of combinations on l -homogeneous l-partite hypergraphs is formulated, the upper bound for the cardinality of the SAS for problems with equal and different powers of parts is obtained.

The results obtained in the work can be used to develop a decision support system. They can be used in the process of modeling problems in different subject areas under multicriterion conditions, the upper bound for the cardinality of the set of admissible solutions both for problems with equal powers of parts and for problems with of different cardinality of fractions is obtained.

An example of this approach can be found in [20]. In the future, it is planned to program the problem of covering the hypergraph with stars.

References

1. R. T. Rockafellar, R. J.-B. Wets, Generalized liner-quadratic problems of deterministic optimal control in discrete time, SIAM Journal Control and Optimization **28** (4) (1990) 810–822. doi:10.1137/0328046
2. C. Berge, Graphs and Hypergraphs, North-Holland, 1973 528 p.
3. C. Berge, Hypergraphs: combinatorics of finite sets. North-Holland, 1989 267 p.
4. K. Thulasiraman and M. Swamy, Graphs: Theory and Algorithms, John Wiley & Sons, New York, 1992 480 p.
5. A. Bretto, Hypergraph Theory, Mathematical Engineering, Springer, Heidelberg, 2013, 136 p. doi:10.1007/978-3-319-00080-0
6. I. Holyer, The NP-completeness of some edge-partition problems, SIAM Journal on Computing **10** (4) (1981) 713–717. doi:10.1137/0210054
7. M. R. Garey, D. S. Johnson and L. Stockmeyer, Some simplified NP-complete graph problems, Theoretical Computer Science, **1** (3) (1976) 237–267. doi:10.1016/0304-3975(76)90059-1

8. M. R. Garey, D. S. Johnson, Computers and Intractability: A Guide to NP-Completeness, W. H. Freeman & Co. NY USA, 1990, 338 p.
9. P. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on quantum computer, SIAM Journal on Computing **26** (5) (1997) 1481–1509. doi:10.1137/S0097539795293172
10. S. Chung, H. Hamacher, F. Maffioli and K. Murty, Note on combinatorial optimization with max-linear objective functions, Discrete Applied Mathematics **42** (2–3) (1993) 139–145. doi:10.1016/0166-218X(93)90043-N
11. D. Luc, Theory of vector optimization, volume 319 of Lecture Notes in Economics and Mathematical Systems, Springer Verlag, Berlin, 1989, 173 p. doi:10.1007/978-3-642-50280-4
12. A. Warburton, Quasiconcave vector maximization: Connectedness of the sets of Pareto-optimal and weak Pareto-optimal alternatives, Journal of Optimization Theory and Applications, **40** (4) (1983) 537–557. doi:10.1007/BF00933970
13. A. Warburton, Approximation of Pareto optima in multiple-objective shortest-path problems. Operations Research **35** (1) (1987) 70–79. https://doi.org/10.1287/opre.35.1.70 doi:10.1287/opre.35.1.70
14. Y. Sawaragi, H. Nakayama and T. Tanino, Theory of Multiobjective Optimization. Academic Press, Orlando, FL., 1985, 311 p.
15. A. Pascolett and P. Serafini, Scalarizing vector optimization problems, Journal of Optimization Theory and Applications **42** (4) (1984) 499–524. doi:10.1007/BF00934564
16. V. A. Emelichev, V. A. Perepelica, Complexity of discrete multicriterial problems, Discrete Mathematics and Applications **4** (2) (1994) 89–117. doi:10.1515/dma.1994.4.2.89
17. C. H. Papadimitriou and K. Steglitz, Combinatorial optimization: Algorithms and Complexity, Prentice Hall, New York, 1982, 496 p.
18. D. M. Moyles and G. L. Thompson, An algorithm for finding the minimum equivalent graph of a digraph, Amer. Math. Soc., **16** (3) (1969) 455–460. doi:10.1145/321526.321534
19. R. Z. Norman and M. O. Rabin, An algorithm for a minimum cover of a graph, Proc. Amer. Math. Soc. **10** (1959) 315–319. doi:10.1090/S0002-9939-1959-0106853-5
20. S. I. Salpagarov, Yu. D. Isaev, An optimization model of distribution of P2P-TV data streams on hypergraphs, CEUR Workshop Proceedings **2064** (2017) 130–135.