Contributions to Game Theory and Management, XV, 178–188

Algorithm of Hierarchical Matrix Clusterization and Its Applications

Elena A. Lezhnina and Elizaveta A. Kalinina

St. Petersburg State University,
Faculty of Applied Mathematics and Control Processes,
7/9, Universitetskaya nab., St. Petersburg, 199034, Russia
E-mail: e.lezhnina@spbu.ru; e.kalinina@spbu.ru

Abstract In this article, the problem of hierarchial matrix clusterization is discussed. For this, the influence of individuals on the community was used. The problem of dividing the community into groups of related participants has been solved, an appropriate algorithm for finding the most influential community agents has been proposed. Clustering was carried out using an algorithm for reducing the adjacency matrix of a directed graph with nodes representing members of a social network and edges representing relationships between them. The applications to the problems of working groups, advertising in social networks and complex technical systems are considered.

Keywords: hierarchial matrix clusterization, influence of agents, working groups, advertising in social networks.

1. Introduction

The article presents an algorithm that solves the problem of recognizing the hierarchical structures in big groups, and also discusses the possibilities of its application. We offer an overview of the models in which the proposed algorithm is applied. The first model solves the problem of dividing a large team into small working groups. The internal connections of groups are represented by an adjacency matrix. Next, the clustering algorithm is used for the task of dividing the work team into effective small groups. The effectiveness of groups is calculated based on the characteristic function of the group, which takes into account interpersonal relationships. In the second model, communities in social networks are analyzed and the problem of advertising distribution within groups is discussed. Practical and numerical experiments were carried out to verify the adequacy of the models.

Hierarchical organizations include all systems where there is control over lower levels. Hierarchical data structures and hierarchical management structures are used in databases, management theory, decision theory, organization theory, big data, sociology and psychology, social network analysis. The hierarchical system makes it possible to simplify the management of complex systems (Li and Mavris, 2012; Mitrishkin et al., 2009; Chmúrny, D. and R. Chmúrny, 1996). In the theory of decision making, Thomas L. Saati in the 1970s developed the analytic hierarchy process (AHP), a structured technique for organizing and analyzing complex decisions (Saaty, 1980). However, determining the existing hierarchical structures on a set of objects is often not an easy task. The existing methods of hierarchical clustering solve the problem of taxonomy, combining objects into clusters according to a certain set of features (Vijaya and Ritika, 2017). In (Parygin et al., 2017), the task

https://doi.org/10.21638/11701/spbu31.2022.13

of identifying the hierarchical structure is solved by experts, in several iterations. We propose an algorithm that solves the problem of analyzing relationships between objects.

2. Algorithm of Hierarchical Clusterization

Let there be a set $\mathcal{N}$. The elements of a set are nodes of a complex system or members of a large team. The connections between the elements of the set can be one-sided, two-sided, or absent. Thus, the set $\mathcal{N}$ can be represented as a directed graph, the edges of which indicate the influence of elements on each other. For a group of $\mathcal{G}$ elements, we make an adjacency matrix $A = [a_{ij}]_{n\times n}, n = |\mathcal{G}|$ which reflects the influence of the elements of the set on each other. This matrix is used to divide a large group into clusters according to the following algorithm.

Let $A = [a_{ij}]_{n\times n}$ be a nonnegative matrix. For $n > 1$ matrix A is reducible if for some permutation matrix P

$$\tilde{A} = PAP^{\top} = \begin{bmatrix} B & C \\ \mathbb{O} & D \end{bmatrix}$$

where B and D are square matrices and the superscript "$\top$" means "transpose". Otherwise, A is *irreducible*. Irreducibility can be tested using the associated directed graph. A directed graph G is a set of n nodes and a set of directed edges joining two nodes. A directed graph is strongly connected if any two distinct nodes are joined by an oriented path. It is unilaterally connected if it contains a directed path from j to k or a directed path from k to j for every pair of nodes j, k. A directed graph is weakly connected if the undirected underlying graph obtained by replacing all directed edges of the graph with undirected edges is a connected graph. The directed graph $G(A)$ associated with an $n \times n$ nonnegative matrix A (the adjacency matrix) is a directed graph of n nodes, such that there exists an edge (j, k) leading from node j to node k if and only if $a_{kj} > 0$. We have the following property: a matrix A is irreducible iff $G(A)$ is strongly connected (Guo, Li and Shuai, 2006). It is known that a graph can be considered as a binary relation on the set of its nodes. The transitive closure of a binary relation R on a set X is the smallest relation on X that contains R and is transitive. By constructing a transitive closure of the graph, we can answer the reachability questions. Let T be the matrix of a transitive graph, and $T = PTP^{\top}$ is its canonical form (Savitskaya, 2008). Then there are the following cases:

1) T is a matrix of units. Then the graph $G(A)$ is strongly connected, and the matrix A is irreducible;

2) T is a block upper triangular matrix without any zeros, symmetric with respect to the main diagonal. Then the matrix A is weakly reducible, and the graph $G(A)$ is one-way connected;

3) T is a block upper triangular matrix with some zero elements symmetric with respect to the main diagonal. Then the matrix A is weakly reducible, and the graph $G(A)$ is weakly connected and has at least a pair of nodes inaccessible to each other;

4) none of the above is true. This means that the graph $G(A)$ is disconnected.

To find the transitive closure of graph $G(A)$, we will use the Floyd–Warshall algorithm. First, let the entry $t_{ij} = 1$ if the nodes i and j of the graph $G(A)$ are connected, and otherwise $t_{ij} = 0$. Then we construct extra edges by the following formula: $t_{ij} = t_{ij} \vee (t_{ik} \wedge t_{kj})$ where $\vee$ and $\wedge$ stand correspondingly for logical operations of conjunction and disjunction.

Now we have the following problem. Given a $(0, 1)$-matrix T that is generally non-symmetric (this is the matrix of the transitive closure of $G(A)$), find its canonical form. We will rely on the algorithm presented in Savitskaya, 2008.

Algorithm 1

For a given nonnegative matrix A, we want to find out whether it is irreducible.

Step 1. Construct the matrix T of the transitive closure of the graph $G(A)$ associated with A.

Step 2. If T is the matrix of ones, then A is irreducible. Otherwise, go to issue **3**.

Step 3. Compute the column S of row sums of T, then construct a permutation matrix P such that $PS = \tilde{S}$ where $\tilde{S}$ is the column with entries sorted in descending order. The column $\tilde{S}$ contains k groups of equal entries with n_j $(j = 1, 2, \ldots, k)$ entries in j-th group, $\sum_{j=1}^{k} n_j = n = \dim(A)$. Hence, we have the matrix

$$\mathcal{T} = PTP^\top = \begin{bmatrix} T_1 & * & \ldots & * \\ 0 & T_2 & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & T_k \end{bmatrix}$$

where T_j is a square matrix of size n_j.

Step 4. In the matrix $\mathcal{T}$, find zeroes symmetric w.r.t. its main diagonal. If $\mathcal{T}$ does not have such zeroes, then the matrix A is associated with unilaterally connected graph $G(A)$. Otherwise, the graph $G(A)$ is weakly connected, and we get the pairs of nodes that are not connected.

Step 5. Find the matrix $\tilde{A} = PAP^\top$.

As a result we have m clusters with n_i $(i = 1, 2, \ldots, m)$ elements in each. Let's denote clusters as $C_1, C_2, \ldots, C_m$. The elements of the first cluster are connected with the participants of the subsequent clusters - the second, third, etc. The elements of the second cluster are related to the elements of the third, fourth, etc., but not the first. And in general elements from , cluster C_k are connected with elements of clusters $C_{k+1}, C_{k+2}, \ldots, C_m$ and do not affect elements of cluster $C_{k-1}, C_{k-2}, \ldots, C_1$. The proposed algorithm can be used to identify hierarchical structures in complex networks.

To illustrate the algorithm, consider a group of elements. The graph in Figure 1 shows a diagram of the mutual influence of the group elements on each other. Despite the small size of the group, it is quite difficult to determine the hierarchy of relations here.

Algorithm 1 gives the following partition:

$$\{3, 4\} \to \{1, 2, 11\} \to \{5, 6, 7, 8, 10, 12\} \to \{9\}. \tag{1}$$

There are programs and applications that graphically illustrate the connections within groups. For small groups of objects, they can help analyze hierarchical relationships within a group. For example, the service edotor.net for a group with Figure 1 gives the tree presented in Figure 2.

The presented tree still allows you to select a group of elements $\{3, 4\}$ as the most influential. However, when increasing the group to 30 elements with internal

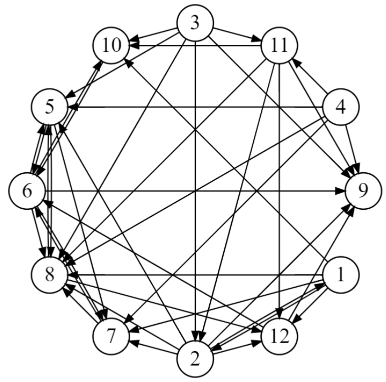

Fig. 1. The diagram of the mutual influence of the group of 12 elements

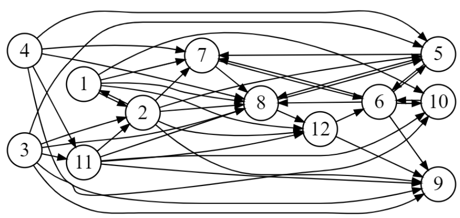

Fig. 2. The tree of the mutual influence of the group in Figure 1

connections, represented in the graph in Figure 3, edotor.net gives the tree represented in Figure 4. In this group, the most influential element is $\{11\}$, but it is difficult to do this by visual analysis of the tree.

Interesting applications of this algorithm appear depending on the semantic load of the connections between the two nodes of the graph. For example, one can interpret the relationship as the influence of one element on another and determine the characteristic function of a coalition of two elements. The methods of forming a characteristic function can be different, depending on the problem being solved.

3. Applications

3.1. Model 1. Effective working groups

To demonstrate the work of the presented algorithm, we used it to solve the problem of splitting a large team into effective small working groups.

One of the most popular and effective approaches to organizing activities is the organizing of temporary working groups. An important indicator is the effectiveness of such groups, which strongly depends on the internal structure of the working group and the relationship between team members (Moreno, 1941; Campion et al., 1993, Gutiérrez et al., 2016). In (Gubanov et al., 2011) is shown that the most optimal management in the working group has a hierarchical structure. In order to quickly and conflict-free build such an "ideal" hierarchical structure, it is necessary to carefully form the composition of working groups. We will interpret the connection between two elements as the power of influence of one element on the other in the sense of whether one element (a member of the group) can be a leader for the other. We introduce the characteristic function $f(i,j)$. This function is not

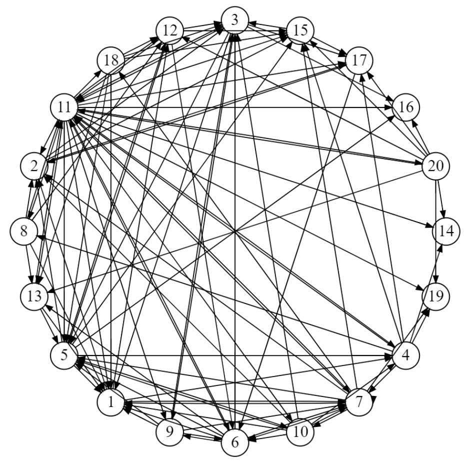

Fig. 3. The diagram of the mutual influence of the group of 20 elements

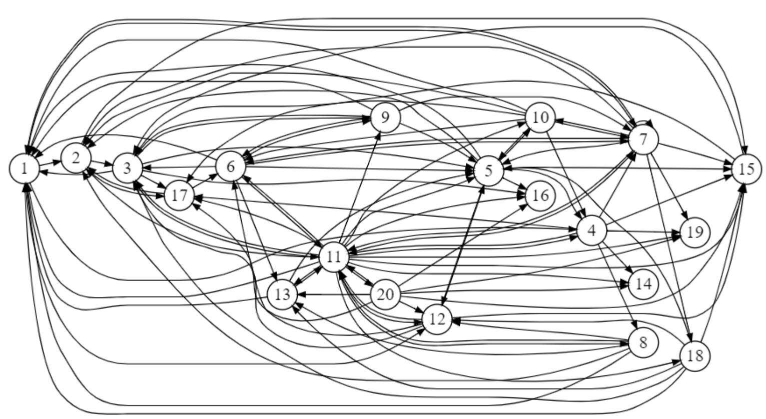

Fig. 4. The tree of the mutual influence of the group in Figure 3

symmetrical, and its asymmetry makes it possible to determine the "direction" of influence. Such a function can also be interpreted as an assessment of the effectiveness of the pair. You can determine the values of this function, for example, using questionnaires compiled by psychologists. To calculate the efficiency of a group S of several participants, we define the effectiveness function as follows:

$$F(S) = \sum_{\{i,j\} \subset S, f(i,j) > f(j,i)} f(i,j).$$

The algorithm for forming working groups. Let there be a group of participants $\mathcal{G}$ with n people. Let's assume that a large group is already divided into influence groups. To simplify the description, we will call them clusters. We have m clusters, each with n_i people. Let's denote them $C_1, C_2, \ldots, C_m$. The members of the first cluster influence (can lead) the participants of the subsequent clusters — of the second, third, etc. The members of the second cluster influence the members of the third, fourth, etc., but not the first. And in general, cluster members C_r affect cluster members $C_{r+1}, C_{r+2}, \ldots, C_m$ and do not affect cluster members $C_{r-1}, C_{r-2}, \ldots, C_1$. It is necessary to make q collectives with maximum efficiency. Let's assume for simplicity that the teams should have the same number of participants and the number of clusters coincides with the required number of participants in workgroups. In this case, we take one person from each cluster. The important thing here is to maximize the effectiveness of the group. That is, we are looking for such a combination of group members that gives the maximum of the effectiveness function $F(S)$. The task of forming such groups can be solved in various ways. If the initial group $\mathcal{G}$ consists of a small number of participants, then you can stop at a simple search. For large groups, we propose the following algorithm.

Algorithm 2

Input: Matrix $\tilde{A}$ constructed by **Algorithm 1**
Result: Division into working groups
$i = 1$
Step 1. If $i \leq m$, **then go to Step 2 otherwise** output group members
Step 2. From cluster C_2, take the element j which gives $max_{j \in C_2} f(i,j)$.
Step 3. From cluster C_3, take the element k which gives $max_{j \in C_2} F(\{i,j,k\})$.
Step 4. Etc. next across all clusters.
$i = i + 1$
Go to Step 1.

This algorithm is used when the number of participants in clusters is the same and coincides with the required number of groups. In the case when the number of participants in each cluster is different, it is necessary to reformulate clusters to new m clusters as follows: 1) If the number elements of the cluster is less than m, then we combine this cluster with the next one. 2) If there are more elements in the first cluster of "leaders" than m, we leave in the cluster C_1 those elements i that give first m maximal values of index $I_i^1 = \sum_{j \in C_2} f(i,j)$. If there are more than m elements in a cluster with a number greater than 1, then we leave in the cluster those elements j that give first m maximal values $I_j^2 = \sum_{i \in C_1} f(j,i)$ and transfer the rest participants to subsequent clusters. In cases where an unequal number of participants is required in groups, or a certain level of efficiency is required, the algorithm changes accordingly.

Numerical experiments. Practical experiments were conducted to demonstrate the proposed algorithms. Three groups of university students attending geometry classes took part in the experiments. The first experiment involved 35 students from two different groups. The internal connections in this combined group are heterogeneous. Some students have never communicated with each other, some have already had experience working together on projects. As a result of applying the algorithms, the students were divided into 7 groups with presumably the same efficiency. Then the students were asked to solve five problems of increased complexity in groups. The authors of the article observed the interactions of students within groups. After working together, the students named the team leaders. In six groups, the true leaders matched those predicted by the algorithm. In one group, the leader's place was taken by a member of the group from the second cluster. All groups demonstrated uniform efficiency. In the second experiment, the third group of 20 students was divided into five working groups of 4 students. All students agreed with the proposed division into groups and the appointment of group leaders. To study the effectiveness of the developed algorithms, 50 relationship matrices were generated for teams of 200 people. A greedy algorithm, an algorithm that selects commands in an arbitrary way, and a new algorithm combining algorithms 1 and 2 of this article were applied. According to the values of the effectiveness of the groups, the new algorithm showed the best results. In terms of speed, the new algorithm loses. However, it should be noted that the operation of the new algorithm in these experiments did not exceed several minutes when implemented on a medium-power personal computer. Considering how often such tasks have to be solved, such a lag is not critical.

Remark 1. The results of our study are published in (Lezhnina et al., 2023).

3.2. Model 2. Distribution of advertising on social networks

Computer social networks occupy an important place in people's lives. Interacting in them, some agents influence the opinions of others, thereby inducing the latter to certain actions. In recent years, advertising has become the main commercial activity on the Internet. Users of a social network are logically grouped according to one or several criteria, such as: friendship, values, interests, ideas, etc. Since these communications are strong enough, buying a certain product, the user can influence his environment, which will lead to an increase sales probabilities. To carry out advertising on a social network, a company, first of all, develops a strategy. It can be a traditional or Internet-oriented platform, targeted or non-targeted. In any case, its main goal is to attract a small group of users to purchase the product. This group will launch a cascading influence reaction, which will lead to a possible increase in the group of users who bought the product as a result. Determining the final number of customers is an important task for an advertising company, because it can develop an effective strategy in order to increase profits. At the same time, the spread of influence depends on many factors that are difficult to formalize. Due to the high level of social influence in the networks, even a small group of buyers can attract the attention of a large number of people to this product. A market based on the principle of "word of mouth" can be much more effective than traditional marketing schemes, since users independently carry out advertising, influencing each other.

Model. Agents within the network influence each other. The degrees of influence of the agents will be specified by an matrix $A = [a_{ij}]_{n\times n}$ with entries a_{ij}, where $a_{ij} > 0$ will denote the strength of the influence of the i-th agent on the j-th agent (or the degree of trust of the j-th agent to the i-th agent). The magnitude of the influence of agents depends on many factors. These factors can be friendships, common interests, etc. In drawing up the model, we will use factors, by which we can determine the degree of influence of agents. We introduce the factors that determine the influence of some agents on others: $x_{ij}^{\ell} \in [0,1]$, $\ell = 1, 2, \ldots, 8$.

Here x_{ij}^1 is the number of messages between the i-th and j-th agents: the more people communicate with each other, the more they influence each other; x_{ij}^2 is the number of dialogs of the i-th and j-th agents, shows how often people communicate with each other; x_{ij}^3 stands for the number of common groups of the i-th and j-th agents. A large number of common groups speaks of the common interests of agents; x_{ij}^4 is the number of common audio recordings of the i-th and j-th agents. As well as general groups, it speaks of common interests and individually does not play a big role, but, in combination with the first two factors, in some cases has a rather large weight; x_{ij}^5 are mutual friends of the i-th and j-th agents; x_{ij}^6 — status of the i-th agent for the j-th (best friends, family ties) (if indicated). This factor can directly indicate the degree of trust between agents; x_{ij}^7 — the number of subscribers of the agent. With a large number of subscriptions, it is likely that the agent is a well-known person and, accordingly, has an increased impact on his subscribers; $x_{ij}^8 \in \{0,1\}$ — use of innovation by the i-th agent (1 — used, 0 — not used). Usually, those who listen to the opinion of a person who already uses the product more than to a person who has only heard about this innovation, therefore this factor is of great importance. Each factor has its own weight in determining the degree of influence between agents. Moreover, this weight in different cases can be determined in different ways. Therefore, to correctly determine the degree of influence between agents, it is necessary to introduce a vector of coefficients that would help in the correct proportions to establish the dependence of the degrees of influence between agents on the above factors. We introduce this vector α of coefficients with components $\alpha_\ell \in [0;1]; \ell = 1, 2, \ldots, 8$. Each component of the coefficient vector corresponds to the components of the factor vector that determine the effect. Moreover, the components of the coefficient vector are subjective quantities that vary depending on innovations. So, for example, with an innovation related to a certain style of music, the coefficient α_4 corresponding to the factor of general audio recordings will be higher than the same coefficient with the innovation related to technological innovations. Now that the main factors and their weights are determined in determining the degree of influence, we can determine the effects of agents on each other. So, the degree of influence of the i-th agent on the j-th will be directly proportional to the components of the vector of factors and the components of the vector of coefficients corresponding to these factors, and is calculated by the formula: Thus, since the i-th agent is affected by n other agents (their influence may be 0), the total effect of all agents that are directly related to agent j on this agent

j will be calculated by the formula:

$$a_{ij} = \sum_{\ell=1}^{8} \alpha_\ell x_{ij}^\ell.$$

We have considered a connected social group via social media website `vk.com`. Firstly, the 127 friends of one person were analyzed. The connections between them were presented as a graph and the adjacency matrix A of this graph was created. Following Algorithm 1, the matrix of the transitive closure T was calculated, using appropriate modification of the Floyd–Warshall algorithm. Since the matrix T contains zeros, matrix A is reducible, and this is confirmed to be true by the **Step 5** of the Algorithm 2. The vector $\tilde{S}$ is the following:

$$\tilde{S} = [105, \ldots, 105, 2, 2, 0, \ldots, 0]^\top,$$

where the cardinalities of each group are 105 for $n_1 = 105$, 2 for $n_2 = 2$ and 20 for $n_3 = 0$.

As a next step, the permutation matrix P was found, and the matrix $\mathcal{T}$ was calculated. We found that it has 2540 zeroes, symmetrical w.r.t. main diagonal, which means, that the graph has to be weakly connected. Since listing all of the not connected pairs of vertices would be problematic due to the quantity, we present here all such pairs for one node. We have that the node 1 is disconnected with nodes 5, 9, 10, 15, 17, 24, 35, 42, 44, 48, 68, 73, 76, 79, 86, 104, 114, 119, 120, 123, 124, 125. Now we can offer product samples to some group of users who have a lot of influence within the group, since the best way is to identify a group of the most influential agents in order to provide them with discounts, benefits or special promotions. That will contribute to the greatest distribution of goods, services or innovations throughout the social network. To determine the proportion of customers, we use the q-influence model presented in Zhao et al., 2009. According to the q-influence model, node i will not acquire the goods if and only if it does not acquire it at will and none of its neighbors who bought the goods have any influence on it. The proportion of customers is determined from the equation

$$1 - E[Y] = (1-\mu)\sum_{k=0}^{\infty} P_1(k+1)(1-qE[Y])^k,$$
$$1 - E[X] = (1-\mu)\sum_{k=1}^{\infty} P_0(k)(1-qE[Y])^k.$$

Here X is a random variable that shows whether the node r bought the product as a result Y is a random variable that shows whether the non-root node i will buy the product only under the influence of advertising and its followers. X and Y are distributed according to Bernoulli distribution. φ_i is a random variable having a Bernoulli distribution, since takes the value 1 if the user decided to purchase the product as he wanted (for example, without the influence of other nodes). ($P(\varphi_i = 1) = p^+$), and 0 if the node acquires the goods under the influence of advertising ($P(\varphi_i = 0) = p^-$). As an advertisement, the company sends free samples $\rho < 1$ share of users.

Then the mathematical expectation φ_i has the form:

$$\mu = \rho p^+(1-\rho)p^-,$$

$P_1(k)$ is the probability that the child node has degree k, $P_0(k)$ is the root node has degree k, q is a threshold showing how many neighbors of a node a product must buy in order for a node to buy it. Next, we can calculate the proportion of users to whom the company will send free samples to reach the customer audience $N = E(X)$.

Now it is necessary to determine to whom exactly it is necessary to send samples, that is, to identify the most influential agents. Algorithms exist for identifying such agents, for example, in Gubanov et al., 2011.

Remark 2. The results of our study are published in Lezhnina et al., 2022.

3.3. Complex technical systems

As a rule, complex technical systems are managed hierarchically. If it is necessary to "recognize" the structure of such a system, then the matrix clustering method can also be used. At the experimental stand, the nodes of the system are tested: the nodes are affected in turn, after which the connections are indicated: if changes occur in some nodes when one of the nodes is affected, then their connections are denoted 1, the connection with nodes that do not respond to the impact on the tested node is denoted 0. The resulting matrix is clustered, identifying the main control nodes. A similar approach can be applied to the diagnosis of complex biological systems. In group decision, the task of choosing the best alternative (candidate) out of many is solved. Condorcet method Moulin, 1989 uses a pairwise comparison of all alternatives to solve such problems. In the case of a large number of alternatives, when the option "indifferent" is allowed, this task becomes time-consuming. However, if we interpret the relationship between the elements as "better" or "worse", we can build an adjacency matrix based on the principle: if i better j, then $a_{ij} = 1$, otherwise $a_{ij} = 0$. As a result, we will get one or more elements that are better than all the others. Comparing the multiple elements is not a difficult problem.

4. Conclusion

The paper presents an overview of mathematical models in which the algorithm of hierarchical clustering of sparse matrices is used and discusses the possibilities of its application.

The work of algorithms for the problem of forming effective small working groups and the problem of distributing advertising in social networks is demonstrated. The efficiency of the proposed algorithms was evaluated.

References

Campion, M. A., Medsker, G. J. and Higgs, A. G. (1993). *Relations between workgroup characteristics and effectiveness: implications for designing effective workgroups.* Pers. Psychol., **46(4)**, 823–847.

Chmúrny, D. and Chmúrny, R. (1996). *Hierarchical control of biotechnological complex systems.* Bioprocess Engineering, **15**, 171–179.

Gubanov, D. A., Novikov, D. A. and Chkhartishvili, A. G. (2011). *Informational influence and informational control models in social networks.* Autom. Remote Control, **72**, 1557–1567.

Gutiérrez, J. H., Astudillo, C. A., Ballesteros-Pérez, P., Mora-Meliá, D. and Candia-Véjar, A. (2016). *The multiple team formation problem using sociometry.* Computers & Operations Research, **75**, 150–162.

Guo, H., Li, M. I. and Shuai, Zh. (2006). *Global Stability of the Endemic Equilibrium of Multigroup SIR Epidemic Models.* Can. Appl. Math. Q., **14**, 259–284.

Lezhnina, E. A., Kalinina, E. A., Karpenko, V. A. and Kononov, I. A. (2022). *Strategy for maximizing information influence in social networks.* AIP Conference Proceedings (to appear)

Lezhnina, E. A., Khitrov, G. M., Kalinina, E. A. and Kononov, I. A. (2022). *Working groups: creating and management.* Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science (to appear)

Li, Y. and Mavris, D. (2012). *A hierarchical control architecture for resource allocation.* 12th International Conference on Control Automation Robotics & Vision (ICARCV), pp. 955–959. https://doi.org/10.1109/ICARCV.2012.6485286

Mitrishkin, Y., Haber, G. and Rodolfo, E. (2009). *Intelligent Hierarchical Control System for Complex Processes – Three Levels Control System.* ICINCO 2009 — 6th International Conference on Informatics in Control, Automation and Robotics, Proceedings. **1**, 333–336.

Moreno, J. L. (1941). *Foundations of sociometry: an introduction.* Sociometry, **4(1)**, 15–35.

Moulin, H. (1989). *Axioms of Cooperative Decision Making.* Cambridge University Press: Camridge.

Parygin, D. S., Sadovnikova, N. P., Shabalina, O. A. and Korneev, I. M. (2017). *Formation of an Objectives Hierarchy in Tasks of Organizational Systems Management.* Ontol. proekt. Vol. 7, **4(26)**, 496–509. https://doi.org/10.18287/2223-9537-2017-7-4-496-509 (in Russian)

Saaty, T. L. (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resources Allocation.* Mcgraw-Hill: New York.

Savitskaya, D. V. (2008). *Normal form of (0,1)-matrices and algorithms of its construction.* Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes, **4**, 85–97. https://doi.org/10.1007/BF00369478

Vijaya, A. S. and Ritika, B. (2017). *A Review on Hierarchical Clustering Algorithms.* Journal of Engineering and Applied Sciences, **12**, 7501–7507.

Zhao, B., Li, Y. K., Lui, J. C. S. and Chiu, D.-M. (2009). *Mathematical Modeling of Advertisement and Influence Spread in Social Networks.* In NetEcon: Workshop on the Economics of Networks, Systems and Computation.