DOI: 10.17323/2587-814X.2024.1.52.64

Nonparametric procedure for comparing the performance of divisions of a network organization

Petr A. Koldanov [a]
E-mail: pkoldanov@hse.ru

Vladimir A. Koldanov [b]
E-mail: vlad.kold@gmail.com

[a] Laboratory of Algorithms and Technologies for Network Analysis, HSE University
Address: 136, Rodionova St., Nizhny Novgorod 603093, Russia

[b] Nizhny Novgorod State University of Engineering and Economics
Address: 22A, building 2, Oktyabrskaya St., Nizhny Novgorod Region, Knyaginino 606340, Russia

Abstract

To solve the problem of comparative efficiency analysis of branch operations for a small volume of randomly observed data, a non-parametric approach is relevant, since it does not require a probabilistic model of observations. Comparing the results of the non-parametric approach with the results obtained within the traditionally used Gaussian model is also relevant. Additionally, obtaining a consistent comparison of a group (of no less than three) branches is important. Currently, the non-parametric approach and the corresponding comparison with the known results of solving the problem considered in this work obtained within the framework of the normal model are absent. In addition, insufficient attention is paid to the search for methods of obtaining consistent solutions. This work to some extent fills these gaps. This work uses non-parametric statistical methods and theory of simultaneous hypothesis testing to address these problems. This paper proposes a procedure for comparative analysis of the efficiency of several units within a network organization with a small volume of observations based on the Mann–Whitney tests. We carry out a comparison of the results obtained from the proposed non-parametric procedure with results based on extensions of Student's *t*-tests. We propose a method for reducing the number of compatibility problems based on the search for an appropriate significance level. We provide an example of a fully consistent comparison of the efficiency of branch operations.

Keywords: network organization, efficiency of branch operations, Mann–Whitney tests, incompatibility problem

Citation: Koldanov P.A., Koldanov V.A. (2024) Nonparametric procedure for comparing the performance of divisions of a network organization. *Business Informatics*, vol. 18, no. 1, pp. 52–64.
DOI: 10.17323/2587-814X.2024.1.52.64

Introduction

Various aspects of comparing the effectiveness of organizations are discussed in numerous scientific papers, for example [1–3]. As a rule, comparisons are based on many indicators. Obviously, the success of such comparisons depends on how adequately and qualitatively comparisons on individual indicators can be made, especially if these indicators are of a random nature. In this latter case, the methods of mathematical statistics are typically used. Such methods are divided into:

- parametric, relying on a specific probabilistic model of the analyzed indicators. In this case, the normal distribution is most often used as a probabilistic model [4, 5],
- non-parametric, free from a detailed probabilistic model, and sometimes from the assumption of the random nature of the analyzed data as well [3, 6–9].

Many tasks, including the one this paper is aimed at, can be considered within the framework of both approaches. In this case, it becomes necessary to compare the conclusions based on parametric model procedures to the conclusions of nonparametric procedures. Such a comparison is one of the goals of this work.

There are many known results comparing parametric and nonparametric tests of hypothesis checking against an alternative. With a finite number of observations, such a comparison is made on the basis of the analysis of the test power function determined by the probability of rejecting the hypothesis. With an unlimited number of observations, the comparison of hypothesis checking tests against an alternative is based on the calculation of asymptotic efficiency indicators [10]. The specificity of the problem considered in this work lies in the need to select one of many solutions based on a small number of observations.

This paper proposes a procedure for comparative analysis of the performance of divisions of a network organization. The results of applying such a procedure can be used to make informed management decisions by the managers of a network organization. In this case, the efficiency of a division is understood as the ratio of the number of sales of a certain product (for example, the number of cars) to the number of potential buyers. A network organization is understood as a set of units operating according to a common scheme. Examples of such organizations are a network of branches of a large automobile company or a chain of Pyaterochka stores, etc.

In this paper, as an illustrative example, we consider the problem of supporting management decisions by the managers of a network of branches of a large university. Everyone is welcome to attend the preparatory courses of such branches. A natural characteristic of the efficiency of the staff of such branches is the ratio of the number of students attending preparatory courses to the number of potential applicants. Information on comparative effectiveness is the basis for making strategic decisions on the development of the branch network.

It is convenient to present a network organization in the form of a graph. The vertices of this graph correspond to divisions. The specificity of the graph we are considering is that it can have both directed and undirected edges. An undirected edge between vertices i, j is added to the graph if and only if it is decided that the i-th and j-th divisions work equally efficiently. A directed edge from vertex i to vertex j is added to the graph if and only if it is settled that the i-th unit is more efficient than the j-th unit. Note that usually either only directed edges or only undirected ones are used in graphs [11–15]. We will use them both simply because they better allow us to emphasize some of the structures of the graph under analysis, which characterize the specifics of the analyzed network organization. Among such structures there will be cliques [16] (a set of vertices, any two of which are connected by a non-directed edge), which characterize a set of divisions working equally efficiently. In what follows, we will call such cliques undifferentiated classes. Another example of a structure is a complete subgraph with only directed edges. Such a structure will be called a structure of ordering or dominance.

In many problems, in particular in the example we are analyzing, it is natural to consider the number of sales as a random variable. At the same time, analysis of real data, especially when there is little of it, can lead to contradictory conclusions. In this case, the corresponding graph contains some logically contradictory structures, for example, subgraphs of three vertices, two edges between which are undirected, and one is directed. This type of problem arose in [17] when discussing the problem of testing hypotheses of homogeneity of at least three populations and was called the incompatibility problem. Applied tasks in which the incompatibility problem arises (the problem of inconsistently combining the results of comparing effectiveness of the pairs of departments) were considered in [4, 5] within the framework of the normal model. The solution to the incompatibility problem was based on the introduction of an additional parameter Δ and the transition to tasks of comparing the effectiveness of two divisions with accuracy Δ. Moreover, if the efficiencies of divisions i, j differed by less than Δ, then it was decided that their effectiveness was the same (with an accuracy of Δ). This technique allows us to solve the incompatibility problem, but leads to an additional problem of choosing Δ. Another goal of this work is to find ways to solve the incompatibility problem without introducing an auxiliary parameter Δ.

In this work, unlike [4, 5], the assumption of a normal distribution of the number of sales is not used. Pairwise comparison of the effectiveness of two units is based on the use of Mann–Whitney tests. The procedure for comparative analysis of divisions by efficiency is based on a combination of nonparametric tests of pairwise comparison of two divisions. This uses a graphical representation, which is convenient for visualizing emerging incompatibility problems. The proposed nonparametric procedure is applied to the analysis of data reported in [4], and an example is given in which the inconsistency problem is overcome by analyzing p-values and appropriately selecting significance levels for pairwise comparison tests. A comparison is made with the results obtained within the normal model.

This article is organized as follows: Section 1 provides the basic notation and formulation of the problem; Section 2 describes the nonparametric procedure for comparative analysis of departments by efficiency, and its graphical representation; Section 3 provides an illustrative example, an example of solving the incompatibility problem, and compares it with the results obtained in [4].

1. Formulation of the problem

It is convenient to present data on the number of sales in the form of a matrix $\| x_{ji} \|$, where $\| x_{ji} \|$ – the ratio of the number of sales to the number of potential buyers in division j in the i-th time period, $j = 1, ..., N$, where N is the number of divisions of the network organization, $i = 1, ..., m_j$, where m_j is the number of analyzed time periods of work j-th division. We will assume that observations x_{ji} represent the values of random variables X_{ji}, which describe the ratio of the number of sales to the number of potential buyers in division j in time period i. Let us assume that all time periods are the same, and the random variables X_{ji} are independent for all $j = 1, ..., N$; $i = 1, ..., m_j$ and for a fixed j are equally distributed as X_j. Let $F_j(x)$ be the distribution function of the random variable X_j.

The problem considered in this work consists in constructing and applying for the analysis of specific data a statistical procedure for distinguishing hypotheses of the form:

$$\begin{aligned} &H_1 : F_1(x) = F_2(x) = \cdots = F_N(x), \quad \forall x \\ &H_2 : F_1(x) < F_2(x) = \cdots = F_N(x), \quad \forall x \\ &\vdots \\ &H_L : F_1(x) < F_2(x) < \cdots < F_N(x), \quad \forall x \end{aligned} \tag{1}$$

Here, hypothesis H_1 means that the efficiency of all divisions is the same, hypothesis H_2 means that division 1 is more efficient than other divisions whose efficiency is the same, etc. Note that relations (1) do not describe all possible relations between the distribution functions $F_j(x)$, $j = 1, ..., N$. We limit ourselves to considering only these hypotheses, since we are only interested in the presence of a systematic shift, which can result from different performance levels of different divisions of a network organization.

Like in [4], we will use the method of constructing procedures with many solutions proposed in [17]. This method is based on reducing a multi-alternative problem to a set of appropriately selected two-alternative generating problems. In our case, to distinguish (1), it is natural to consider two-alternative hypothesis testing problems $h_{ij} : F_{ij}(x) \geq F(x), \forall x, \forall i, j = 1, \ldots, N$.

For fixed i, j combining both the tests φ_{ij}, φ_{ji} of simultaneous testing the hypotheses h_{ij} and h_{ji} with non-zero probability can lead to a logically untenable (for a given x) decision to reject both hypotheses, i.e. to the incompatibility problem. As shown in [17], to eliminate such a contradiction, it is sufficient to require that the significance levels α_{ij}, α_{ji} of the tests φ_{ij}, φ_{ji} satisfy the condition $\alpha_{ij} + \alpha_{ji} < 1$. In this case, such a combination of the tests φ_{ij}, φ_{ji} leads to a joint procedure for distinguishing three hypotheses:

$$\begin{aligned} &H_{ij}^1 : F_i(x) < F_j(x); \\ &H_{ij}^2 : F_i(x) = F_j(x); \\ &H_{ij}^3 : F_i(x) > F_j(x). \end{aligned} \tag{2}$$

However, combining such procedures with three solutions for different i, j can lead to a contradiction, namely: with non-zero probability it can be decided that (for example):

$$F_1(x) = F_2(x) \text{ and } F_2(x) = F_3(x), \text{ but } F_1(x) \neq F_3(x).$$

To eliminate this contradiction, in [4], following the proposal of [17], a slightly modified system of generating hypotheses was considered. However, the studies were limited to the case when $F_j(x)$ is a normal distribution. In the notation of this work, the modified system of generating hypotheses has the form:

$$h'_{ij} : F_{ij}(x) + \Delta \geq F(x), \forall x, \forall i, j = 1, \ldots, N.$$

When combining tests φ'_{ij}, φ'_{jii} for simultaneous testing of hypotheses h'_{ij}, h'_{ji} we obtain a procedure for distinguishing three hypotheses:

$$\begin{aligned} &H'^1_{ij} : F_i(x) + \Delta < F_j(x); \\ &H'^2_{ij} : |F_i(x) - F_j(x)| < \Delta; \\ &H'^3_{ij} : F_i(x) > F_j(x) + \Delta. \end{aligned} \tag{3}$$

In this case, the problem of obtaining contradictory conclusions does not arise. At the same time, the introduction of the Δ parameter formally changes the original problem.

In this work, the Δ parameter is not introduced and the assumption of a normal distribution is not made.

At the same time, one of the interesting questions is to find options for consistently combining statistical rules with three solutions without introducing the Δ parameter.

Note that at present, in the intensively developing theory of simultaneous testing of many hypotheses, no emphasis is placed on the need to solve the problem of incompatibility [18–21]. Moreover, starting from [22], the problem of incompatibility is considered as too strong a requirement imposed on the procedure for simultaneous testing of many hypotheses. Within the framework of the theory of simultaneous testing of many hypotheses, approaches to the construction of procedures that control the probability of at least one error of the first type, the proportion of errors of the first type, and some others are mainly studied. In this work, on the contrary, we focus on solving the incompatibility problem, which allows for an appropriate comparison with the results obtained in [4].

2. Nonparametric comparative analysis procedure and its visualization

2.1. Procedure with three solutions

One of the most effective nonparametric procedures for distinguishing hypotheses (2) is based on the Mann–Whitney statistics [6, 9]. The Mann–Whitney statistic looks like this:

$$W_{ij}(x_i, x_j) = \sum_{r=1}^{m_j} \sum_{s=1}^{m_i} I(x_{is} < x_{jr}), \tag{4}$$

where $I(A)$ is the indicator of event A.

For fixed i, j the procedure with three solutions for distinguishing hypotheses (2) in terms of p-values can be written as

$$\varphi(i,j) = \begin{cases} d_{ij}^1 & p_{ij}^1 \le \alpha_1 \\ d_{ij}^3 & p_{ij}^3 \le \alpha_3 \\ d_{ij}^2 & p_{ij}^1 > \alpha_1,\ p_{ij}^3 > \alpha_3, \end{cases} \tag{5}$$

where d_{ij}^k is the decision to accept the hypothesis H_{ij}^k $(k = 1, 2, 3)$;

p_{ij}^1, p_{ij}^3 are the corresponding p-values, namely:

$$p_{ij}^1 = P_{F_i = F_j}\left(W_{ij}(X_i, X_j) \le W_{ij}(x_i, x_j)\right)$$
$$p_{ij}^3 = P_{F_i = F_j}\left(W_{ij}(X_i, X_j) > W_{ij}(x_i, x_j)\right), \tag{6}$$

where α_1, α – significance levels of tests for hypothesis testing H_{ij}^1 and H_{ij}^3, respectively.

It is assumed that among the observed values and there are no equals. The necessary adjustments in the case of equal observations can be made based on the methodology outlined in [9].

Tables of distribution of statistics (4) for small m_i, m_j are given in [9]. For large m_i, m_j, one can use the normal distribution

$$N\left(\frac{m_i \cdot m_j}{2}, \frac{m_i \cdot m_j \cdot (m_i + m_j + 1)}{12}\right),$$

which is recommended to be used when

$$\min(m_i, m_j) > 50 \ [9].$$

From (6) it is obvious that for fixed i, j, $p_{ij}^1 + p_{ij}^3 = 1$. Therefore, to apply the procedure with three solutions (5), information about the minimum p-value is sufficient:

$$p_{ij} = \min(p_{ij}^1, p_{ij}^3). \tag{7}$$

2.2. Procedure with many solutions and its graphical representation

We will obtain the procedure for distinguishing hypotheses (1) by combining procedures (5). This procedure can be written as:

$$\delta = \left(\varphi(1,2), \varphi(2,3), \ldots, \varphi(N-1, N)\right). \tag{8}$$

Let denote the efficiency of the i-th department. In the problem under consideration, two types of relationships are possible between the performance of divisions of a network organization (dominance or equivalence). The entry means that the i-th unit works more efficiently (dominates) than the j-th unit. The entry means that the i-th and j-th divisions work equally efficiently (equivalence). To visually analyze the results of applying procedure (8), we will use the technique proposed in [23–25].

For a given vector $(f_1, ..., f_N)$, we introduce matrix D with elements

$$d_{ij} = \begin{cases} 1, & f_i > f_j \\ 0, & f_i = f_j \\ -1, & f_i < f_j \end{cases}$$

and a matrix B with elements

$$b_{ij} = \begin{cases} 1, & f_i > f_j \\ 0, & \text{otherwise}\,. \end{cases}$$

It is easy to show [24] that the matrix B is related to the matrix D by the relation

$$D = B - B^T,$$

where B^T is the transposed matrix B.

The relationships described by the matrix B are more easily interpreted if the rows and columns of the matrix B are rearranged in such a way as to obtain an upper triangular shape (i.e., to collect, if possible, all the ones above the main diagonal of the matrix B). To obtain the upper triangular form of the matrix B, one can arrange in descending order the rows (and columns) of the matrix D by the sums of the row elements. Replacing the -1 elements of the resulting matrix with 0, we obtain matrix B, which is most consistent with the upper triangular shape. Matrix B, which is most consistent with the upper triangular form, allows us to identify the so-called "undifferentiated classes" [25]. The term "undifferentiated class" will be used to designate the largest set of units whose performance levels are not significantly different from each other. The term "largest" means that for any unit i that does not belong to a given undifferentiated class, there is at least one unit j from that class such that the performance efficiencies of units i and j are meaningfully distinguishable. The matrix B, which is most consistent with the upper triangular form, shows the undifferentiated classes as square sub-matrices, symmetric about the main diagonal, with all elements equal to 0.

Obviously, the overlapping undifferentiated classes resulting from the delta procedure mean that there is an incompatibility problem. It is convenient to visualize the matrix D in the form of a graph $G = (V, E)$, where $V = \{1, 2, ..., N\}$ is the set of vertices of the graph, $E = \{e_{ij}\}$ is the set of edges of the graph. If element d_{ij} of matrix D is equal to 1, then a directed edge from vertex i to vertex j is added to the graph G. In this case, vertex i dominates vertex j. In [11], vertex i is called the parent of vertex j, and vertex j is called the child of vertex i. If vertex i is connected by a directed path of length greater than 1 to some vertex k, then vertex i is called the ancestor of vertex k, and vertex k is called a descendant of vertex i. If element d_{ij} of matrix D is equal to 0, then an undirected edge between vertices i and j is added to graph G. If the element $d_{ij} = -1$, then, since $d_{ij} = -d_{ji}$, the graph G already contains a directed edge from vertex j to vertex i. It is obvious that all the vertices of this graph, corresponding to divisions from a certain undifferentiated class, are connected to each other by undirected edges and therefore represent cliques of the graph G. Below we will separately depict sub-graphs with only directed edges and only with undirected edges.

Note that the proposed representation will clearly reflect the existence of problems of incompatibility of the obtained conclusions, if any. Obviously, if in a representation different undifferentiated classes contain the same vertices, then the incompatibility problem occurs.

3. Illustrative example

Let us consider the task of comparative analysis of the performance of university branches, which was briefly described in the introduction. Let us denote 1*f* as the first branch of the university, 2*f* – the second branch of the university, etc. Data for analysis are borrowed from [4] and are shown i*n Table 1.*

Minimum p-values (7) of tests (5) are given in *Table 2.*

3.1 Construction of undifferentiated classes

Let us first consider the traditional significance level. Matrix $D_{0.05}$ is shown in *Table 3.*

Table 1.

Data on the number of students attending preparatory courses in various branches

$1f$	$2f$	$3f$	$4f$	$5f$	$6f$	$7f$	$8f$
103	131	187	154				
92	212	262	92	151	99	235	
122	197	376	129	164	268	338	77
48	143	283	146	141	217	239	63
86	95	231	125	140	231	187	59
89	70	203	127	173	175	123	78
147	92	276	183	141	137	139	82
134	95	258	213	187	242	185	28
Number of potential applicants in the i-th branch							
6390	7090	28900	6320	6320	11130	4660	2530

Table 2.

Minimum p-values

	$1f$	$2f$	$3f$	$4f$	$5f$	$6f$	$7f$
$2f$	p^1_{12} 0.4392						
$3f$	p^3_{13} 0.0023	p^3_{23} 0.0103					
$4f$	p^1_{14} 0.0364	p^1_{24} 0.0652	p^1_{34} 0.0001				
$5f$	p^1_{15} 0.0006	p^1_{25} 0.0469	p^1_{35} 0.0002	p^1_{45} 0.1678			
$6f$	p^1_{16} 0.3063	p^1_{26} 0.4775	p^1_{36} 0.0070	p^3_{46} 0.0760	p^3_{56} 0.0020		
$7f$	p^1_{17} 0.0002	p^1_{27} 0.0011	p^1_{37} 0.0002	p^1_{47} 0.0200	p^1_{57} 0.0012	p^1_{67} 0.0003	
$8f$	p^3_{18} 0.0147	p^3_{28} 0.0539	p^1_{38} 0.0007	p^1_{48} 0.1725	p^1_{58} 0.1830	p^1_{68} 0.0256	p^3_{78} 0.0175

Table 3.

Matrix $D_{0.05}$

	1*f*	2*f*	3*f*	4*f*	5*f*	6*f*	7*f*	8*f*	Σ
1*f*	–	0	1	–1	–1	0	–1	–1	–3
2*f*	0	–	1	0	–1	0	–1	0	–1
3*f*	–1	–1	–	–1	–1	–1	–1	–1	–7
4*f*	1	0	1	–	0	0	–1	0	1
5*f*	1	1	1	0	–	1	–1	0	3
6*f*	0	0	1	0	–1	–	–1	–1	–2
7*f*	1	1	1	1	1	1	–	1	7
8*f*	1	0	1	0	0	1	–1	–	2

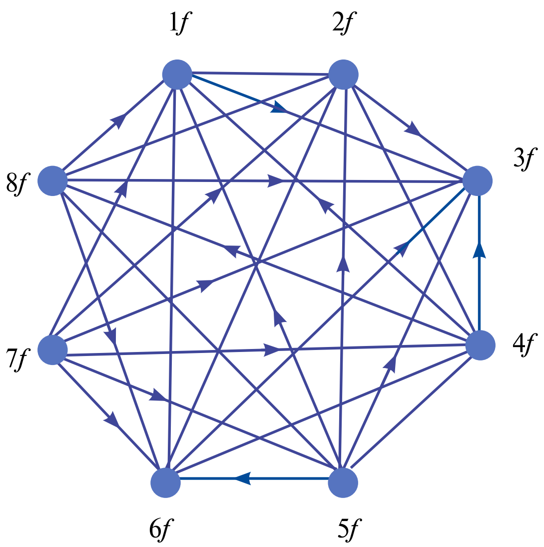

Fig. 1. Graphical representation of the matrix $D_{0,05}$.

A graphical representation of the $D_{0.05}$ matrix is shown in *Fig. 1*.

Matrix $B_{0.05}$ obtained from matrix $D_{0.05}$ reduced to upper triangular form is shown in *Table 4*.

Table 4.

Matrix $B_{0.05}$, educed to upper triangular form

	7*f*	5*f*	8*f*	4*f*	2*f*	6*f*	1*f*	3*f*
7*f*	–	1	1	1	1	1	1	1
5*f*		–	0	0	1	1	1	1
8*f*			–	0	0	1	1	1
4*f*				–	0	0	1	1
2*f*					–	0	0	1
6*f*						–	0	1
1*f*							–	1
3*f*								–

A graphical representation of the matrix $B_{0.05}$ reduced to the upper triangular form is shown in *Fig. 2*.

In *Fig. 2* it is easy to see that there are 6 cliques in this graph: {7}, {4, 5, 8}, {2, 4, 8}, {2, 4, 6}, {2, 6, 1}, {3}. Note that these cliques have common vertices, for example, vertex 4 belongs to the 3rd cliques. This indicates that there is an incompatibility problem. Thus, for $\alpha_{ij} = 0.05$, $\forall\, i, j = 1, \ldots, 8$ disjoint undifferentiated classes are not distinguished.

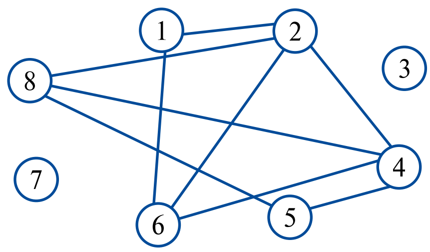

Fig. 2. Graphical representation of the matrix $B_{0.05}$ reduced to upper triangular form. The branch number is indicated by a number.

It is well known that there are no strict rules for choosing the significance level α. A possible way to reduce the number of incompatibility problems arising could be based on changing the significance level α. Analysis of the p-values given in *Table 2* suggests the advisability of constructing and studying graphical models when choosing α_{ij}, $\forall i, j = 1, ..., 8$ from the interval (p^3_{46} =0.076; p^1_{45} =0.1678). For definiteness, we choose $\alpha_{ij} = 0.1$, $\forall i, j = 1, ..., 8$. The matrix $D_{0.1}$ is given in *Table 5*.

Table 5.

Matrix $D_{0.1}$

	1*f*	2*f*	3*f*	4*f*	5*f*	6*f*	7*f*	8*f*	Σ
1*f*	–	0	1	–1	–1	0	–1	–1	–3
2*f*	0	–	1	–1	–1	0	–1	–1	–3
3*f*	–1	–1	–	–1	–1	–1	–1	–1	–7
4*f*	1	1	1	–	0	1	–1	0	3
5*f*	1	1	1	0	–	1	–1	0	3
6*f*	0	0	1	–1	–1	–	–1	–1	–3
7*f*	1	1	1	1	1	1	–	1	7
8*f*	1	1	1	0	0	1	–1	–	3

A graphical representation of the matrix $D_{0.1}$ is shown in *Fig. 3*.

The matrix $B_{0.1}$ obtained from the matrix $D_{0.1}$ reduced to upper triangular form is shown in *Table 6*.

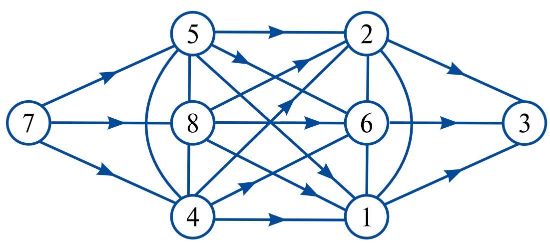

Fig. 3. Graphical representation of the matrix $D_{0.1}$.
The branch number is indicated by a number.

Table 6.

Matrix $B_{0.1}$ reduced to upper triangular form

	7*f*	5*f*	8*f*	4*f*	2*f*	6*f*	1*f*	3*f*
7*f*	–	1	1	1	1	1	1	1
5*f*		–	0	0	1	1	1	1
8*f*			–	0	1	1	1	1
4*f*				–	1	1	1	1
2*f*					–	0	0	1
6*f*						–	0	1
1*f*							–	1
3*f*								–

A graphical representation of the matrix $B_{0.1}$ reduced to the upper triangular form is shown in *Fig. 4*.

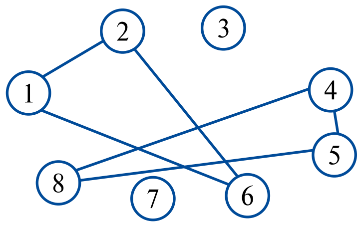

Fig. 4. Graphical representation of the matrix $B_{0.1}$ reduced to upper triangular form.
The branch number is indicated by a number.

In *Fig. 4* it is easy to see that in this graph there are 4 cliques {7}, {4, 5, 8}, {2, 6, 1}, {3} and these cliques do not have common vertices. This indicates that at the problem of incompatibility does not arise, i.e., the undifferentiated classes do not intersect.

3.2 Ordering construction of structures

A graphical representation of the matrix $D_{0.05}$ shows the structure of ordering branches according to the efficiency of their work, shown in *Fig. 5*.

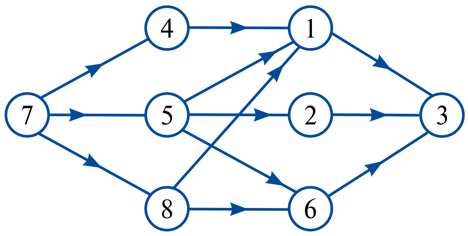

Fig. 5. Nonparametric ordering $\alpha = 0.05$.
The branch number is indicated by a number.

Only dominance relationships are shown; equivalence relationships and ancestor-descendant relationships are not shown. In particular, there is no directed edge between vertices 7 and 3, because there is a directed path $7 \to 4 \to 1 \to 3$ from the ancestor "$7f$" to the descendant "$3f$", meaning strict ordering: $7f$ is more effective than $4f$, $4f$ is more effective than $1f$, $1f$ is more effective than $3f$. Note that there are no directed paths in this graph $7 \to 4 \to 2 \to 3$; $7 \to 4 \to 6 \to 3$; $7 \to 8 \to 2 \to 3$. This indicates the absence of complete ordering between the operating efficiencies of branches at $\alpha = 0.05$. Considering that $\{4f, 5f, 8f\}$ belong to the same undifferentiated class, the absence of such paths leads to logical contradictions. We emphasize that at $\alpha = 0.1$ complete ordering takes place (see *Fig. 6*) and logical contradictions do not arise.

3.3. Comparison

In *Fig. 7* shows graphs constructed from the results of comparing the operating efficiencies of branches both under the assumption of a normal distribution of the random variables under study [4] and in a nonparametric formulation.

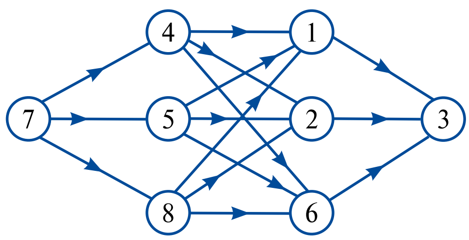

Fig. 6. Nonparametric ordering $\alpha = 0.1$.
The branch number is indicated by a number.

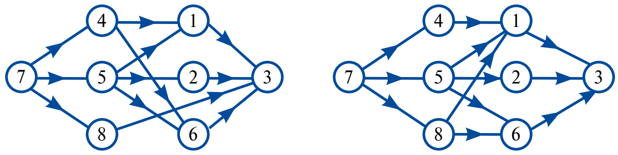

Fig. 7. Parametric (left),
nonparametric (right) ordering at α =0.05.
The branch number is indicated by a number.

The ordering graphs shown in *Fig. 7*, differ in three edges, namely: edge (4, 6) is present in parametric ordering, and absent in non-parametric ordering; edges (8, 1); (8, 6) is present with nonparametric ordering, and absent with parametric ordering.

Linear ordering, constructed according to the scheme proposed in [4], corresponding to nonparametric ordering (*Fig. 7*, right), has the form:

$$f_3 < f_1 \le f_6 \le f_2 \le f_4 \le f_8 \le f_5 < f_7 \text{ with precision } \Delta. \quad (9)$$

Obtaining such a linear ordering, formally proposed in [4] in a slightly different formulation (ordering with accuracy Δ), is based on analyzing the number of directed links leaving a specific vertex or entering a specific vertex. Namely, since vertices 4, 5, 8 (*Fig. 7*, right) are not connected by directed edges, at first glance it seems that the solution $f_4 = f_8 = f_5$ can be made with an accuracy of Δ. However, since vertex 4 dominates only one vertex 1 (vertex 4 has one directed edge going out), vertex 8 dominates two vertices 1 and 6 (vertex 8 has two directed edges going out), and vertex 5 dominates three vertices 1, 2, 6 (three directed edges emerge from vertex 5), then the solution $f_4 \le f_8 \le f_5$ is obtained with an accuracy of Δ. Writing $f_4 \le f_8$ with precision Δ means that $f_4 + \Delta < f_8$ or $|f_4 - f_8| < \Delta$. Similarly, since vertices 1, 2, 6 (*Fig. 7*, right) are not connected by directed edges, at first glance it seems that a decision $f_4 \le f_8 \le f_5$ with an accuracy of Δ can be made. However, since vertex 1 is a descendant of all vertices 4, 5, 8 (vertex 1 includes three directed edges), vertex 6 is a descendant of two vertices 5 and 8 (vertex 6 includes two directed edges), vertex 2 is a descendant of one vertex 5 (vertex 2 includes one directed edge), then the solution $f_1 \le f_6 \le f_2$ is made with accuracy Δ.

The linear ordering obtained in [4] has the form:

$$f_3 < f_1 = f_6 \le f_2 \le f_8 \le f_4 \le f_5 \le f_7 \text{ with precision } \Delta. \quad (10)$$

Orderings (9) and (10) differ very slightly. In fact, the extreme elements of the orderings coincide, the sign of $f_1 \le f_6$ has changed compared to the sign of $f_1 = f_6$, and the non-strict ordering $f_4 \le f_8$ has changed to the ordering $f_8 \le f_4$. But in both cases, no significant difference in the operating efficiency of branches f_8 and f_4 was found. This indirectly indicates the acceptability of the normal model proposed in [4].

The undifferentiated classes presented in *Fig. 8* are distinguished by three edges, namely: the edge (4, 6) is present in the nonparametric construction of undifferentiated classes, and is absent in the parametric construction; edges (8, 1); (8, 6) is present in the parametric construction of undifferentiated classes, and is absent in the nonparametric construction. This is quite consistent with the comparison of ordering structures (see *Fig. 7*).

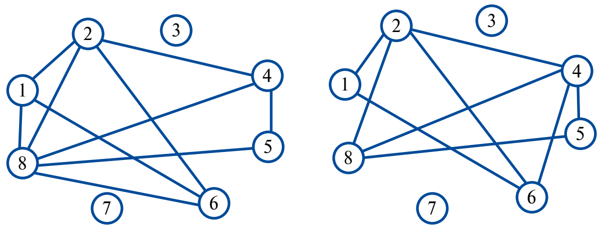

Fig. 8. Parametric (left) and nonparametric (right) undifferentiated classes at $\alpha = 0.05$. The branch number is indicated by a number.

Conclusion

In this work, a nonparametric procedure for comparative analysis of the performance of several divisions of a network organization based on a small volume of observations has been constructed. We give an example of the application of the proposed approach to a comparative analysis of the performance of university branches. The results of the comparative analysis obtained by the proposed nonparametric procedure are compared with the results obtained within the framework of the normal model [4]. It is shown that the results of nonparametric ordering without introducing an additional uncertainty parameter Δ and the ordering results obtained within the normal model with the introduction of Δ are quite close. An example of a completely consistent comparison of the performance of several divisions of a network organization is provided. ■

Acknowledgments

The results of sections 1 and 2 were prepared within the framework of the Fundamental Research Program of the National Research University "Higher School of Economics" (HSE University). The results of section 3 were prepared with the support of a grant from the Russian Science Foundation (project 22-11-00073).

References

1. Aleskerov F.T., Belousova V. (2007) Effective development of the branch network of a commercial bank. *Upravleniye v kreditnoy organizatsii (Management in a credit institution)*, no. 6, pp. 23–34 (in Russian).
2. Kryukov A.M. (2010) Efficiency analysis of branch and division activities is a prerequisite for business stability. *Strategic decisions and risk management*, vol. 4, pp. 84–87 (in Russian). https://doi.org/10.17747/2078-8886-2010-4-84-87
3. Myznikova M.A. (2022) Quality of strategic management under ambiguity: Assessment within the framework of sustainable development. *Business Informatics*, vol. 16, no. 3, pp. 36–52. https://doi.org/10.17323/2587-814X.2022.3.36.52
4. Koldanov A.P., Koldanov P.A. (2012) Multiple decision procedures for the analysis of higher school entry selection results. *Business Informatics (Biznes-informatika)*, vol. 19, no. 1, pp. 24–31 (in Russian).

5. Koldanov P.A. (2013) Efficiency analysis of branch network. *Models, Algorithms, and Technologies for Network Analysis* (eds. B.I. Goldengorin, V.A. Kalyagin, P.M. Pardalos). Springer Proceedings in Mathematics & Statistics, vol. 59, pp. 71–83. https://doi.org/10.1007/978-1-4614-8588-9_5

6. Hettmansperger T. (1987) *Statistical Inference Based on Ranks*. Moscow: Finance and Statistics (in Russian).

7. Kendall M. (1975) *Rank correlations*. Moscow: Statistics (in Russian).

8. Hollander M., Wolfe D. (1983) *Nonparametric statistical methods*. Moscow: Finance and Statistics (in Russian).

9. Lehmann E.L. (1975) *Nonparametrics: Statistical methods based on ranks*. San Francisco: Hoden-Day. https://doi.org/10.1002/zamm.19770570922

10. Kendall M., Stuart A. (1973) *Statistical Inferences and relationships*. Moscow: Science (in Russian).

11. Jordan M.I. (2004) Graphical models. *Statistical Science*, vol. 19, no. 1, pp. 140–155. https://doi.org/10.1214/088342304000000026

12. Newman M.E.J. (2010) *Networks. An introduction*. New York: Oxford University Press.

13. Opsahl T., Panzarasa P. (2009) Clustering in weighted networks. *Social Networks*, vol. 31, no. 2, pp. 155–163.

14. Horvath S. (2011) *Weighted network analysis. Applications in genomics and systems biology*. New York: Springer. https://doi.org/10.1007/978-1-4419-8819-5

15. Li S., He J., Zhuang Y. (2010) A network model of the interbank market. *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 24, pp. 5587–5593. https://doi.org/10.1016/j.physa.2010.08.057

16. Boginski V., Butenko S., Pardalos P.M. (2005) Statistical analysis of financial networks. *Computational Statistics & Data Analysis*, vol.48, no. 2, pp. 431–443. https://doi.org/10.1016/j.csda.2004.02.004

17. Lehmann E.L. (1957) A theory of some multiple decision problems, I. *The Annals of Mathematical Statistics*, vol. 28, no. 1, pp. 1–25. https://doi.org/10.1214/aoms/1177707034

18. Bretz F., Hothorn T., Westfall P. (2010) *Multiple comparisons using R*. New York: Chapman and Hall/CRC. https://doi.org/10.1201/9781420010909

19. Hochberg Y., Tamhane A.C. (1987) *Multiple comparison procedures.* New York: John Wiley & Sons. https://doi.org/10.1002/9780470316672

20. *Handbook of Multiple Comparisons* (2021) (eds. X. Cui, T. Dickhaus, Y. Ding, J.C. Hsu). New York: Chapman and Hall/CRC. https://doi.org/10.1201/9780429030888

21. Dickhaus T. (2014) *Simultaneous statistical inference. With applications in the life sciences*. Heidelberg: Springer Berlin. https://doi.org/10.1007/978-3-642-45182-9

22. Gabriel K.R. (1969) Simultaneous test procedures – some theory of multiple comparisons. *The Annals of Mathematical Statistics*, vol. 40, no. 1, pp. 224–250. https://doi.org/10.1214/aoms/1177697819

23. Cliff N. (1975) Complete orders from incomplete data: Interactive ordering and tailored testing. *Psychological Bulletin*, vol. 82, no. 2, pp. 289–302. https://doi.org/10.1037/h0076373

24. Cliff N. (1993) Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, vol. 114, no. 3, pp. 494–509. https://doi.org/10.1037/0033-2909.114.3.494

25. Donoghue J.R. (2004) Implementing Shaffer's multiple comparison procedure for a large number of groups. *Recent Developments in Multiple Comparison Procedures* (Eds. Benjamini, Y., Bretz, F. and Sarkar, S.), Institute of Mathematical Statistics, Beachwood, Ohio, pp. 1–23.

About the authors

Petr A. Koldanov

Dr. Sci. (Phys.-Math.);

Senior Research Fellow, Laboratory of Algorithms and Technologies for Networks Analysis, HSE University (campus in Nizhny Novgorod), 136, Rodionova St., Nizhny Novgorod 603093, Russia;

E-mail: pkoldanov@hse.ru

ORCID: 0000-0001-5961-0282

Vladimir A. Koldanov

Cand. Sci. (Phys.-Math.);

Associate Professor, Department of Mathematics and Computer Science, Nizhny Novgorod State University of Engineering and Economics, 22A, building 2, Oktyabrskaya St., Nizhny Novgorod Region, Knyaginino 606340, Russia;

E-mail: vlad.kold@gmail.com

ORCID: 0009-0006-1924-746X