THE RESEARCH OF POSSIBILITIES FOR FAST CALCULATION OF MEDIAN CONSENSUS RANKINGS

Досліджено можливість швидкого обчислення колективних експертних оцінок медіанного типу. Незважаючи на широке застосування для розрахунку колективних експертних оцінок медіан Кемені-Снелла і Кука-Сейфорда, недостатньо досліджені можливості скорочення часу обчислення медіанного консенсусного ранжирування шляхом застосуванням задачі про призначення і відомих алгоритмів її рішення. На відміну від більшості відомих методів пропонований в статті метод не є наближеним і зберігає вихідну медіанну аксіоматику Кемені.

Досліджено можливість розрахунку медіан Кемені-Снелла і Кука-Сейфорда із застосуванням задачі про призначення методами комп'ютерного експерименту. Оцінено час рахунку медіанних ранжирувань чотирма різними алгоритмами рішення задачі про призначення. Встановлено, що запропонованим методом при помірній кількості альтернатив (n<50) медіанні ранжирування розраховуються в режимі часу, близькому до реального. Показано також, що на відміну від інших методів час обчислення медіанних ранжирувань за допомогою задачі про призначення не залежить від ступеня узгодженості індивідуальних експертних ранжирувань.

Отримані результати корисні для практичного застосування дослідженої процедури в мережевих експертних системах. У таких системах час обчислення консенсусного ранжирування має бути близьким до реального. Крім того, в мережевих системах експертизи за рахунок випадкового комплектування колективу експертів можливий низький рівень узгодженості індивідуальних ранжирування. Для дослідженої процедури це не впливає на тривалість розрахунку. Це дозволяє рекомендувати розроблену обчислювальну процедуру швидкого пошуку медіанних консенсусних ранжирувань за Кемені-Снеллом і Куком-Сейфордом для практичного застосування в системах колективної мережевої експертизи

Ключові слова: колективне експертне оцінювання, медіанні консенсусні ранжирування, задача про призначення

V. Boltenkov
PhD, Associate Professor*
E-mail: vaboltenkov@gmail.com
V. Kuvaieva
Postgraduate student*
E-mail: vkuvayeva@opu.ua
O. Galchonkov
PhD, Associate Professor*
E-mail: o.n.galchenkov@gmail.com
A. Ishchenko
Senior Lecturer
Department of Applied Mathematics and Information Technologies**
E-mail: alesya.ishchenko@gmail.com
*Department of Information Systems**
**Institute of Computer Systems
Odessa National Polytechnic University
Shevchenka ave., 1, Odessa, Ukraine, 65044

1. Introduction

The task on searching for an aggregate expert opinion along a ranking scale based on individual rankings of a group of experts, initiated as the classic problem about collective expert estimation, has been widely used in various technical and socio-economic applications. These include the tasks on social choice (or vote), multiagent resource allocation problems, a multicriteria choice [1]. The applied interest in aggregate ranking of preferences received new impetus with the beginning of the 21th century with the emergence of search engines and meta-search engines. In such problems, rankings of large arrays of data from electronic sources are aggregated (specifically global and local networks) [2].

One of the fields of wide application of the aggregation of expert preferences is the network expertise [3], in which participants of social networking and professional networking communities act as experts [4, 5]. For the tasks of network expertise, it is a very principal requirement to obtain an accurate aggregated expert estimate in real time peer.

The most theoretically justified method for constructing aggregated rankings is the median method, initiated in paper [6] and named the Kemeny's median. Median aggregation in ranking scales was further developed in studies [7, 8], in which the Cook-Seiford median was proposed. The main problem in calculating the median aggregate rankings is that all the medians are the NP-complete problems and require the non-polynomial cost of computing time. Despite the existence of multiple heuristic algorithms to reduce computing time for the problems on the median aggregation based on ranking scales, this issue cannot be considered to be completely resolved. In this context, it is promising, though unfairly disregarded by specialists, to find the median aggregation by stating it as the assignment problem (AP), which is a linear problem of integer programming, well known in the study of operations [9]. The assignment problem has a polynomial complexity, a number of efficient algorithms were developed in detail to solve it. Thus, research into economical alternatives for the calculation of strict aggregated rankings applying the assignment problem is rather relevant.

2. Literature review and problem statement

The search for a consensus ranking implies the following: K experts are offered to consider n alternatives that the experts organize based on personal preferences. In accordance with positions that the alternatives take based on the individual preferences of experts, the alternatives are numbered along a ranking scale, forming an individual ranking. In the process of aggregation (merging, based on any rule, denoted as an aggregation function) of individual rankings of experts, there should be formed a collective ranking, commonly referred to as a consensus if it expresses opinions of all K experts in the best fashion. Historically, the first methods for consensus rankings are the methods of social choice or voting, based on the principles of «fair» vote [10]. Methods of social choice are entirely based on heuristic rules [11]. These methods lead to well-known controversial collective decisions, known as the «voting paradoxes» [12]. Specifically, it is a well-known possible violation of the transitivity of a result – the Condorcet paradox [13]. Therefore, methods of social choice cannot be considered sufficiently substantiated mathematically. The first attempt at a rigorous mathematical formalization of the problem was paper [6], which formulated the axiomatic method for aggregating the rankings based on the distance between them. The method of aggregation, proposed in [6], was named the Kemeny-Snell median, and immediately drew the attention of specialists as consensus devoid of weaknesses inherent to the algorithms of social choice. Note that paper [6] did not estimate the computational complexity for the authors' proposed median consensus ranking. Later, [7] pointed out that the Kemeny-Snell median is a problem of the non-polynomial complexity, and [8] proved that its computational complexity is estimated to be $O(n!)$. Paper [14] gave impressive estimates for the volume of computing the Kemeny-Snell median when solving a problem by the brute-force method. Specifically, it is shown that at $n = 17$ the total number of considered rankings is expressed by a figure with 18-th decimal digits.

The non-polynomial complexity of the Kemeny-Snell problem has led to an intensive search for the methods to reduce it to the polynomial problem. At present, a number of algorithms have been proposed that solve the problem on the median consensus ranking by reducing it to the problem of integer linear programming. It is proposed to subsequently solve this problem using a branch and bound method [15].

In [16], the Kemeny's median is computed by the evolutionary algorithm. Genetic algorithm for the economical calculation of the Kemeny-Snell median is shown in paper [17]; it stresses that the algorithm is heuristic. Paper [18] proposed the algorithm of genetic maps to solve the problem. Two heuristic variants of the branch and bound method is presented in [19] to find the Kemeny-Snell median. Another approach to reduce the complexity of computing the median ranking by searching based on a tree with constraints is described in paper [20]. In [21], the median ranking is found applying the algorithm of ant colonies. [15] pointed out that by the time that paper was published, there were 104 known calculation algorithms for the Kemeny-Snell median. The main disadvantage of all these approaches to reducing the computational complexity in the calculation of the Kemeny-Snell median is that all these algorithms disrupt the original axiomatics and are more or less approximate.

An alternative median consensus ranking is the Cook-Seiford median [7]. In their work, Cook and Seiford proposed not only the new median ranking, satisfying the axiomatics of Kemeny-Snell, but also showed that the proposed median reduces to the assignment problem, well-known in the study of operations. We shall emphasize this point. The purpose of solving an AP is the allocation of n indivisible resources to n objects at a minimum cost (or a maximum profit) as a result of the allocation. One object can only be assigned with a single resource. For the first time, AP and its solving algorithm were formulated in paper [22]. Because AP to a large extent was based on studies by Hungarian mathematicians in 1930s, in [22] the proposed algorithm for solving AP was named the «Hungarian method». Half a century of AP use in applied mathematics helped investigate it in detail [23]. It is proven that the Hungarian algorithm for solving AP has a polynomial complexity of not larger than $O(n^4)$ [24]. The proof that the complexity of solving AP could be reduced to $O(n^3)$ was given in [9]. Paper [25] showed that the polynomial complexity could be lowered below 3. It was also proven that the Kemeny-Snell median could also be stated as the assignment problem [26].

The above analysis allows us to argue on the following. Median consensus rankings are the only precise solution to the problem on collective expert estimation. However, both the Kemeny-Snell median and the Cook-Seiford median are the problems of the non-polynomial complexity. A large number of the proposed algorithms to reduce the computational complexity of calculating the medians do not meet the axiomatics of Kemeny and are approximate. At the same time, it is known that the median rankings can be reduced without any approximations to the assignment problem. A given problem can be solved over a polynomial time. However, this issue has not been sufficiently investigated, neither theoretically nor practically. The research into this issue has necessitated our article.

3. The aim and objectives of the study

The aim of this study is to investigate the possibilities to effectively solve the problem on computing the median consensus ranking applying the assignment problem and known algorithms for solving it.

To accomplish the aim, the following tasks have been set:
– to formalize the median rankings and representation of medians in the form of AP;
– to analyze AP and algorithms for solving it;
– to run a computational experiment aimed at estimating the time required to calculate median rankings formulated in the form of AP.

4. Median rankings and reducing the median rankings to the assignment problem

We shall state the problem on aggregating the individual ranking preferences in the following form [14, 27]. Let there be a set of alternatives $\boldsymbol{A} = \{A_1, A_2,..., A_n\}$, which are arranged by a group of K experts. Each expert from k ($k = 1,..., K$) arranges alternatives based on personal preference and represents individual ranking:

$$\mathbf{P}^k = \left\{ A_{k_1}, A_{k_2},..., A_{k_n} \right\}. \tag{1}$$

Each individual ranking (1) can be represented as:

$$\mathbf{P}^k = \left\{ q_1^k, q_2^k, ..., q_n^k \right\}, \qquad (2)$$

where q_i^k is the position taken by the i-th alternative in the ranking of the k-th expert.

Collective consensus ranking, based on the concept of distance, implies determining the ranking $\dot{\mathbf{P}}$ that is closest to all individual rankings according to a certain introduced measure. In this case, result $\dot{\mathbf{P}}$ is called a median consensus ranking.

The Kemeny-Snell median.

In paper [6], Kemeny and Snell formulated four axioms that must be met for any measure of distance between two rankings, and proved the existence of a distance metric, which satisfies all these axioms, and its uniqueness.

The Kemeny-Snell axioms.

Axiom 1: $d(\mathbf{P}^1; \mathbf{P}^2)$ satisfies three properties of the standard metric (or distance):

1) Positivity: equality $d(\mathbf{P}^1; \mathbf{P}^2) \geq 0$, holds if and only if $\mathbf{P}^1 = \mathbf{P}^2$;

2) Symmetry: $d(\mathbf{P}^1; \mathbf{P}^2) = d(\mathbf{P}^2; \mathbf{P}^1)$;

3) Triangle inequality:

$$d(\mathbf{P}^1; \mathbf{P}^3) \leq d(\mathbf{P}^1; \mathbf{P}^2) + d(\mathbf{P}^2; \mathbf{P}^3)$$

for any three rankings $\mathbf{P}^1$, $\mathbf{P}^2$, $\mathbf{P}^3$, and the equality hods if and only if ranking $\mathbf{P}^2$ is located between $\mathbf{P}^1$ and $\mathbf{P}^3$.

Axiom 2 (invariance): distance d is invariant relative to designations, that is, at identical permutations of objects within rankings $\mathbf{P}^1$ and $\mathbf{P}^2$ the distance between new rankings $\mathbf{P}^{1,1}$ and $\mathbf{P}^{2,1}$ equals $d(\mathbf{P}^1; \mathbf{P}^2)$.

Axiom 3 (coherence): If two rankings $\mathbf{P}^1$ and $\mathbf{P}^2$ are the same, except for the set $\mathbf{S}$ of k elements, which is a segment of both, then $d(\mathbf{P}^1; \mathbf{P}^2)$ equals the distance between the rankings of only those k elements.

Axiom 4 (scaling): minimal positive distance d between rankings is equal to unity.

It was proven in [6] that axioms 1–4 uniquely determine distance $d(\mathbf{P}^1; \mathbf{P}^2)$ at any length of rankings $n \geq 2$, and expression:

$$Med\_KemSn\left(\mathbf{P}^1, \mathbf{P}^2, ..., \mathbf{P}^K\right) = \arg\min_{\mathbf{P}} \sum_{i=1}^{K} d\left(\mathbf{P}, \mathbf{P}^{(i)}\right) \qquad (3)$$

defines the only one distance d that satisfies axioms 1–4, which is named the Kemeny-Snell median.

The Cook-Seiford median.

An alternative axiomatical collective ranking is the median that was proposed in a paper by Cook-Seiford [7]. The distance between two rankings by Cook-Seiford is determined in the space of positions taken by alternatives in individual rankings:

$$d\left(\mathbf{P}^{k_1}, \mathbf{P}^{k_2}\right) = \sum_{i=1}^{n} \left| q_i^{k_1} - q_i^{k_2} \right|, \qquad (4)$$

where q_i^k is the position taken by the i-th alternative in the k-th ranking of an expert.

Papers [7, 8] show that distance (4) satisfies the axiom system by Kemeny-Snell. The distance of consensus vector $\mathbf{P}$ from the entire set of expert rankings $\{\mathbf{P}^k\}$ is equal to:

$$d\left(\mathbf{P}, \mathbf{P}^{(k)}\right) = \sum_{k=1}^{K} d\left(\mathbf{P}, \mathbf{P}^k\right) = \sum_{k=1}^{K} \sum_{i=1}^{n} \left| q_i^k - q_i \right|, \qquad (5)$$

where q_i is the position of alternative A_i in ranking $\mathbf{P}$.

The optimization problem by Cook-Seiford can be formulated in the form: it is required to find such a ranking $\dot{\mathbf{P}}$ (the Cook-Seiford median) for which:

$$Med\_CookSeif\left(\mathbf{P}^1, \mathbf{P}^2, ..., \mathbf{P}^K\right) =$$

$$= \arg\min_{\mathbf{P}} \sum_{k=1}^{K} d\left(\mathbf{P}^k, \dot{\mathbf{P}}\right) = \arg\min_{\mathbf{P}} \sum_{k=1}^{K} \sum_{i=1}^{n} \left| q_i^k - q_i \right|. \qquad (6)$$

In [8], it was demonstrated that the computation of the proposed median is equivalent to solving a problem of linear integer programming.

If we assume that alternative A_i may take the j-th position in ranking $\mathbf{P}$ ($j = 1, 2, ..., n$), then distance (6) can be written as:

$$d\left(\dot{\mathbf{P}}, \mathbf{P}^{(k)}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} y_{ij}, \qquad (7)$$

where

$$d_{ij} = \sum_{k=1}^{K} \left| d_i^k - j \right|,$$

$$y_{ij} = \begin{cases} 1, & \text{if } A_i \text{ takes a place } j \text{ in ranking } \mathbf{P}, \\ 0, & \text{otherwise.} \end{cases}$$

The minimization of expression (7) can be represented as the assignment problem, known in the linear discrete programming [9]:

$$\arg\min_{y_{ij}} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} y_{ij} \qquad (8)$$

with constraints:

$$\mathop{\forall}_{i=1,...,n} \sum_{j=1}^{n} y_{ij} = 1 \qquad (9)$$

(one alternative may take only one position)

$$\mathop{\forall}_{i=1,...,n} \sum_{i=1}^{n} y_{ij} = 1 \qquad (10)$$

(in one position, there may be only one alternative).

We shall consider the problem on finding the Kemeny median on the set of preference vectors [26].

Let ($\mathbf{P}^1$, $\mathbf{P}^2$,..., $\mathbf{P}^K$) be the individual rankings defined by the experts, and $\mathbf{P}$ is an arbitrary ranking. We shall assign to it a n-dimensional vector $\Pi = (\pi_1, \pi_2, ..., \pi_n)$ whose i-th component is equal to the number of alternatives that are preferable than A_i. In this case, the Kemeny median is such ranking Π^* that:

$$\sum_{i=1}^{K} d\left(\Pi^*, \Pi^{(i)}\right) = \arg\min \sum_{i=1}^{K} d\left(\Pi, \Pi^{(i)}\right). \qquad (11)$$

It follows from (11):

$$\sum_{i=1}^{K} d(\Pi^*, \Pi^{(i)}) = \sum_{i=1}^{K} \sum_{j=1}^{n} \left| \Pi_i, \Pi_i^{(j)} \right|.$$

Let alternative A in ranking $\mathbf{P}$ be located at the j-th position. We introduce the loss matrix:

$$r_{ij} = \sum_{j=1}^{n} \left| \Pi_i, \Pi_i^{(j)} \right|.$$

Considering the rankings in which random alternative A_i, $i \in \{1,..., n\}$ is located sequentially from position 1 to n, it can be argued that the loss matrix r_{ij} characterizes the «disagreement» of experts with assigning the alternative A_i to the j-th position in the resulting ranking. Introduce variable:

$$x_{ij} = \begin{cases} 1, & \text{if the alternative } A_i \text{ is assigned to } j\text{-th place,} \\ 0, & \text{otherwise.} \end{cases}$$

Vector $\mathbf{X} = (x_{11}, x_{22},..., x_{nn})$ matches a certain ranking if and only if, when:

$$\sum_{j=1}^{n} x_{ij} = 1, \quad j \in \{1,...,n\}.$$

The Kemeny median is the ranking at which:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} x_{ij}$$

reaches a minimum. Thus, the problem on finding the Kemeny median can be stated in the form of the following integer optimization problem:

$$\arg\min_{\Pi} \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} x_{ij} \qquad (12)$$

with constraints

$$\sum_{j=1}^{n} x_{ij} = 1, \quad i \in \{1,..., n\},$$

$$\sum_{i=1}^{n} x_{ij} = 1, \quad j \in \{1,..., n\},$$

$$x_{ij} = \begin{cases} 0, \\ 1, \end{cases} \quad i \in \{1,..., n\},$$

that is, the assignment problem.

5. The assignment problem and algorithms for solving it

An analysis of methods for solving AP shows that many existing solving algorithms, despite the same theoretical computational complexity, demonstrate in practice different indicators of AP solving speed at different values for n. We have chosen for the computational experiment, described below, on estimating the speed of calculations four of the most popular AP solving algorithms:
– the Hungarian algorithm;
– the Mack's algorithm [28, 29];
– the AP solving algorithm as a particular case of the transportation problem with a zero supply and demand [30];
– an algorithm for finding the minimal matching in a bipartite weighted graph [31].
We shall represent AP in the following way. Let there be n resources (in the original statement – employees) that can

be utilized to perform n operations; in this case, exploiting the i-th employee for the j-th operation requires c_{ij} cost (i, $j = 1,..., n$). It is required to instruct each employee to perform a certain work to minimize total costs. Introduce variables:

$$x_{ij} = \begin{cases} 1, & \text{if } i\text{-th employee does } j\text{-th operation,} \\ 0, & \text{otherwise.} \end{cases}$$

Given that each employee performs one task and somebody does each work, we obtain the following problem of integer linear programming:

$$\arg\min \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \qquad (13)$$

with constraints

$$\sum_{i=1}^{n} x_{ij} = 1, \quad j = 1,...,n,$$

$$\sum_{j=1}^{n} x_{ij} = 1, \quad i = 1,...,n,$$

$$x_{ij} \in \{0,1\}, \quad i,j = 1,...,n.$$

The Hungarian algorithm for AP solving is based on the transformation of the cost matrix. A system of zero matrix elements, with the property that no two of them lie in the same line or column, is called the system of independent zeros (SIZ). The Hungarian method implies converting the original problem with matrix $\mathbf{C}$ to the equivalent minimization problem with matrix $\mathbf{D}$. If matrix $\mathbf{D}$ contains a system of n independent zeros, a solution to the problem then is the matrix $\mathbf{X}$, in which positions that correspond to independent zeros are taken by unities, while other elements are zero. Thus, the Hungarian algorithm aims to build SIZ through equivalent transformations of matrix $\mathbf{D}$. Basic equivalent transform is applied: if the elements of matrices $\mathbf{C}_{n \times n}$ and $\mathbf{D}_{n \times n}$ are related through equalities

$$d_{ij} = c_{ij} + \alpha_i + \beta_j (\alpha_i \geq 0, \beta_j \geq 0, i,j = 1,...,n),$$

the assignment problems are then equivalent with these matrices, that is, the sets of solutions (optimal points) in the matrices are the same. The original Kuhn's algorithm and its numerous modifications [32, 33] defines the sequence of steps that lead to forming the SIZ of assignment matrix.

The same transformations underlie the algorithm to solve AP proposed by Mack [28]. Paper [29] indicated that the Mack's algorithm largely focuses on program implementation, and, given this, it has a better computational efficiency.

The assignment problem is a particular case of the classic problem of linear programming – a transportation problem, in which the number of departure points equals the number of destinations, that is, a transportation table is shaped like a square. In addition, at each destination, the required volume is 1, and the proposed quantity at each production point is equal to 1. Therefore, the assignment problem can be solved employing the algorithm for solving a transportation problem, based on the method of potentials [30].

Another group of algorithms to solve AP are the algorithms based on the equivalence between AP and the problem on searching for a minimal matching in a weighted bipartite graph [31]. The vertices of the graph correspond

to the lines and columns of the cost matrix, and the edges have equal weights to the elements of the Kmatrix [32]. This problem can be effectively solved employing the Hopcroft-Karp algorithm [33]. Its computational complexity is $O(n^{2.5})$ [34].

6. Experimental study into computational efficiency of calculating the consensus median rankings as the assignment problem

To study the time required to compute consensus median rankings by Kemeny-Snell and Cook-Seiford using the methods to solve AP of various dimensionality n by different algorithms, we performed a computational experiment whose general scheme is shown in Fig 1. We investigated the dependence of time required to solve AP on dimensionality of n (number of alternatives). In order to solve the problem by employing different algorithms, we generated a representative statistical sample of pseudorandom sets of individual rankings for which we calculated the Kemeny-Snell and Cook-Seiford medians. When calculating, the median rankings were stated in the form of AP based on relations (8) and (12). Next, AP was solved by the above four algorithms:
– the Hungarian algorithm;
– the Mack's algorithm;
– the algorithm for solving a transportation problem using the method of potentials;
– the Hopcroft-Karp algorithm to search for a minimal matching in a weighted bipartite graph.

```
┌─────────────────────┐
│ Generation of set of│
│ individual rankings │
└─────────────────────┘
```

Collective ranking in the form of the Kemeny median

Collective ranking in the form of the Cook-Seiford median

Statement of the assignment problem based on relation (12)

Statement of the assignment problem based on relation (8)

Solving AP by the Hungarian method

Solving AP as the transportation problem

Solving AP by the Mack method

Solving AP by searching for a minimal matching in a weighted bipartite graph
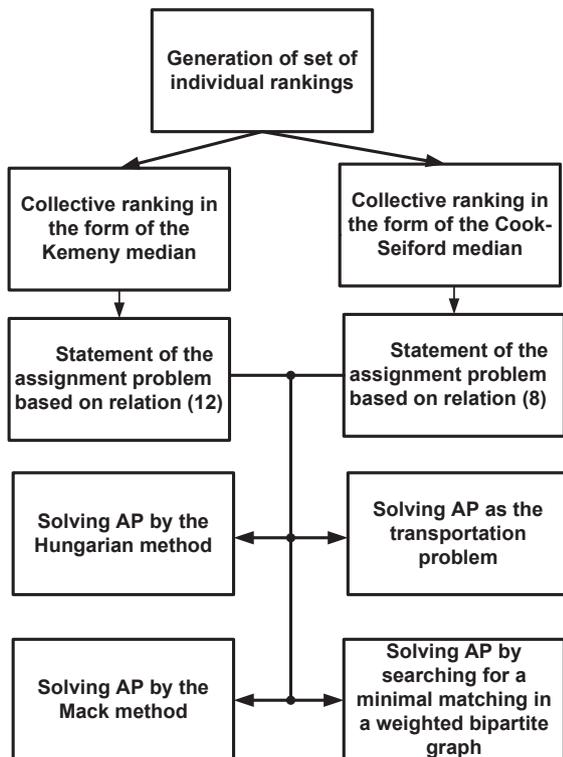
Fig. 1. Scheme of computational experiment

Each of the above algorithms was implemented in the software of computer algebra system Scilab. The experiment was conducted at the hardware-software platform: CPU Intel Core i5 2450M 2.5 Ghz; RAM 6 GB DDR3 1,300 MHz;

OS Win7. During calculations, we estimated the average count time for all algorithms. We shall describe in detail the generation of pseudo-random rankings. Paper [35] applied, in order to generate pseudo-random rankings, rather complicated algorithms based on ranking statistics [36]. The experiment reported here utilized the productive idea of generation, proposed in paper [37], based on which we formulated the following simplified algorithm for generating the rankings. To generate a synthesized matrix of collective rankings the size of $M \times N$, the following algorithm was applied. Standard operator *rand* generates M independent arrays of N pseudo-random numbers uniformly distributed in the range of [0,1]. The arrays are ordered line by line, the values for the elements of arrays are replaced with a number that these elements accept in the ordered row (that is, rank). As a result, in each line, there forms a pseudorandom sequence along the ranking scale the size of [1, N] with evenly distributed ranks. This sequence is considered as an individual ranking by expert number i, $i \in [1,M]$, and the entire generated ranking matrix as the set of individual rankings by M experts. The following values were selected for the dimensionality of the problem (the power of the set of alternatives) $n = \{5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$. For clarity, the number of experts K in each synthesized matrix was taken to be equal to $2n$. The sample size of each dimensionality is 150 elements.

Results of the experiment are shown in Fig. 2.

Papers [36, 38] that addressed the calculation of median consensus rankings using a branch and bound method and the decomposition combinatorial techniques, highlighted the important fact. It was established that the time required to solve the problem on the search for a consensus ranking significantly depends on the coherence degree of individual expert rankings.

This coherence is estimated based on the Kendall's concordance coefficient [26, 39], which shall be calculated from formula:

$$W = \frac{12S}{K^2(n^3 - n)}, \tag{14}$$

where S is the variance estimate; K is the number of experts; n is the number of alternatives.

Estimation of variance:

$$S = \sum_{i=1}^{n}\left(\sum_{k=1}^{K} rnk_{ik} - rnk_0\right)^2, \tag{15}$$

where k is the index of an expert ($k = 1, 2,..., K$), rnk_{ik} is the rank value, assigned by the k-th expert to the i-th alternative; rnk_0 is the estimate of the mean value of ranks based on alternatives, derived from:

$$rnk_0 = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K} rnk_{ik}.$$

The concordance coefficient is equal to 1 if all experts' rankings in a group are the same, and is equal to 0 if all rankings are different. According to [40], the coherence of individual expert rankings is considered to be weak if $W \leq 0.3$, moderate if $0.3 \leq W < 0.5$, strong if $W > 0.5$, respectively. To verify the dependence of calculation time of median rankings as AP on the coherence degree of individual rankings by experts, we performed a separate computational experiment.
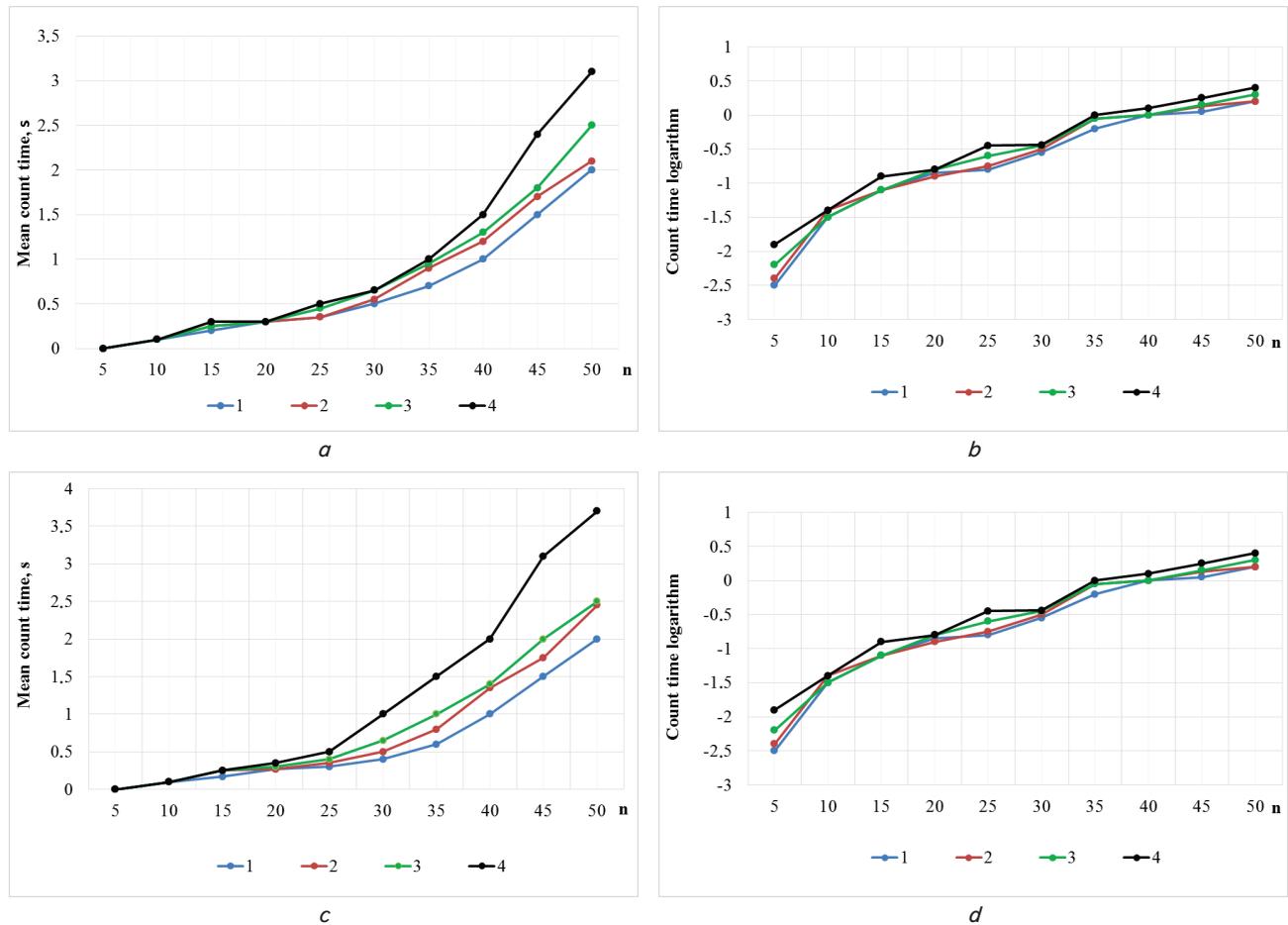
Fig. 2. Dependence of mean count time on dimensionality of the problem for different algorithms to solve AP:
a − Cook-Seiford median, *c* − Kemeny-Snell median, *b, d* − same dependences with a timeline based on the logarithmic scale
(1 − Mack's algorithm, 2 − Hungarian algorithm, 3 − Hopcroft-Karp algorithm to search for a minimal matching in a weighted
bipartite graph, 4 − algorithm for solving a transportation problem using the method of potentials)

For the sets of individual expert rankings, synthesized by the above algorithm, upon generating each matrix, we estimated its coherence degree based on a concordance coefficient by testing its statistical significance for criterion χ^2. At $W<0.3$, the generated matrix was assigned to sample «1 − Weak coherence»; at $0,3 \leq W < 0,5$ – to sample «2 − Moderate coherence»; at $W \geq 0.5$ – to sample «3 − Strong coherence». Each sample was generated to a volume of 150 matrices for the above set n – 50 matrices of each dimensionality and three levels of coherence. Next, we calculated the Kemeny-Snell and Cook-Seiford medians as AP in the formed samples by the four algorithms described above, and the estimated the mean calculation time. Table 1 gives an example of the summary results of the experiment, for calculation based on the Hungarian algorithm.

We applied to the constructed tables the non-parametric Friedman-Babington-Smith test (hereinafter abbreviated as the Friedman test) – a known ranking nonparametric statistics [41]. The Friedman test null hypothesis H_0 is considered to be the following: «differences between the values for rows in the table are random». The alternative hypothesis is H_1 – «there are significant differences in the rows of the non-random character». The confidence level for the Friedman test was accepted equal to 0.95. For two of the four tables, at the required significance level, the Friedman statistics do not yield a definite answer. We applied, in order to analyze these two tables, more powerful non-parametric statistics – the Anderson-Kahneman-Schach test [42] and the Dana Quade test [43]. By using these tests, we obtained for all four tables an unambiguous result with the above-specified confidence probability – the differences between the table's rows are random.

Consider the results shown in Fig. 2. At a moderate number of alternatives ($n<50$), calculating the both median rankings using AP requires low computational cost that are well within the real time scale (1−2 s), regardless of the algorithm employed. The charts of the count time, given in

Table 1

**Mean calculation time of the Kemeny-Snell median (ms) by the Hungarian method
for different degrees of coherence of individual expert rankings**

Coherence degree	Problem dimensionality									
	$N=5$	$N=10$	$N=15$	$N=20$	$N=25$	$N=30$	$N=35$	$N=40$	$N=45$	$N=50$
Weak	11.7	42.0	85.3	152.3	272.1	426.3	531.7	925.3	1,420.3	1,850.2
Moderate	17.3	37.4	93.9	156.9	293.2	441.1	523.2	996.0	1,355.1	2,009.7
Strong	15.3	47.4	86.7	147.5	307.5	412.8	545.2	1,014.5	1,483.5	1,943.5

a logarithmic scale, really demonstrate the computational complexity of the problem, not exceeding $O(n^3)$. The shape of dependences corresponds to the power function logarithm. Computational costs for calculating the Kemeny-Snell and Cook-Seiford medians almost do not differ. At a small number of alternatives ($n<15$), all algorithms require a roughly equal time. With the growing number of alternatives, the algorithm for solving AP as the transportation problem runs noticeably slower than the rest. The highest performance is demonstrated by the Mack's, although its high-speed performance is almost comparable to the Hungarian algorithm and the Hopcroft-Karp algorithm.

Results of the second part of the computational experiment demonstrate another important result. When calculating the median consensus rankings as AP, the computational complexity does not depend on coherence degree of individual expert rankings. Given this, the proposed procedure differs favorably from the solution to the problem on collective ranking applying a branch and bound method. At such a solution, at a low coherence of individual expert opinions, the time required to solve the problem increases.

7. Discussion of results of estimating the speed of computation of median consensus rankings

In the process of study, we showed the computational efficiency of applying the assignment problem for calculating collective consensus rankings in the form of the Kemeny-Snell and Cook-Seiford medians.

Numerical experiment has revealed that the three practically investigated algorithms for solving the assignment problem out of four are practically equivalent in terms of high-speed performance; however, with the increasing number of alternatives, the Mack's algorithm demonstrates the highest speed performance. The calculation time based on a medium power platform using the system Scilab for the number of alternatives $n<50$ employing the Mack's algorithm does not exceed 2 s.

This result is useful for applying in collective systems for expert estimation of real time. Specifically, such systems are the network expertise systems.

The proposed method for the calculation of median rankings has an important positive quality. It is established that the calculation time of consensus rankings, in contrast to other methods for reducing the computational complexity of a problem, does not depend on coherence degree of individual expert rankings. This favorably distinguishes it from, for example, solving a problem on medians using a branch and bound method and by combinatorial methods. This result is important for the application of the method in network expert systems with a random selection of the contingent of experts.

The result of the computational experiment is also useful for practical application of the proposed procedure at network expert systems. In network expertise systems, due to a random gathering of the team of experts, there is a possi-

bility for a low coherence level of individual rankings. In such cases, the application of the proposed method does not affect duration of the calculation.

The reported results of the study are not exhaustive. Specifically, during expert evaluation, there is a possible situation when experts assign to different alternatives the same ranks: this is the situation of so-called tied ranks [26]. In the presence of associated ranks, the ratios, given in our work, should be improved or adjusted. Study into computational efficiency of the proposed method in the presence of associated ranks in the individual experts' rankings is the subject of further research.

8. Conclusions

1. Calculation of a collective expert opinion based on individual experts' rankings has two axiomatically substantiated variants – the Kemeny-Snell median and the Cook-Seiford median. Since both options are the problems of the non-polynomial computational complexity, in order to decrease labor intensity of the problem, we suggested reducing the median rankings to the problem of integer programming – the assignment problem. In contrast to many approximated and heuristic methods for solving the problem on median rankings, reducing median rankings to AP retains the original axiomatics of Kemeny-Snell and is, therefore, the precise solution to the original problem. In this case, the computational complexity of solution does not exceed $O(n^3)$, which produces a significant gain in time during computation without violating the strict axiomatics of the problem.

2. The assignment problem to which the Kemeny-Snell and Cook-Seiford medians are reduced, has been sufficiently explored. A number of algorithms to solve the assignment problem exist. The most popular are the Mack's algorithm, the Hungarian algorithm, the Hopcroft-Karp algorithm. All these algorithms yield a polynomial computational complexity with a power not exceeding 3. The algorithms were implemented in software and investigated experimentally.

3. The numerical experiment showed that the use of algorithms to solve the assignment problem when evaluating the median consensus rankings makes it possible to fast compute the ranking medians. For the number of alternatives not exceeding 50, at a typical computing platform, the count time does not exceed 2–3 s, that is, the computation is executed almost in real time. It was also found that the median calculation time in this case, in contrast to other methods, does not depend on coherence degree of individual experts' rankings.

Results of the conducted research allow us to recommend the proposed rapid computational procedure for the search of median consensus rankings by Kemeny-Snell and Cook-Seiford for practical application in the network examination systems. The investigated computational procedure combines the axiomatic accuracy, computational efficiency, and independence from the coherence degree of individual experts' rankings.

References

1. Dopazo E., Martínez-Céspedes M. L. Rank aggregation methods dealing with ordinal uncertain preferences // Expert Systems with Applications. 2017. Vol. 78. P. 103–109. doi: https://doi.org/10.1016/j.eswa.2017.01.051

2. Raspoznavanie informacionnyh operaciy / Dodonov A. G., Lande D. V., Cyganok V. V., Andreychuk O. V., Kadenko S. V., Grayvoronskaya A. N. Kyiv: OOO «Inzhiniring», 2017. 282 p.

3. E-Expertise: modern collective intelligence / Gubanov D., Korgin N., Novikov D., Raikov A. Springer, 2014. 147 p. doi: https://doi.org/10.1007/978-3-319-06770-4

4. Finding experts using internet-based discussions in online communities and associated social networks / Breslin J., Bojars U., Aleman-Meza B., Boley H., Mochol M., Nixon L. J. B. et. al. // Proceedings of the 1st International Expert Finder Workshop at Knowledge Web General Assembly. 2007. P. 38–47.

5. Finding experts in online forums for enhancing knowledge sharing and accessibility / Wei C.-P., Lin W.-B., Chen H.-C., An W.-Y., Yeh W.-C. // Computers in Human Behavior. 2015. Vol. 51. P. 325–335. doi: https://doi.org/10.1016/j.chb.2015.04.055

6. Kemeny J. G., Snell L. J. Preference ranking: an axiomatric approach, in: Mathematical Models in the Social Sciences. Ginn and Company, New York, 1962. P. 9–23.

7. Cook W. D., Seiford L. M. Priority Ranking and Consensus Formation // Management Science. 1978. Vol. 24, Issue 16. P. 1721–1732. doi: https://doi.org/10.1287/mnsc.24.16.1721

8. Armstrong, R. D., Cook, W. D., Seiford L. M. Priority Ranking and Consensus Formation: The Case of Ties // Management Science. 1982. Vol. 28, Issue 6. P. 638–645. doi: https://doi.org/10.1287/mnsc.28.6.638

9. Burkard R. Assignment Problems. Society for Industrial and Applied Mathematics. Philadelphia, 2009. 382 p.

10. Procaccia A. D. Computational social choice // XRDS: Crossroads, The ACM Magazine for Students. 2011. Vol. 18, Issue 2. P. 31. doi: https://doi.org/10.1145/2043236.2043249

11. Moulin H. The strategy of social choice. Vol. 18. Advanced Textbooks in Economics. Elsevier, 1983. 226 p.

12. Heckelman J., Miller N. Handbook of social choice and voting. Edward Elgar Publishing, 2015. 424 p. doi: https://doi.org/10.4337/9781783470730

13. Bartholdi J., Tovey C. A., Trick M. A. Voting schemes for which it can be difficult to tell who won the election // Social Choice and Welfare. 1989. Vol. 6, Issue 2. P. 157–165. doi: https://doi.org/10.1007/bf00303169

14. Bury H., Wagner D. Group judgement with ties. Distance-based methods // New Approaches in Automation and Robotics. 2008. P. 153–172. doi: https://doi.org/10.5772/5393

15. Ali A., Meilă M. Experiments with Kemeny ranking: What works when? // Mathematical Social Sciences. 2012. Vol. 64, Issue 1. P. 28–40. doi: https://doi.org/10.1016/j.mathsocsci.2011.08.008

16. A differential evolution algorithm for finding the median ranking under the Kemeny axiomatic approach / D'Ambrosio A., Mazzeo G., Iorio C., Siciliano R. // Computers & Operations Research. 2017. Vol. 82. P. 126–138. doi: https://doi.org/10.1016/j.cor.2017.01.017

17. Aledo J. A., Gámez J. A., Molina D. Tackling the rank aggregation problem with evolutionary algorithms // Applied Mathematics and Computation. 2013. Vol. 222. P. 632–644. doi: https://doi.org/10.1016/j.amc.2013.07.081

18. Jackson B. N., Schnable P. S., Aluru S. Consensus Genetic Maps as Median Orders from Inconsistent Sources // IEEE/ACM Transactions on Computational Biology and Bioinformatics. 2008. Vol. 5, Issue 2. P. 161–171. doi: https://doi.org/10.1109/tcbb.2007.70221

19. D'Ambrosio A., Amodio S., Iorio C. Two algorithms for finding optimal solutions of the Kemeny rank aggregation problem for full rankings // Electronic Journal of Applied Statistical Analysis. 2015. Vol. 8, Issue 2. P. 198–213. doi: https://doi.org/10.1285/i20705948v8n2p198

20. Fixed-parameter algorithms for Kemeny rankings / Betzler N., Fellows M. R., Guo J., Niedermeier R., Rosamond F. A. // Theoretical Computer Science. 2009. Vol. 410, Issue 45. P. 4554–4570. doi: https://doi.org/10.1016/j.tcs.2009.08.033

21. Aggregation of Partial Rankings – An Approach Based on the Kemeny Ranking Problem / Nápoles G., Dikopoulou Z., Papageorgiou E., Bello R., Vanhoof K. // Lecture Notes in Computer Science. 2015. P. 343–355. doi: https://doi.org/10.1007/978-3-319-19222-2_29

22. Kuhn H. W. The Hungarian method for the assignment problem // Naval Research Logistics Quarterly. 1955. Vol. 2, Issue 1-2. P. 83–97. doi: https://doi.org/10.1002/nav.3800020109

23. Pentico D. W. Assignment problems: A golden anniversary survey // European Journal of Operational Research. 2007. Vol. 176, Issue 2. P. 774–793. doi: https://doi.org/10.1016/j.ejor.2005.09.014

24. Munkres J. Algorithms for the Assignment and Transportation Problems // Journal of the Society for Industrial and Applied Mathematics. 1957. Vol. 5, Issue 1. P. 32–38. doi: https://doi.org/10.1137/0105003

25. Lelyakova L. V., Haritonova A. G., Chernyshova G. D. Prikladnye zadachi o naznacheniyah (modeli, algoritmy resheniya) // Vestnik VGU, Seriya: Sistemniy analiz i informacionnye tekhnologii. 2017. Issue 2. P. 22–27.

26. Samohvalov Yu. Ya., Naumenko E. M. Ekspertnoe ocenivanie. Metodicheskiy aspekt: monografiya. Kyiv: DUIKT, 2007. 262 p.

27. Boltenkov V. A., Kuvaieva V. I., Pozniak A. V. Analiz mediannyh metodov konsensusnogo agregirovaniya rangovyh predpochteniy // Informatyka ta matematychni metody v modeliuvanni. 2017. Vol. 7, Issue 4. P. 307–317.

28. Mack C. The Bradford method for the assignment problem // New Journal of Statistics and Operational Research. 1969. Vol. 1. P. 17–29.

29. Jonker R. Teaching linear assignment by Mack's algorithm // Twenty-Five Years of Operations Research in the Netherlands: Papers Dedicated to Gijs de Leve, CWI Tract. 1989. Vol. 70. P. 54–60.

30. Eddous M., Stensfild R. Metody prinyatiya resheniy / I. I. Eliseeva (Ed.). Moscow: Audit, YuNITI, 1997. 590 p.

31. Ramshaw L., Tarjan R. E. On Minimum-Cost Assignments in Unbalanced Bipartite Graphs // HP Laboratories, HPL-2012-40R1. 2012. 91 p.

32. Jonker R., Volgenant A. A shortest augmenting path algorithm for dense and sparse linear assignment problems // Computing. 1987. Vol. 38, Issue 4. P. 325–340. doi: https://doi.org/10.1007/bf02278710

33. A complete assignment algorithm and its application in constraint declarative languages / Brauner N., Echahed R., Finke G., Gregor H., Prost F. // Cahier du laboratoire Leibniz No. 111. 2004. 60 p.

34. Matching Algorithms Are Fast in Sparse Random Graphs / Bast H., Mehlhorn K., Schafer G., Tamaki H. // Theory of Computing Systems. 2005. Vol. 39, Issue 1. P. 3–14. doi: https://doi.org/10.1007/s00224-005-1254-y

35. Amodio S., D'Ambrosio A., Siciliano R. Accurate algorithms for identifying the median ranking when dealing with weak and partial rankings under the Kemeny axiomatic approach // European Journal of Operational Research. 2016. Vol. 249, Issue 2. P. 667–676. doi: https://doi.org/10.1016/j.ejor.2015.08.048

36. Critchlow D. E., Fligner M. A., Verducci J. S. Probability models on rankings // Journal of Mathematical Psychology. 1991. Vol. 35, Issue 3. P. 294–318. doi: https://doi.org/10.1016/0022-2496(91)90050-4

37. Ocenka vzaimosvyazi tekhniko-ekonomicheskih pokazateley ob'ektov EES / Farhadzade E. M., Muradaliev A. Z., Farzaliev Yu. Z., Rafieva T. K., Abdullaeva S. A. // Elektronnoe modelirovanie. 2017. Vol. 39, Issue 6. P. 93–106.

38. Antosiak P. P. Dekompozytsiyni protsedury u zadachi znakhodzhennia strohoho rezultuiuchoho ranzhuvannia obiektiv u vyhliadi mediany Kemeni-Snella // Naukovyi visnyk Uzhhorodskoho universytetu. 2008. Issue 17. P. 29–36.

39. Sprent P., Smeeton N. C. Applied nonparametric statistical methods. CRC Press Taylor & Francis Group, 2007. 544 p.

40. Kraska-Miller M. Nonparametric statistics for social and behavioral sciences. CRC Press, 2013. 260 p. doi: https://doi.org/10.1201/b16188

41. Kobzar' A. I. Prikladnaya matematicheskaya statistika. Dlya inzhenerov i nauchnyh rabotnikov. Moscow: FIZMATLIT, 2006. 816 p.

42. Schach S. An Alternative to the Friedman Test with Certain Optimality Properties // The Annals of Statistics. 1979. Vol. 7, Issue 3. P. 537–550. doi: https://doi.org/10.1214/aos/1176344675

43. Quade D. Using Weighted Rankings in the Analysis of Complete Blocks with Additive Block Effects // Journal of the American Statistical Association. 1979. Vol. 74, Issue 367. P. 680. doi: https://doi.org/10.2307/2286991