A new coefficient of rankings similarity in decision-making problems

Wojciech Sałabun[1][0000−0001−7076−2519] and Karol Urbaniak[1][0000−0002−5373−9760]

[1]Research Team on Intelligent Decision Support Systems,
Department of Artificial Intelligence and Applied Mathematics,
Faculty of Computer Science and Information Technology,
West Pomeranian University of Technology in Szczecin,
ul. Żołnierska 49, 71-210 Szczecin, Poland
wojciech.salabun@zut.edu.pl

Abstract. Multi-criteria decision-making methods are tools that facilitate and help to make better and more responsible decisions. Their main objective is usually to establish a ranking of alternatives, where the best solution is in the first place and the worst in the last place. However, using different techniques to solve the same decisional problem may result in rankings that are not the same. How can we test their similarity? For this purpose, scientists most often use different correlation measures, which unfortunately do not fully meet their objective.
In this paper, we identify the shortcomings of currently used coefficients to measure the similarity of two rankings in decision-making problems. Afterward, we present a new coefficient that is much better suited to compare the reference ranking and the tested rankings. In our proposal, positions at the top of the ranking have a more significant impact on the similarity than those further away, which is right in the decision-making domain. Finally, we show a set of numerical examples, where this new coefficient is presented as an efficient tool to compare rankings in the decision-making field.

Keywords: Decision analysis · decision making · decision theory · measurment uncertainty · ranking .

1 Introduction

Decision making is an integral part of human life. Every day, every person is faced with different kinds of decision-making problems, which can affect both professional and private life. An example of a decision-making problem can be a change of legal regulations in the state, choice of university, purchase of a new car, determination of the amount of personal income tax, selection of a suitable location for the construction of a nuclear power plant, adoption of a plan of research, or the sale or purchase of stock exchange shares.

In the majority of cases, decision-making problems are based on many, often contradictory, decision-making criteria. Therefore multi-criteria decision-analysis

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

(MCDA) methods and decision support systems enjoy deep interest both in the world of business and science. Almost in every case, a reliable decision requires the analysis of many alternatives. Each of them should be assessed from the perspective of all the criteria characterizing its acceptability. As the complexity of the problem increases, it becomes more and more challenging to make the optimal decision. An additional complication is that there is no complete mathematical form depending on the criteria and the expected consequences. In particularly important problems, the role of the decision-maker is entrusted to an expert in a given field or to a group of experts who will help to identify the best solution. We talk about individual or group decision making, respectively. Even then, it can often be problematic for an individual expert as well as for collegiate bodies to determine the right decision. In this case, MCDA methods can be helpful.

MCDA methods are great tools to support the decision-maker in the decision-making process. We can identify two main groups of MCDA methods, i.e., American and European schools [33]. Methods of the American school of decision support are based on the utility or value function [5, 16]. The most important methods belonging to this family are: analytic hierarchy process (AHP) [34], analytic network process (ANP) [35], utility theory additive (UTA) [21], simple multi-attribute rating technique (SMART) [28], technique for order preference by similarity to ideal solution (TOPSIS) [3, 32], or measuring attractiveness by a categorical based evaluation technique (MACBETH) [2]. Methods of European school of decision support use outranking relation in the preference aggregation process, where the most popular are ELECTRE family [1, 36] and PROMETHEE methods [8, 15]. Additionally, we can indicate the set of techniques based strictly on the rules of decision making. These methods use the fuzzy sets theory (COMET) [12, 25–27] and the rough set theory (DRSA) [29].

Generally, MCDA methods help to create a ranking of decision variants where the most preferred alternative comes first [4]. The problem arises when we use more than one MCDA method, and the rankings obtained are not identical. Then the question arises on how to compare the received rankings? Currently, the most popular approach is the analysis based on the correlation between the two or more rankings [7, 12, 19, 24]. However, we are going to show that this analysis is insufficient in the decision support domain. An appropriate approach should ensure that a better ranking in terms of the order can be identified. Then, with a proper benchmark, it would be possible to assess the correctness of the MCDA methods in terms of the rankings generated [22].

In this paper, we identify the shortcomings of currently used coefficients to measure the similarity of two rankings. The most significant contribution is the WS coefficient, which depends strictly on the position on which the difference in the ranking occurred. Afterward, three linguistic terms are identified by using trapezoidal fuzzy numbers, i.e., low, medium, and high similarity. We compare the proposed coefficient with ρ Spearman, τ Kendall, and γ Goodman-Kruskal coefficients, which are commonly used to measure rankings similarity in MCDA problems [9, 12, 17, 18, 23]. In addition, the proposed approach is compared with

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

the similar coefficients presented in [6, 10, 13]. For this purpose, numerical experiments are discussed.

The rest of the paper is organized as follows: In section 2, some basic preliminary concepts are discussed. Section 3 introduces a new coefficient of rankings similarity in the decision-making problems. In Section 4, the practical feasibility study of the WS coefficient is discussed. In Section 5, we present the summary and conclusions.

2 Preliminaries

An important issue is how to compare the correctness of the order of the two rankings. The simplest method is to check whether the rankings are consistent or inconsistent. Such an approach is not sufficient and can be used almost exclusively for 2 or 3 elementary rankings [27]. The much more common approach is to use one of the coefficients of monotonous dependence of two variables, where the obtained rankings for a set of considered alternatives are our variables. The most commonly used symmetrical coefficient of such dependence is the Spearman's coefficient [9, 18, 17, 23], which is expressed by the following formula (1):

$$r_s = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)} \tag{1}$$

where d_i is defined as the difference between the ranks $d_i = R_{xi} - R_{yi}$ and n is the number of elements in the ranking. The Spearman's coefficient is interpreted as a percentage of the rank variance of one variable, which is explained by the other variable [31].

The most frequently used asymmetrical monotonous coefficients of two variables are Kendall [12, 20] and Goodman-Kruskal coefficients [12, 14]. They are expressed in formulas (2) and (3) respectively:

$$\tau = 2 \cdot \frac{N_s - N_d}{n \cdot (n - 1)} \tag{2}$$

$$G = \frac{N_s - N_d}{N_s + N_d} \tag{3}$$

where N_s is the number of compatible pairs, N_d is the number of non-compliant pairs, and n is the number of all pairs. The Kendall and Goodman-Kruskal coefficients, unlike Spearman, are interpreted in terms of probability. They represent the difference between the probability that the compared variables will be in the same order for both variables and the probability that they will be in the opposite order.

The presented coefficients are the most frequently used measures of the analysis of the rankings similarity in decision-making problems [9, 12, 17, 18, 23]. However, we want to indicate a significant shortcoming, which is related to the place of difference occurrence. The idea of measuring the rankings similarity is not new and has been the subject of many works [11, 30]. Particularly interesting

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

in the context of the presented approach are works related to Blest's measure of rank correlation v and the weighted rank measure of correlation r_w [6, 10, 13]. They are expressed in formulas (4) and (5) respectively:

$$r_w = 1 - \frac{6\sum_{i=1}^{n}\left(R_{xi} - R_{yi}\right)^2\left((n - R_{xi} + 1) + (n - R_{yi} + 1)\right)}{n^4 + n^3 - n^2 - n} \tag{4}$$

$$v = 1 - \frac{12\sum_{i=1}^{n}(n + 1 - R_{xi})^2 \cdot R_{yi} - n(n+1)^2(n+2)}{n(n+1)^2(n-1)} \tag{5}$$

The presented coefficients (1-3) are regardless of whether the error occurs at the top or bottom; the values of the factors will be identical. In Table 1, the simple example shows five rankings, including one reference (R_x) and four test rankings ($R_y^{(1)} - R_y^{(4)}$). The test rankings were created by a change in the correct ranking of the two adjacent alternatives. We want to remind that the rankings are determined to choose the best possible solution, and the value of the preferences decreases with each position in the ranking. The difference at the top should be more significant than an error at the bottom of the ranking. The exchange of alternative locations from the first and second position is a more considerable error than the swap of the second and third position. However, the values of the coefficients indicate that similarity of the test rankings to the reference ranking is the same for all test sets.

3 WS coefficient of rankings similarity

The new ranking similarity factor should be resistant to the situation described in the previous section, and at the same time, should be sensitive to significant changes in the ranking. Besides, this factor should be easy to interpret, and its values should be limited to a specific interval.

We assumed that the new indicator should be strongly related to the difference between two rankings on particular positions. An additional assumption is that the top has a more significant influence on similarity than the bottom of the ranking. Based on these assumptions, a new indicator was developed, which can be presented as (6):

$$WS = 1 - \sum_{i=1}^{n}\left(2^{-R_{xi}} \cdot \frac{|R_{xi} - R_{yi}|}{max\{|1 - R_{xi}|, |N - R_{xi}|\}}\right) \tag{6}$$

where WS is a value of similarity coefficient, N is a length of ranking, R_{xi} and R_{yi} mean the place in the ranking for $i - th$ element in respectively ranking x and ranking y.

The proof of convergence for the WS factor is quite simple. The formula (6) can be divided into two main components. The first one (7) is responsible for making the WS value dependent on the position in the reference ranking (R_x).

$$2^{-R_{xi}} \tag{7}$$

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

Table 1. Summary of the test with reference ranking (Rx) and four test rankings ($R_y^{(1)} - R_y^{(4)}$) with the calculated correlation factors and proposed WS coefficient for the set of five alternatives ($A_1 - A_5$), each having a different position in the ranking.

A_i	R_x	$R_y^{(1)}$	$R_y^{(2)}$	$R_y^{(3)}$	$R_y^{(4)}$
A_1	1	2	1	1	1
A_2	2	1	3	2	2
A_3	3	3	2	4	3
A_4	4	4	4	3	5
A_5	5	5	5	5	4
coefficients	r_s	0.9000	0.9000	0.9000	0.9000
	τ	0.8000	0.8000	0.8000	0.8000
	G	0.8000	0.8000	0.8000	0.8000
	r_w	0.8500	0.8833	0.9167	0.9500
	v	0.8500	0.8833	0.9167	0.9500
	WS	0.7917	0.8542	0.9167	0.9714

We are dealing with a geometric series which is convergent. As proof, we can calculate a trivial limit.

$$\lim_{n\to\infty} \sum_{i=1}^{n} (2)^{-R_{xi}} = 1 \tag{8}$$

The second component (9) determines to what extent the difference in rankings affects the similarity of rankings. This value can be obtained from zero (the positions are identical) to one.

$$\frac{|R_{xi} - R_{yi}|}{max\{|1 - R_{xi}|, |N - R_{xi}|\}} \tag{9}$$

If we multiply the (7) by (9) then this series cannot be higher than one. Therefore, it is clear that the WS coefficient can only take values from zero to one. We can compare all coefficients for a simple example in Table 1. The WS, r_w, and v coefficients take into account the position of the error occurrence, and the rest of them remain the same regardless of where the error occurs. In the next section, other tests comparing the performance of the indicators will be presented and discussed.

4 Results and Discussion

4.1 Analysis of five-element rankings

The first experiment here presents tied ranks, i.e., the same values in the ranking. It happens when two alternatives get the same place. For example, if two decision variants receive the first place together, the ranking will contain a value of 1.5 for

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

Table 2. Summary of the test with reference ranking (Rx) and four test rankings ($R_y^{(1)} - R_y^{(4)}$) with the calculated correlation factors and proposed WS coefficient for the set of five alternatives ($A_1 - A_5$), where one pair has the same position in the ranking.

A_i	R_x	$R_y^{(1)}$	$R_y^{(2)}$	$R_y^{(3)}$	$R_y^{(4)}$
A_1	1	1.5	1	1	1
A_2	2	1.5	2.5	2	2
A_3	3	3	2.5	3.5	3
A_4	4	4	4	3.5	4.5
A_5	5	5	5	5	4.5
coefficients	r_s	0.9747	0.9747	0.9747	0.9747
	τ	0.9487	0.9487	0.9487	0.9487
	G	1.0000	1.0000	1.0000	1.0000
	r_w	0.9625	0.9708	0.9792	0.9875
	v	0.9250	0.9417	0.9583	0.9750
	WS	0.8958	0.9271	0.9583	0.9857

both (the average of their positions). Table 2 shows the results of calculations for the five-element ranking, where the different location of tied pairs is considered.

One more again, WS, r_w, and v coefficients show the change of value together with the change of position on which there are the tied pairs. It is a property that was identified as a significant drawback of the currently used methods, i.e., ρ Spearman, τ Kendall, and γ Goodman-Kruskal. Ranking $R_y^{(4)}$ is more similar than $R_y^{(1)}$ because full correctness occurs on the first three positions and not on the last three.

Another simple experiment consists in creating test rankings, where successive rankings differ from the base ranking by the alternative indicated as the best. The results for the five-element ranking are shown in Table 3. Replacing the best option with the worst in all coefficients results in a negative value result (except WS). It means a negative correlation, which is not trivial to interpret in decision-making problems. Besides, interesting is the case of the $R_y^{(3)}$ ranking because it obtained a total lack of correlation (for τ and G coefficients). It means that the order of the base and second rankings is utterly unrelated to each other. It is a confirmation that these classical rank coefficients do not examine the similarity of the two rankings thoroughly. In general, all coefficients assess test rankings against the base ranking in a somewhat similar way. The rationale for this is that the three positions in the ranking have been indicated flawlessly.

The last example in this subsection examines the r_w coefficients in two cases, i.e., for test rankings $R_y^{(1)} - -R_y^{(2)}$ and $R_y^{(3)} - -R_y^{(4)}$). Once again, the R_x ranking is used as a reference point. The detailed results are presented in Table 4.

Rankings $R_y^{(1)}$ and $R_y^{(2)}$ have equal values for most of the coefficients, where only WS and v are exceptions. Ranking $R_y^{(1)}$ is significantly better than ranking

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

Table 3. Summary of the test with reference ranking (Rx) and four test rankings ($R_y^{(1)} - R_y^{(4)}$) with the calculated correlation factors and proposed WS coefficient for the set of five alternatives ($A_1 - A_5$), where each ranking has a different position error.

A_i	R_x	$R_y^{(1)}$	$R_y^{(2)}$	$R_y^{(3)}$	$R_y^{(4)}$
A_1	1	2	3	4	5
A_2	2	1	2	2	2
A_3	3	3	1	3	3
A_4	4	4	4	1	4
A_5	5	5	5	5	1
coefficients	r_s	0.9000	0.6000	0.1000	-0.6000
	τ	0.8000	0.4000	0.0000	-0.4000
	G	0.8000	0.4000	0.0000	-0.4000
	r_w	0.8500	0.4667	-0.0500	-0.6000
	v	0.8500	0.4667	-0.0500	-0.6000
	WS	0.7917	0.6250	0.5625	0.4688

$R_y^{(2)}$. Even though the A_5 alternative has been identified as the best in $R_y^{(1)}$. The rest of this ranking has been correctly identified according to the right order. However, in the $R_y^{(2)}$, the best alternative was wrongly rated as the worst. Therefore, the best alternative (A_1) has a chance to be chosen in the first case and not in the second. These rankings cannot be evaluated as being the same. This shows the superiority of WS and v coefficients in decision-making ranks analysis. Rankings $R_y^{(3)}$ and $R_y^{(4)}$ show greater variability of coefficients, i.e., the ranking $R_y^{(3)}$ has a coefficient value less, equal to or greater than the ranking $R_y^{(4)}$. It all depends on which coefficient is taken into account, but WS and v again point to the superiority of the ranking $R_y^{(3)}$.

4.2 Influence of a ranking size on coefficients

In this subsection, we want to indicate the impact of the ranking size on the achieved value of the indicator. Figure 1 shows comparisons of WS, r_s, r_w, and v coefficients. Only alternatives on the first and second positions have been replaced. It is a consequence of the conclusion drawn from Table 1. We take into account the rankings with the number from 5 to 50 elements. We can observe that with the increased ranking size, the similarity with the assumed assumptions increases. The Ws coefficient is characterized by the greatest variability, depending on the size of the ranking. Figure 2 shows the changes in the WS value when replacing the best elements with the second, third, fourth, and fifth ones (the position of one adjacent pair is swapped). As we can see, the WS values decrease accordingly, as the quality of the rankings decreases as the best solution moves away from the top of the ranking.

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

Table 4. Summary of the test with reference ranking (Rx) and four test rankings ($R_y^{(1)} - R_y^{(4)}$) with the calculated correlation factors and proposed WS coefficient for the set of five alternatives ($A_1 - A_5$), where the change of coefficients is investigated

A_i	R_x	$R_y^{(1)}$	$R_y^{(2)}$	$R_y^{(3)}$	$R_y^{(4)}$
A_1	1	2	5	2	4
A_2	2	3	1	3	2.5
A_3	3	4	2	5	1
A_4	4	5	3	4	5
A_5	5	1	4	1	2.5
coefficients	r_s	0.0000	0.0000	-0.1000	-0.0513
	τ	0.2000	0.2000	0.0000	-0.1054
	G	0.2000	0.2000	0.0000	-0.1111
	r_w	0.0000	0.0000	-0.0667	-0.0667
	v	0.1667	-0.1667	0.0833	-0.1083
	WS	0.6771	0.3225	0.6354	0.4180

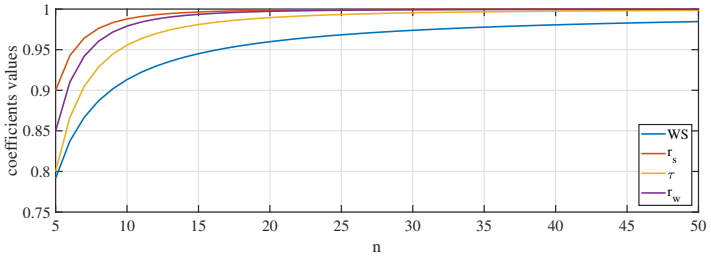

Fig. 1. The value of the coefficients depending on the length of the ranking (n), where occurs one error (change of the first and second position in the ranking) and the converted positions in the ranking.

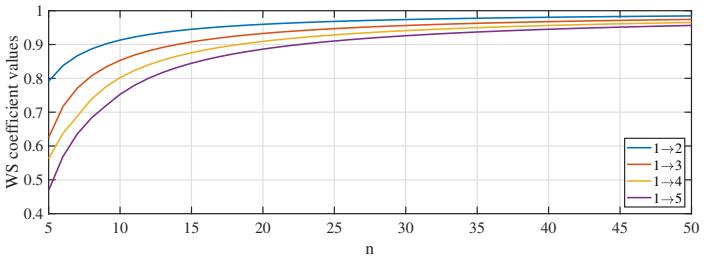

Fig. 2. The value of the WS coefficient depending on the length of the ranking (n) and the converted positions in the ranking.

4.3 Distribution of coefficients values

In the next step, we attempted to visualize the distributions for three indicators. Figure 3 presents the distribution of the τ Kendall factor for all possible

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

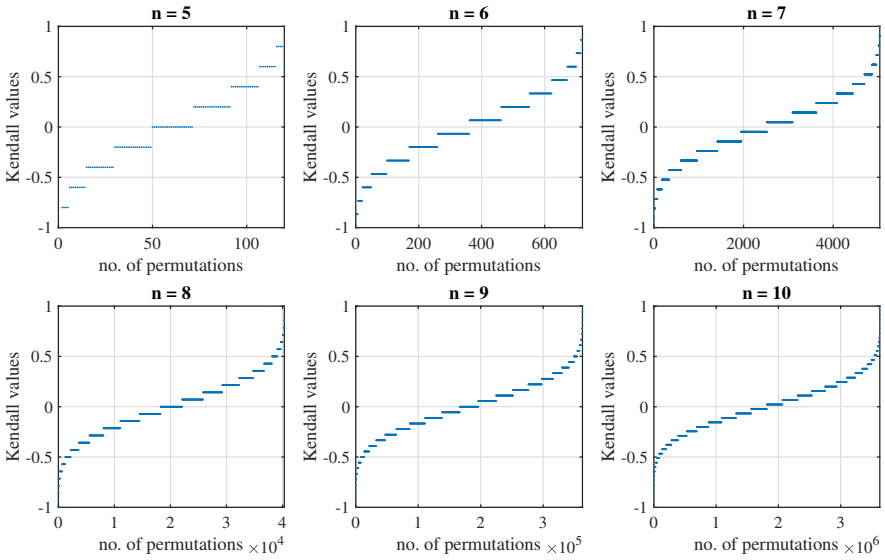

Fig. 3. Sorted distribution of all values of the Kednall coefficient in relation to the length of the ranking (n).

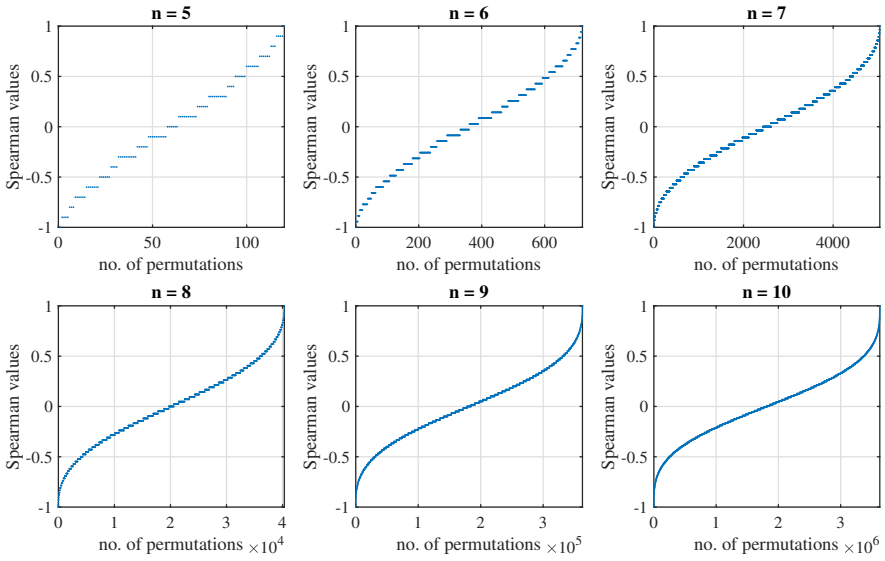

Fig. 4. Sorted distribution of all values of the Spearman coefficient in relation to the length of the ranking (n).

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

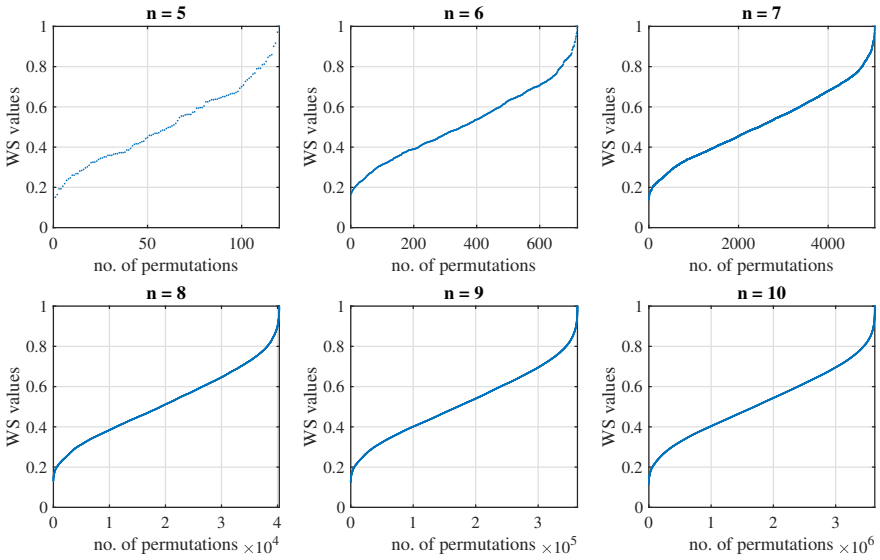

Fig. 5. Sorted distribution of all values of the WS coefficient in relation to the length of the ranking (n).

permutations of the sets of five, six, seven, eight, nine, and ten elements. Figures 4 and 5 show the distribution of the ρ Spearman coefficient and the WS coefficient, respectively. The shape of the ρ Spearman values is smoother than the τ Kendall. Both indicators have a symmetrical distribution, unlike the WS coefficient. The problem may be the interpretation of the WS value, because it is a new approach. However, the question arises when the similarity of the WS coefficient is low, medium, and high. A statistical analysis of the distribution of the WS should be carried out to define three appropriate linguistic terms and answer on this research question.

4.4 Definition of rankings similarity

All possible permutations and values of the WS coefficient are determined for ranks of the size from 3 to 10 elements. Based on the obtained values, we calculated basic statistics, which are presented in Table 5. For larger rankings, statistics are based on random samples of 100,000 rankings. Both population and random sampling data are used. Note the convergence of the arithmetic mean, standard deviation, and typical value ranges. The biggest differences concern the arithmetic mean, and it is equal to 0.207 (for a ranking of 10 and 1000 elements). Based on the analysis of typical values, i.e. interval of $[\bar{x} - S_x; \bar{x} + S_x]$, we identified the linguistic terms low, medium and high similarity of rankings.

It can indeed be said that if the WS is less than 0.234, then the similarity is low. If the value is higher than 0.808, then the similarity is high. The medium of

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

Table 5. A summary of the basic statistics of the WS coefficient for all possible permutations, where n length of the ranking.

n	$\bar{x}$	S_x	$\bar{x} - S_x$	$\bar{x} + S_x$	x_{min}	x_{max}
3	0.5208	0.2869	0.2339	0.8077	0.1875	1.0000
4	0.5313	0.2164	0.3149	0.7477	0.2083	1.0000
5	0.5135	0.1938	0.3197	0.7073	0.1510	1.0000
6	0.5195	0.1817	0.3378	0.7012	0.1656	1.0000
7	0.5164	0.1757	0.3407	0.6921	0.1383	1.0000
8	0.5197	0.1721	0.3476	0.6918	0.1314	1.0000
9	0.5193	0.1700	0.3493	0.6893	0.1252	1.0000
10	0.5208	0.1688	0.3520	0.6896	0.1144	1.0000

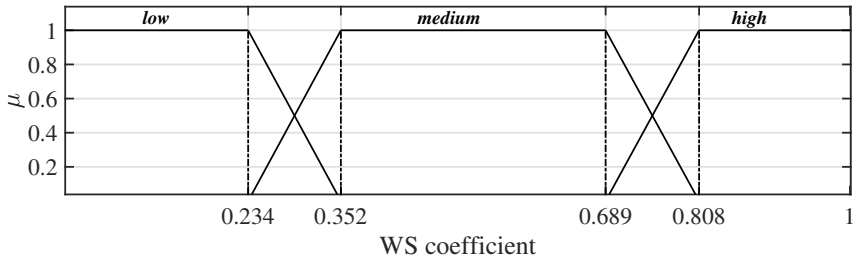

Fig. 6. The definitions of three linguistic terms, i.e., low, medium, and high similarity of rankings by using trapezoidal fuzzy numbers.

likeness, which corresponds to a typical value, belongs to the range from 0.352 to 0.689. The remaining values are values where we can talk about a partial belonging to linguistic concepts according to the theory of fuzzy sets, or just low/medium and medium/high concept can be used. Detailed definitions are presented in Figure 6. Linguistic values are important because they can be used to evaluate the adjustment of the reference and test rankings.

5 Conclusions

The main contribution of the paper is a proposal of the new coefficient of the rankings similarity. For this purpose, the short analysis of classical factors are presented, and some of their shortcomings are emphasized. The most critical is the equality of the values of the classical coefficients in case the ranking error concerns the replacement of a pair of adjacent alternatives (Table 1). The paper presents a theoretical foundation of proposed WS coefficient, which ensures that a new factor is free of identified shortcomings.

The results of numerical experiments compare all analyzed coefficients and their correctness, i.e., ρ Spearman, τ Kendall, G Goodman-Kruskal, and WS coefficients. Then, the distributions of τ Kendall, ρ Spearman, and WS coefficients were compared. WS values can be used to measure similarity of rankings.

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

Finally, three linguistic concepts were formulated for the low, medium, and high similarity of the two rankings. The properties of the WS coefficient indicate that it is a useful tool for comparing the similarity of rankings and is better suited for this purpose than the currently used correlation coefficients.

During the research, some improvement areas have been identified. The future work directions should concentrate on:

- further comparing between existing coefficients and the proposed WS;
- testing the use of the WS coefficient in real-life examples;
- detection and correction of WS coefficient shortcomings;
- adaptation of the proposed coefficient to uncertain (fuzzy) rankings.

Acknowledgments: The work was supported by the National Science Centre, Decision number UMO-2018/29/B/HS4/02725.

References

1. de Almeida, A.: Multicriteria modelling for a repair contract problem based on utility and the ELECTRE I method. IMA Journal of Management Mathematics **13**(1), 29–37 (2002). https://doi.org/10.1093/imaman/13.1.29
2. de Andrade, G., Alves, L., Andrade, F., de Mello, J.: Evaluation of power plants technologies using multicriteria methodology MACBETH. IEEE Latin America Transactions **14**(1), 188–198 (2016). https://doi.org/10.1109/TLA.2016.7430079
3. Ashraf, Q., Habaebi, M., Islam, M.R.: TOPSIS-based service arbitration for autonomic internet of things. IEEE Access **4**, 1313–1320 (2016). https://doi.org/10.1109/ACCESS.2016.2545741
4. Bandyopadhyay, S.: Ranking of suppliers with mcda technique and probabilistic criteria. In: International Conference on Data Science and Engineering. pp. 1–5. IEEE (August 2016). https://doi.org/10.1109/ICDSE.2016.7823948
5. Bandyopadhyay, S.: Appication of fuzzy probabilistic TOPSIS on a multi-criteria decision making problem. In: Second International Conference on Electrical, Computer and Communication Technologies. pp. 1–3. IEEE (February 2017). https://doi.org/10.1109/ICECCT.2017.8118038
6. Blest, D.C.: Theory & methods: Rank correlation — an alternative measure. Australian & New Zealand Journal of Statistics **42**(1), 101–111 (2000). https://doi.org/10.1111/1467-842X.00110
7. Brazdil, P.B., Soares, C.: A comparison of ranking methods for classification algorithm selection. In: Machine Learning: ECML 2000. pp. 63–75. Springer Berlin Heidelberg, Berlin, Heidelberg (2000). https://doi.org/10.1007/3-540-45164-1_8
8. Cavalcante, V., Alexandre, C., Ferreira, R.P., de Almeida, A.T.: A preventive maintenance decision model based on multicriteria method PROMETHEE II integrated with bayesian approach. IMA Journal of Management Mathematics **21**(4), 333–348 (2010). https://doi.org/10.1093/imaman/dpn017
9. Ceballos, B., Lamata, M.T., Pelta, D.A.: A comparative analysis of multi-criteria decision-making methods. Progress in Artificial Intelligence **5**, 315–322 (2016). https://doi.org/10.1007/s13748-016-0093-1
10. Pinto da Costa, J., Soares, C.: A weighted rank measure of correlation. Australian & New Zealand Journal of Statistics **47**(4), 515–529 (2005). https://doi.org/10.1111/j.1467-842X.2005.00413.x

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

11. Fagin, R., Kumar, R., Sivakumar, D.: Comparing top k lists. In: Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms. p. 28–36. SODA '03, Society for Industrial and Applied Mathematics, USA (2003). https://doi.org/10.1137/S0895480102412856
12. Faizi, S., Rashid, T., Sałabun, W., Zafar, S., Wątróbski, J.: Decision making with uncertainty using hesitant fuzzy sets. International Journal of Fuzzy Systems **20**(1), 93–103 (2018). https://doi.org/10.1007/s40815-017-0313-2
13. Genest, C., Plante, J.F.: On blest's measure of rank correlation. Canadian Journal of Statistics **31**(1), 35–52 (2003). https://doi.org/10.2307/3315902
14. Goodman, L., Kruskal, W.: Measures of association for cross classifications. Journal of the American Statistical Association **49**(268), 732—764 (1954). https://doi.org/10.1080/01621459.1954.10501231
15. Haddad, M., Sanders, D.: Selecting a best compromise direction for a powered wheelchair using PROMETHEE. IEEE Transactions on Neural Systems and Rehabilitation Engineering **27**(2), 228–235 (2019). https://doi.org/10.1109/TNSRE.2019.2892587
16. Hemili, M., Laouar, M.R.: Use of multi-criteria decision analysis to make collection management decisions. In: 3rd International Conference on Pattern Analysis and Intelligent Systems. pp. 1–5. IEEE (October 2018). https://doi.org/10.1109/PAIS.2018.8598495
17. Ishizaka, A., Siraj, S.: Are multi-criteria decision-making tools useful? an experimental comparative study of three methods. European Journal of Operational Research **264**(2), 462 – 471 (2018). https://doi.org/10.1016/j.ejor.2017.05.041
18. Ivlev, I., Jablonsky, J., P., P.K.: Multiple-criteria comparative analysis of magnetic resonance imaging systems. International Journal of Medical Engineering and Informatics **8**(2), 124–141 (2016). https://doi.org/10.1504/IJMEI.2016.075757
19. Jeremic, V.M., Radojicic, Z.: A new approach in the evaluation of team chess championships rankings. Journal of Quantitative Analysis in Sports **6**(3), 1–11 (2010). https://doi.org/10.2202/1559-0410.1257
20. Kendall, M.G.: A new measure of rank correlation. Biometrika **30**(1/2), 81–93 (1938). https://doi.org/10.2307/2332226
21. Luo, H.C., Sun, Z.X.: A study on stock ranking and selection strategy based on UTA method under the condition of inconsistence. In: 2014 International Conference on Management Science & Engineering 21th Annual Conference Proceedings. pp. 1347–1353. IEEE (August 2014). https://doi.org/10.1109/ICMSE.2014.6930387
22. de Monti, A., Toro, P.D., Droste-Franke, B., Omann, I., Stagl, S.: Assessing the quality of different mcda methods. In: Getzner, M., Spash, C., Stagl, S. (eds.) Alternatives for environmental evaluation, chap. 5, pp. 115–149. Routledge (2004). https://doi.org/10.4324/9780203412879
23. Mulliner, E., Malys, N., Maliene, V.: Comparative analysis of mcdm methods for the assessment of sustainable housing affordability. Omega **59**, 146 – 156 (2016). https://doi.org/10.1016/j.omega.2015.05.013
24. Ray, T., Triantaphyllou, E.: Evaluation of rankings with regard to the possible number of agreements and conflicts. European Journal of Operational Research **106**(1), 129–136 (1998). https://doi.org/10.1016/S0377-2217(97)00304-4
25. Sałabun, W.: The characteristic objects method: A new distance-based approach to multicriteria decision-making problems. Journal of Multi-Criteria Decision Analysis **22**(1-2), 37–50 (2015). https://doi.org/10.1002/mcda.1525

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47

26. Sałabun, W., Karczmarczyk, A., Wątróbski, J., Jankowski, J.: Handling data uncertainty in decision making with comet. In: IEEE Symposium Series on Computational Intelligence. pp. 1478–1484. IEEE (November 2018). https://doi.org/10.1109/SSCI.2018.8628934
27. Sałabun, W., Piegat, A.: Comparative analysis of mcdm methods for the assessment of mortality in patients with acute coronary syndrome. Artificial Intelligence Review **48**, 557–571 (2017). https://doi.org/10.1007/s10462-016-9511-9
28. Sari, J., Gernowo, R., Suseno, J.: Deciding endemic area of dengue fever using simple multi attribute rating technique exploiting ranks. In: 10th International Conference on Information Technology and Electrical Engineering. pp. 482–487. IEEE (July 2018). https://doi.org/10.1109/ICITEED.2018.8534882
29. Shen, K., Tzeng., G.: A refined drsa model for the financial performance prediction of commercial banks. In: International Conference on Fuzzy Theory and Its Applications. pp. 352–357. IEEE (December 2013). https://doi.org/10.1109/iFuzzy.2013.6825463
30. Shieh, G.S.: A weighted Kendall's tau statistic. Statistics & Probability Letters **39**(1), 17 – 24 (1998). https://doi.org/10.1016/S0167-7152(98)00006-6
31. Spearman, C.: The proof and measurement of association between two things. The American Journal of Psychology **15**(1), 72–101 (1904). https://doi.org/10.2307/1422689
32. Tian, G., Zhang, H., Zhou, M., Li, Z.: AHP, gray correlation, and TOPSIS combined approach to green performance evaluation of design alternatives. IEEE Transactions on Systems, Man, and Cybernetics: Systems. Part A Syst. Humans **48**(7), 1093–1105 (2017). https://doi.org/10.1109/TSMC.2016.2640179
33. Wątróbski, J., Jankowski, J., Ziemba, P., Karczmarczyk, A., Zioło, M.: Generalised framework for multi-criteria method selection. Omega **86**, 107–124 (2019). https://doi.org/10.1016/j.omega.2018.07.004
34. Yaraghi, N., Tabesh, P., Guan, P., Zhuang, J.: Comparison of AHP and Monte Carlo AHP under different levels of uncertainty. IEEE Transactions on Engineering Management **62**(1), 122–132 (2015). https://doi.org/10.1109/TEM.2014.2360082
35. Zhang, C., Liu, X., Jin, J.G., Liu, Y.: A stochastic ANP-GCE approach for vulnerability assessment in the water supply system with uncertainties. IEEE Transactions on Engineering Management **63**(1), 78–90 (2015). https://doi.org/10.1109/TEM.2015.2501651
36. Zhang, P., Yao, H., Qiu, C., Liu, Y.: Virtual network embedding using node multiple metrics based on simplified ELECTRE method. IEEE Access **6**, 37314–37327 (2018). https://doi.org/10.1109/ACCESS.2018.2847910

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_47