APPLICATION OF PAGERANK CENTRALITY IN MULTI-CRITERIA DECISION MAKING

Nikola Kadoić
University of Zagreb, Faculty of organization and informatics, Pavlinska 2, HR-42000 Varaždin
E-mail: nkadoic@foi.hr

Nina Begičević Ređep
University of Zagreb, Faculty of organization and informatics, Pavlinska 2, HR-42000 Varaždin
E-mail: nbegicev@foi.hr

Blaženka Divjak
University of Zagreb, Faculty of organization and informatics, Pavlinska 2, HR-42000 Varaždin
E-mail: bdivjak@foi.hr

Abstract: Multicriterial analysis is a highly developed area since there is a large number of multi-criteria decision-making (MCDM) methods. The multi-criteria analysis enables a precise problem analysis and ensures the rationality of decision that is made. However, all problem analysis using different methods can give different results (decision), so it is important to recognize which MCDM method is appropriate for a particular situation. There are MCDM methods by using which we can model dependencies and influences between the criteria in decision-making problem. One of the most used MCDM methods that are used in terms of problem analysis is the analytic network process (ANP). Previous researches discussed some problems related to using the ANP in decision-making. As a solution to those problems using the PageRank centrality can be considered. In this paper, we are presenting several possibilities of applying the PageRank centrality for multi-criteria analysis. Presented possibilities are compared and discussed. As a result, using the weighted PageRank centrality is proposed as the optimal solution for multi-criteria analysis when dependencies (influences) between the criteria are examined.

Keywords: criteria, multi-criteria decision-making, MCDM, ANP, PageRank centrality, PageRank, influences, dependencies

1 INTRODUCTION

There are many methods that can be used in terms of multi-criteria analysis. Each of them models the problem differently, but with the main goal – to find an optimal solution to a problem that has been analysed. In this paper, we are analysing decision methods from the perspective of modelling the influences (dependencies) between the criteria. Those two concepts have the opposite meaning [1]: if the first criterion influences the second criterion, then the second criterion depends on the first criterion.

This paper is motivated with the research in the scope of the project "Development of a methodological framework for strategic decision-making in higher education – a case of open and distance learning (ODL) implementation." As a part of the research on the project, different MCDM methods were analysed from the position of applicability in the area of higher education. It is concluded that the area of higher education is characterized by the existence of influences between the criteria [2]. However, literature review analysis resulted with the conclusion that, in the analysis of MCDM problems in the area of higher education, methods which do not support modelling influences (dependencies) between the criteria (such as analytic hierarchy process (AHP)) are much more often used instead of methods which support this feature. In this project, special attention is given to modelling influences between the criteria and analysis of the method analytic network process (ANP). The ANP is the most often used method for modelling the influences (dependencies) between the criteria. However, it has many disadvantages, and this is the reason for such literature review results.

In the second section of this paper, we will shortly present the ANP method and discuss some of its characteristics. In the third section, we will present several types of PageRank centrality and its possibility for using in terms of multi-criteria analysis. Finally, we will discuss and compare the presented types of PageRank centrality and list the advantages of using PageRank comparing to the ANP.

2 THE ANALYTIC NETWORK PROCESS (ANP)

We presented the decision-making process using the ANP in our last SOR paper [3], and CJOR paper [4] which followed the SOR paper, and our further analysis will be demonstrated on the decision-making problem that is discussed in those papers. The problem is related to the evaluation of senior researchers (scientists). Senior researchers are active in both the research and teaching fields. In this analysis, we will not include alternatives. The decision-making problem is presented in Figure 1.

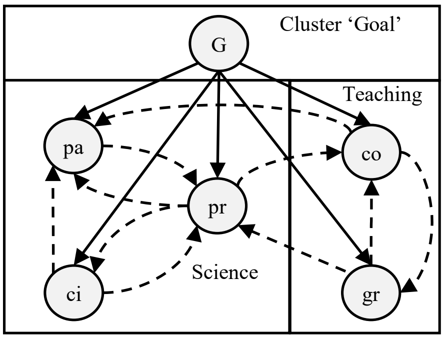

Figure 1: Network Structure of the problem evaluation of the scientists

The steps in ANP (also adapted from [5], [6]):

- Problem structuring phase: it is related to the creation of the network structure. It is presented in Figure 1. The model consists of five criteria (papers, pa; projects, pr; citations, ci; courseware, co; grades from students, gr) which are grouped into two clusters (Teaching and Science). The model also includes the decision-making goal (node G in cluster Goal). The arrows between the elements represent dependencies in the model.
- Pairwise comparisons procedure and creating the weighted supermatrix (Table 2):
  - Comparing criteria in order to reach unweighted supermatrix (Table 1),
  - Comparing clusters in order to reach clusters' weights which are needed to obtain the weighted supermatrix. In this example, we decided that all clusters are equally important,
  - Combining the unweighted supermatrix with clusters' weights. (Much more detailed procedure description is available in our previous papers [3], [4].)
- Creation the limit matrix (Table 3) by multiplying the weighted supermatrix with itself until it converges.

Table 1: Unweighted Supermatrix

	G	co	gr	pa	ci	pr
G	0	0	0	0	0	0
co	0.33	0	1	0	0	1
gr	0.67	1	0	0	0	0
pa	0.25	1	0	0	0.4	0.6
ci	0.25	0	0	0	0	0.4
pr	0.5	0	1	1	0.6	0

Table 2: Weighted Supermatrix

	G	co	gr	pa	ci	pr
G	0	0	0	0	0	0
co	0.08	0	0.5	0	0	0.5
gr	0.17	0.5	0	0	0	0
pa	0.186	0.5	0	0	0.4	0.3
ci	0.186	0	0	0	0	0.2
pr	0.383	0	0.5	1	0.6	0

Table 3: Limit matrix

	G	co	gr	pa	ci	pr
G	0.00	0.00	0.00	0.00	0.00	0.00
co	0.23	0.23	0.23	0.23	0.23	0.23
gr	0.11	0.11	0.11	0.11	0.11	0.11
pa	0.24	0.24	0.24	0.24	0.24	0.24
ci	0.07	0.07	0.07	0.07	0.07	0.07
pr	0.34	0.34	0.34	0.34	0.34	0.34

The final priorities can be found in any column of the limit matrix.

The ANP is much deeper analysed in paper [7]. In the paper, a list of weak points of the ANP is provided. Those characteristics result from the fact that ANP is much less used than it should be used. Most of them influence the complexity of the ANP implementation, misunderstanding of certain steps of the method, and long duration of the implementation process [1]. However, the most exciting three characteristics are [7]:

- The inseparability of the criteria and alternatives. In some decision-making problems, if there is no directed connection between any of two nodes, then it is possible that at least some nodes will weight 0.0, or even the whole limit matrix is zero-matrix. This is not the case in this decision-making problem.
- The influence of the goal node on the priorities. The interesting and slightly intriguing characteristic of the ANP is the fact that the priorities with respect to the goal (first column in unweighted and weighted supermatrix) do not influence the finale priorities in the limit matrix. So, if we change the numbers in the first column (respecting that the sum of the numbers equals 1), the finale priorities will remain the same. This means that the node goal is not necessary for the model and that only the dependencies between the elements determine the final priorities. Indeed, there is a large number of papers which do not include goal as an element of the network structure. However, the goal node is theoretically defined as a network element (cluster), and it is a necessary element in the AHP, which is a 'weaker' variant of the ANP.
- The stochasticity of the supermatrix in the ANP. The most interesting and the most intriguing characteristic of the ANP is related to the stochasticity of the supermatrix in the ANP. If we look at the connection between the pa and pr, in both, unweighted and weighted supermatrix, we do not know how strong this connection is. The element pr can influence the element pa weakly, strongly or very strongly. So, independently of the intensity of this influence, the final priorities will remain the same. In the DEMATEL method [8], we use scale 0-4 to describe the intensity (level) of the influence (dependency) between two elements. For all real values between 0 and 4, in this case, we will reach the same final priorities. Let us say that there are four elements in the model; each influence any other (not itself). All elements influence the first element with intensity 4, the second element with intensity 3, the third element with intensity 3, and finally the third element with intensity 1. We got the situation that the first element depends on mostly by others, and the last element depends at least by others. However, in ANP, all elements will have the same priority. The reason for that is 'forcing' the stochasticity of supermatrix. It relativizes the problem, and the solution (the decision) might not be optimal.

3 THE PAGERANK CENTRALITY

The PageRank centrality is a special type of eigenvalue centrality. The eigenvalue centrality for undirected and unweighted networks is calculated using Equation 1 [9].

$$\boldsymbol{C_E(i) = \frac{1}{\lambda}\sum_{j\in M(i)} C_E(j) = \frac{1}{\lambda}\sum_{j\in N} a_{ij} C_E(j)} \qquad (1)$$

where $\boldsymbol{M(i)}$ is a set of neighbours of actor $\boldsymbol{i}$, $\boldsymbol{\lambda}$ is a constant (the maximum eigenvalue) and $\boldsymbol{a_{ij}}$ is an element of a matrix of neighbours $\boldsymbol{A}$. PageRank centrality is used for directed networks, and there are variants of this measure in terms of weighted and unweighted graphs.

The PageRank centrality can be calculated using the iterative procedure [10] or using Equation 2.

$$\lim_{k\to\infty} A^k Z_0 = \widetilde{A} \qquad (2)$$

where A is the matrix of neighbours, Z_0 is a one-column matrix which contains elements $\frac{1}{N}$, and $\widetilde{A}$ is a matrix of priorities.

In terms of decision-making with the ANP, matrix A can correlate with weighted supermatrix. Additionally, we can create weighted supermatrix avoiding the pairwise procedure on the node level as described in the paper [4]:

- the starting point is the identification of the intensities of influences between the elements in the network (Table 4),
- then, that matrix can be stochastically normalized using the normalization by sum (Table 6) or transition matrix (function).

Table 4: Matrix of influences intensities between the criteria

	co	gr	pa	ci	pr
co	0	3	0	0	2
gr	2	0	0	0	0
pa	2	0	0	2	3
ci	0	0	0	0	2
pr	0	3	4	3	0

Table 5: (Un)weighted supermatrix

	co	gr	pa	ci	pr
co	0	0.5	0	0	0.5
gr	0.5	0	0	0	0
pa	0.5	0	0	0.4	0.3
ci	0	0	0	0	0.2
pr	0	0.5	1	0.6	0

The problem that appears with the powering the supermatrix is already mentioned earlier when three the most interesting characteristics of the ANP were listed. The solution to that problem, the PageRank calculates the new matrix as in Equation 3 [11], [12].

$$G = \alpha \cdot A + (1 - \alpha) \cdot E \qquad (3)$$

In most cases, $\alpha = 0.85$ [13]. (Note: If a certain column in A contains all 0, then this column has to be replaced with ca column whose values equal $\frac{1}{N}$.)

Adding the E in Equation 3 ensures that original matrix A converges to the non-zero matrix, and now it is no longer possible that we cannot calculate the global priorities. The role of E is making a matrix (G) whose graph is strongly connected – there is a direct connection between any two nodes in G. Additionally, the influence of G on the final priorities is negligible (0.15). This is the first possible application of PageRank to eliminate at least one of the weak points of the ANP (the inseparability of the criteria and alternatives).

The PageRank centrality can be also interesting in terms of eliminating the issues that are the result of the stochasticity of the supermatrix in the ANP [14]. Then, the original PageRank centrality algorithm should be changed in a way that we sum powers of the non-stochastic supermatrix and then aggregate and normalize the results. The sums of the columns of the non-stochastic supermatrix should be less than 1 because - in only that case, it is possible to sum all the powers (using Equation 4).

$$\widetilde{A} = \sum_{k\to\infty} A^k = A \cdot (A - I)^{-1} \qquad (4)$$

Consequently, the original PageRank for directed and weighted graphs (matrices) is transformed as follows:

1. The starting point is a matrix of influences between the criteria (Table 4)
2. In the second step, we are dividing each value in Table 4 with the maximum sum of columns, which is increased by 1. The maximum sum is in column pr and equals 7, which means that all values in Table 4 will be divided by 8. The result is presented in Table 6.
3. Respecting the Equation 3, we have to calculate the matrix $I - A$. The result is presented in Table 7.

4. Now we calculate the inverse of matrix $\boldsymbol{I}-\boldsymbol{A}$. (Table 8)
5. Multiplication of Tables 7 and 8 (Table 9)
6. Calculation of the sum of rows (ΣR) and columns (ΣC) of Table 9 and their difference, $\boldsymbol{d}$ (see Table 9). The difference should then be normalized. There are several ways to do it: using the absolute value of the smallest value (difference), or any other higher number. Increasing the normalization value will result in smaller differences between the priorities on end. When normalization value, $\boldsymbol{n}$, is chosen, it should be added to differences, $\boldsymbol{d}+\boldsymbol{n}$. Now, all values are positive, and it is possible to calculate the criteria weights (normalization by sum).
   In this example, we chose the normalization value as differences between the highest difference and the lowest difference.

Table 6: Step 2

	co	gr	pa	ci	pr
co	0	0.375	0	0	0.25
gr	0.25	0	0	0	0
pa	0.25	0	0	0.25	0.375
ci	0	0	0	0	0.25
pr	0	0.375	0.5	0.375	0

Table 7: Step 3

	co	gr	pa	ci	pr
co	1	-0.375	0	0	-0.25
gr	-0.25	1	0	0	0
pa	-0.25	0	1	-0.25	-0.375
ci	0	0	0	1	-0.25
pr	0	-0.375	-0.5	-0.375	1

Table 8: Step 4

	co	gr	pa	ci	pr
co	1.21	0.62	0.22	0.22	0.44
gr	0.30	1.15	0.05	0.05	0.11
pa	0.47	0.48	1.40	0.65	0.81
ci	0.10	0.19	0.20	1.20	0.40
pr	0.38	0.74	0.80	0.80	1.59

Table 9: Steps 5 and 6

	co	gr	pa	ci	pr	ΣR		ΣC	d	$d+n$	priorities
co	0.21	0.62	0.22	0.22	0.44	1.71		1.46	0.24	2.89	0.22
gr	0.30	0.15	0.05	0.05	0.11	0.68		2.18	-1.50	1.14	0.09
pa	0.47	0.48	0.40	0.65	0.81	2.82		1.68	1.14	3.78	0.29
ci	0.10	0.19	0.20	0.20	0.40	1.08		1.93	-0.85	1.80	0.14
pr	0.38	0.74	0.80	0.80	0.59	3.32		2.35	0.97	3.61	0.27
ΣC	1.46	2.18	1.68	1.93	2.35			highest	1.14	13.22	
								lowers	-1.50		
								norm. value. n	2.64		

If we compare the final priorities with the priorities in Table 3 (from the Limit matrix), we can identify some differences. Even though the ranks of the criteria remained the same, there are absolute differences between the criteria weights. Now, the problem with stochasticity of the supermatrix has been eliminated.

4 CONCLUSIONS

In this paper, we were dealing with the possibilities to use the PageRank centrality to diminish some of the weak points of the method ANP. Using the PageRank centrality in the process of calculating the limit matrix (from the weighted supermatrix), we can directly influence and eliminate the weak point of the ANP related to the converging to zero matrix - inseparability the criteria and alternatives. Indeed, there are real-world requests to calculate the criteria weights when the alternatives are still not known. In those cases, very often, when there is a small number of the connections in the model, some of the criteria, or even all would weight 0.0. If we found those criteria irrelevant; we will not put them into the model at all – so we cannot accept 0.0. as the final criteria weight of certain criteria. PageRank solves this situation.

The other benefit of the PageRank centrality is related to the dealing with stochasticity in supermatrix in ANP. When the matrix is stochastic, it is ensured that it will converge into the limit matrix from which we can directly take the criteria weights. However, we should not force the stochasticity of the supermatrix just because there is a 'great' mathematical property of stochastic matrix in terms of its powering. If we want to use the PageRank approach, which is not stochastic, it is important to have the original matrix of the intensities of the influences between the criteria – not pairwise comparisons priorities. This is not a problem since the pairwise comparisons are also resulting from those intensities between the elements. Additionally, in this approach, when we use original intensities of the influences and avoid making the pairwise comparisons, we lower some other ANP weak points.

The only open issue in terms of ANP characteristics is related to the influence of the goal on the criteria weights. In ANP, and presented approach, criteria weights are consequences of influences between the criteria, not consequences of their importance with respect to the goal, too. To solve those issues, we can use possible aggregate the obtained results with the AHP results by using ex. arithmetic mean.

References

[1] N. Kadoić, N. Begičević Ređep, and B. Divjak, "A new method for strategic decision-making in higher education," *Central European Journal of Operations Research*, no. Special Issue of Croatian Operational Research Society and Collaborators, Oct. 2017.

[2] N. Kadoić, N. Begičević Ređep, and B. Divjak, "E-learning decision making: methods and methodologies," in *Re-Imagining Learning Scenarios*, 2016, vol. CONFERENCE, no. June, p. 24.

[3] N. Kadoić, N. Begičević Ređep, and B. Divjak, "Decision Making with the Analytic Network Process," in *SOR 17 Proceedings*, 2017, pp. 180–186.

[4] N. Kadoić, B. Divjak, and N. Begičević Ređep, "Integrating the DEMATEL with the analytic network process for effective decision-making," *Central European Journal of Operations Research*, vol. 27, no. 3, pp. 653–678, Sep. 2019.

[5] T. L. Saaty and L. G. Vargas, *Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks*. Springer; Softcover reprint of hardcover 1st ed. 2006 edition (December 28, 2009), 2006.

[6] T. L. Saaty and L. G. Vargas, "The Analytic Network Process," in *Iranian journal of operational research*, vol. 1, no. 1, 2013, pp. 1–28.

[7] N. Kadoić, "Characteristics of the Analytic Network Process, a Multi-Criteria Decision-Making Method," *Croatian Operational Research Review*, vol. 9, no. 2, pp. 235–244, 2018.

[8] N. Kadoić, N. Begičević Ređep, and B. Divjak, "Structuring e-Learning Multi-Criteria Decision Making Problems," in *Proceedings of 40th Jubilee International Convention, MIPRO 2017*, 2017, pp. 811–817.

[9] M. Mincer and E. Niewiadomska-Szynkiewicz, "Application of social network analysis to the investigation of interpersonal connections," *Journal of Telecommunications and Information Technology*, vol. 2012, no. 2, pp. 83–91, 2012.

[10] D. Munđar and D. Horvat, "Rangiranje ekipa i prognoziranje ishoda u rukometu korištenjem PageRank algoritma," *Poučak*, vol. 67, pp. 8–15, 2016.

[11] T. Csendes and E. Antal, "PageRank Based Network Algorithms for Weighted Graphs with Applications to," vol. 2, pp. 209–216, 2010.

[12] T. Kumari, A. Gupta, and A. Dixit, "Comparative Study of Page Rank and Weighted Page Rank Algorithm," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, no. 2, pp. 2929–2937, 2014.

[13] W. Xing and A. Ghorbani, "Weighted PageRank algorithm," *Proceedings. Second Annual Conference on Communication Networks and Services Research, 2004.*, pp. 305–314, 2004.

[14] N. Kadoić, "Nova metoda za analizu složenih problema odlučivanja temeljena na analitičkom mrežnom procesu i analizi društvenih mreža," University of Zagreb, 2018.