Node Importance Evaluation of Complex Network Based on M-TOPSIS Method

Laijun Luo\textsuperscript{1 \textsuperscript{a}}, Haiping Ren\textsuperscript{2 \textsuperscript{b}}

\textsuperscript{1}School of Software and Engineering, Jiangxi University of Science and Technology, Nanchang, China
\textsuperscript{2}Teaching Department of Basic Subjects, Jiangxi University of Science and Technology, Nanchang, China
\textsuperscript{a}junbetter@163.com, \textsuperscript{b}chinarhp@163.com

Abstract—The importance evaluation of nodes in complex networks is of great significance. Usually, the importance evaluation of nodes can be based on degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, and structural holes, etc. It is more reliable to evaluate the importance of the nodes with multiple indicators. Therefore, based on the principle of multiple attribute decision making (MADM), a new method for evaluating the importance of nodes based on M-TOPSIS is proposed in this paper, this method can be used to evaluate the nodes by using multiple indicators, and it can avoid the typical problems existing in the traditional method based on TOPSIS, such as the problem of inverted order of evaluation value. Experimental results show that the new method based on M-TOPSIS has good reliability and adaptability.

1. INTRODUCTION
In recent years, the research on complex network has attracted many scholars' attention. The research results of complex networks play an important role in social networks [1], virus spreading networks, marketing networks, biological networks, power networks and other specific complex networks. In related researches, the importance evaluation of complex network nodes is an important research direction. In terms of status and functions, the importance of different nodes in complex network is different [2]. Therefore, it is of great value to evaluate the importance of nodes and find the key nodes. For example, in the online social networks, using the importance evaluation of nodes can quickly find the key nodes [3], and using the key nodes can effectively guide the public opinion, or control the spreading of rumours’ information quickly. In the virus spreading network, it is possible to quickly locate the source of the infection and isolate or treat it, thus effectively preventing the spread of the virus. In the marketing network, the product can be sold quickly by key nodes [4]. In addition, it is of great value and significance to evaluate the importance of nodes in power networks, biological networks and transportation networks.

In general, evaluating the importance of nodes in complex network uses degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, and structural holes [5], these indicators describe the importance of nodes from different aspects. However, the structure of complex network is irregular and has its own characteristics. It is one-sided to evaluate the importance of network nodes with a single indicator, and it is difficult to guarantee the reliability of the results [6].
The importance of nodes in the network is related to several factors, so it is more reliable to evaluate the importance of network nodes with multiple indicators. Therefore, this paper proposes a multiple attribute decision making (MADM) method to evaluate the importance of nodes, the degree centrality, closeness centrality, betweenness centrality, and structural which holes are considered as attributes in multiple attribute decision making.

As a classic multiple attribute decision making method, TOPSIS has attracted many scholars’ attention. But, the TOPSIS method has the problem of inverse order, and the comprehensive evaluation value can only indicate the relative proximity within each evaluation method [7], it can’t represent the true closeness to the ideal optimal method.

Therefore, this paper proposes an improved TOPSIS decision method for the importance evaluation of complex network nodes, called M-TOPSIS method. The experimental results show that this method can accurately evaluate the importance of nodes of the kite network. In addition, the indicators involved in this method can be changed and expanded according to specific requirements, and the method has better reliability and adaptability.

2. PRELIMINARY KNOWLEDGE

The following abstractions are used for a complex network: Graph G(V, E) represent a complex network, using V={v_1, v_2, ..., v_n} represents the set of all nodes in the network, using E={e_1, e_2, ..., e_m}⊆V×V represents the set of all edges in the network. It can be seen that the number of nodes in the graph is n, and the number of edges is m. M = (m_{ij})_{n×n} represents the adjacency matrix of the graph.

If there is an edge between the nodes v_i and v_j, then m_{ij} = 1, otherwise m_{ij} = 0. There are many indicators for evaluating the importance of nodes in complex networks, but, it does not mean the more the better [8]. The key indicators in evaluating importance of nodes are selected and defined as follows:

**Definition 1:** Degree centrality (DC)

The ratio of the number of edges connected with the node v_i to the possible number of edges with the node v_i, defined as follows:

\[ DC_i = K_i / (n - 1) \]

k_i represents the number of edges directly connected with node v_i, degree centrality represents the ability of a node connecting directly with other nodes, the larger value of DC_i, the more important node v_i is [9].

**Definition 2:** Closeness centrality (CC)

The closeness centrality of the node v_i is the reciprocal of the sum of its shortest distances to the other nodes, defined as follows:

\[ CC_i = (n - 1) / \sum_{j=1}^{n} k_{ij} \]

k_{ij} represents the shortest distance from node v_i to node v_j. the larger value of CC_i, the more important the node v_i is in the network, the greater the value of CC_i, the more important the node v_i is [10].

**Definition 3:** Betweenness centrality (BC)

Among all the shortest paths, the more shortest paths passing through node v_i, the more important node v_i is, BC defined as follows:

\[ BC_i = \sum_{j \neq i \neq k \in V} \frac{g_{jk}(i)}{g_{jk}} \]
Here, \( g_{jk} \) is the value of all shortest paths from node \( v_j \) to node \( v_k \) and \( g_{jk}(i) \) represents the number of shortest paths formed by node \( v_j \) and node \( v_k \), and the shortest paths pass through node \( v_i \).

**Definition 4:** Structural holes (C)

If there is no direct connection or no indirect redundancy between the two nodes, then the obstruction between two nodes is the structural hole. Burt proposed a network constraint coefficient for calculating the structural holes, defined as follows:

\[
C_i = \sum_j \left( P_{0j} + \sum_{q \neq j} P_{iq} P_{qj} \right)^2
\]

\( v_q \) is the indirect connection node between node \( v_i \) and node \( v_j \), \( p_{ij} \) is the ratio of the cost that the node \( v_i \) spent on the node \( v_j \) to its total cost, the smaller value of \( C \), the more important the node \( v_i \) is.

### 3. EVALUATION METHOD BASED ON M-TOPSIS

TOPSIS method is a widely used static comprehensive evaluation method, but it has a problem of reverse order [11]. The comprehensive evaluation value of TOPSIS can only embody the relative proximity of the evaluation object, and the range which evaluation value differentiates each evaluation object’s optimal and inferior is limited [12]. Since TOPSIS method is widely used, it is necessary to improve TOPSIS method. Document [8] proposed a new improved TOPSIS method, the new method can not only guarantee the order of evaluation values, but also be more sensitive than the traditional TOPSIS method. In this paper, based on the idea of document [8], a new node importance evaluation method based on M-TOPSIS method is proposed. Suppose \( X = \{x_1, x_2, \cdots, x_m\} \) is the set of \( m \) nodes, \( O = \{o_1, o_2, \cdots, o_n\} \) is the set of selected evaluation indicators. \( a_{ij} \) is the evaluation value of node \( x_i \) on \( o_j \), thus we can get evaluation matrix \( A = (a_{ij})_{m \times n} \).

The steps as follows:

**Step2.** The coefficient of variation is used to determine the weight of the indicator.

First, calculate the coefficient of variation of the indicator \( j \): \( \delta_j = s_j / \bar{x}_j \), among them,

\[
\bar{x}_j = \frac{1}{m} \sum_{i=1}^{m} x_{ij}
\]

is the average value and \( s_j = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (x_{ij} - \bar{x}_j)^2} \) is the standard deviation.

Then, the weight of the coefficient of variation of the \( j \) indicator is defined as

\[
w_j = \frac{\delta_j}{\sum_{j=1}^{n} \delta_j}, \quad j = 1, 2, \ldots, n.
\]

**Step3.** In order to eliminate the impact of different physical dimensions, the linear transformation method is used to standardize the decision matrix into \( R = (r_{ij})_{m \times n} \), among them, if \( o_j \) is the benefit type indicator, then

\[
r_{ij} = \frac{(a_{ij} - \min_i(a_{ij}))}{(\max_i(a_{ij}) - \min_i(a_{ij}))},
\]

if \( o_j \) is cost type indicator, then

\[
r_{ij} = \frac{(\max_i(a_{ij}) - a_{ij})}{(\max_i(a_{ij}) - \min_i(a_{ij}))}.
\]

**Step4.** Determine the ideal point \( x^+ \) and the negative ideal point \( x^- \) as follows:

\[
x^+ = \{r^+_1, r^+_2, \cdots, r^+_n\} = \max\{r_{ij}, j = 1, 2, \cdots, n\},
\]

\[
x^- = \{r^-_1, r^-_2, \cdots, r^-_n\} = \min\{r_{ij}, j = 1, 2, \cdots, n\}.
\]
Step 5. Calculate the distance $D_i^+$ from solution $x_i$ to $x^*$ and calculate the distance $D_i^-$ from $x_i$ to $x^*$ as follows:

$$D_i^+ = \sum_{j=1}^{n} w_j | r_{ij} - r_{i,j^*} |,$$

$$D_i^- = \sum_{j=1}^{n} w_j | r_{ij} - r_{i,j^-} |.$$

Step 6. According to the idea of document [11], use $oD^+$ as the x axis, use $oD^-$ as the y axis, then, establish coordinate plane $D^+ oD^-$. The alternative Point $x_i$ is represented by the point $(D_i^+, D_i^-)$ on the coordinate plane. Assume that the corresponding scheme for

$$(D^+, D^-) = (\min_i\{D_i^+\}, \max_i\{D_i^-\})$$

is the optimal reference scheme $x_i$, then the Euclidean distance between the alternative $x_i$ and the optimal reference scheme $x_i$ is $d(x_i, x) = \sqrt{(D_i^+ - D^+)^2 + (D_i^- - D^-)^2}$.

Step 7. Sort and select alternative scheme according to the order of $d(x_i, x)$ from small to large. The smaller the $d(x_i, x)$, the better the scheme. If there are two scenarios, $x_i$ and $x_j$ ($i \neq j$), make $d(x_i, x) = d(x_j, x)$, then $x_2$ is the second best reference scheme, the corresponding second best reference point is, when calculating

$$C(x_i, x_2) = \sqrt{(D_i^+ - \min_i\{D_i^+\})^2 + (D_i^- - \min_i\{D_i^-\})^2},$$

the evaluation object with smaller value of $C(x_i, x_2)$ is better, that is, in the case of equal distance between the evaluation scheme and the optimal reference point $x_i$, the point closest to the second best reference point $x_2$ is selected as the better one.

4. EXAMPLE ANALYSIS

This example will use the "kite network" as the analysis object, it can be seen from the data of Figure 1 and Table 1, that the number of nodes directly connected to the node $v_4$ is the maximum, and the degree of node $v_4$ is the largest. Node $v_6$ has the same status as node $v_7$ in the network, and the two nodes have the largest number of closeness centrality and the smallest value of structural hole. Node $v_8$ is the necessary node of other nodes to connect node $v_9$ and node $v_{10}$ in the kite network the node $v_8$ has the special status.

Figure 1. Kite network

For simplicity and clarity, the nodes' importance will be evaluated by four indicators: DC (degree centrality), CC (closeness centrality), BC (betweenness centrality) and C (structural holes), among them, DC, CC and BC are benefit indicators and C is a cost indicator.
Table 1. Quantized values of each indicator of all nodes in kite network

| ID | DC  | C   | CC  | BC  |
|----|-----|-----|-----|-----|
| 1  | 0.4444 | 0.5783 | 0.5882 | 1.67 |
| 2  | 0.4444 | 0.5783 | 0.5882 | 1.67 |
| 3  | 0.3333 | 0.7059 | 0.5556 | 0.00 |
| 4  | 0.6667 | 0.4746 | 0.6667 | 7.33 |
| 5  | 0.3333 | 0.7059 | 0.5556 | 0.00 |
| 6  | 0.5556 | 0.4701 | 0.7143 | 16.67 |
| 7  | 0.5556 | 0.4701 | 0.7143 | 16.67 |
| 8  | 0.3333 | 0.4944 | 0.6667 | 28.00 |
| 9  | 0.2222 | 0.5556 | 0.4762 | 16.00 |
| 10 | 0.1111 | 1.25  | 0.3448 | 0.00 |

Next, the new method will be used to sort the importance of nodes, steps are as follows:

**Step1.** Using the coefficient of variation method to determine the indicator weight vector: 

\[ w = (w_1, w_2, w_3, w_4) = (0.1980, 0.1777, 0.0923, 0.5320) . \]

**Step2.** The normalized decision matrix obtained by the linear transformation method is

\[
R = \begin{bmatrix}
0.5999 & 0.1387 & 0.3413 & 0.0596 \\
0.5999 & 0.1387 & 0.3413 & 0.0596 \\
0.3999 & 0.3023 & 0.4295 & 0 \\
1.0000 & 0.0058 & 0.1288 & 0.2618 \\
0.3999 & 0.3023 & 0.4295 & 0 \\
0.8000 & 0 & 0 & 0.5954 \\
0.8000 & 0 & 0 & 0.5954 \\
0.3999 & 0.0312 & 0.1288 & 1.0000 \\
0.2000 & 0.1096 & 0.6444 & 0.5714 \\
0 & 1.0 & 1.0 & 0 \\
0 & 1.0 & 1.0 & 0
\end{bmatrix}
\]

**Step3.** According to the steps 4 to 6 in the previous section, the following optimal reference scheme can be obtained as follows: 

\[ x_1 = (D^*, D^*) = (0.3714, 0.6286) , \]

the Euclidean distance between \( x_i \) and \( x_1 \) is as follows:

\[
d(x_i, x_1) = 0.5967, \quad d(x_2, x_1) = 0.5967, \quad d(x_3, x_1) = 0.6450, \quad d(x_4, x_1) = 0.3938, \quad d(x_5, x_1) = 0.6450, \quad d(x_6, x_1) = 0.2171, \quad d(x_7, x_1) = 0.2171, \quad d(x_8, x_1) = 0, \quad d(x_9, x_1) = 0.2914, \quad d(x_{10}, x_1) = 0.5072
\]

**Step4.** According to the principle that the smaller the \( d(x_i, x) \) value is, the better the scheme is, the ranking result of node importance in Fig1 can be obtained as follows:

\[ x_8 > (x_6 = x_7) > x_9 > x_4 > x_5 > (x_1 = x_4) > (x_2 = x_3) , \]

the detailed explanation can be found in document [4].

5
5. CONCLUSIONS
The importance of nodes in complex networks can’t be described only by one single index, but also by multiple indexes from different perspectives. Document [5] proposed a method of node importance evaluation based on TOPSIS method. However, TOPSIS method has a problem of inverse sequence, its evaluation value $C_i$ can only express the relative proximity within each evaluation object, but can’t express the proximity with the ideal optimal scheme, the role of evaluation value $C_i$ in differentiating the advantages and disadvantages of each evaluation object is limited. To avoid the shortcomings of TOPSIS in sorting the importance of nodes in complex networks, this paper proposes a new improved TOPSIS method for evaluating the importance of nodes in complex networks, named M-TOPSIS method. This method not only has strong order preservation, but also has better accuracy than traditional TOPSIS method. The new method is used to evaluate the importance of the nodes in the kite network, the experimental result shows that the method has good effectiveness.

ACKNOWLEDGMENT
This work was kindly supported by the National Natural Science Foundation of China (No.71661012), and Foundation of Jiangxi Educational Committee (No. GJJ160629 and No. GJJ170496).

REFERENCES
[1] X. L. REN, L. Y. Lü, “Review of ranking nodes in complex networks,” Chinese Science Bulletin, vol. 59, no. 13, pp.1175-1197, January 2014.
[2] Y. X. Du, C. Gao, Y. Hu, et al., “A new method of identifying influential nodes in complex network networks based on Topsis,” Physical A:statistical Mechanics and its applications, vol. 399, pp.57-69, April 2014.
[3] William, Sean, Everton, et al., “Mathematical modeling in social network analysis: using Topsis to find node Influences in a Social network,” Math Syst Sci, vol. 10, pp.531-541, October 2013.
[4] H. Yu, Z. Liu, Y. J. Li, et al., “Key nodes in complex networks identified by multi-attribute decision-making method,” Acta Physica Sinica, vol. 62, no. 2, pp. 020204, January 2013.
[5] J. T. Hu, Y. X. Du, H. M. Mo, et al., “A modified weighted TOPSIS to identify influential nodes in complex networks,” Physica A: Statistical Mechanics and its Applications, vol. 444, pp. 73-85, February 2016.
[6] S. B. Wang, J. L. Zhao., “Multi-attribute integrated measurement of node importance in complex networks,” Chaos, vol. 25, no. 11, pp. 113105, November 2015.
[7] C. S. Zhu, X. Y. Wang, et al., “A novel method of evaluating key nodes in complex networks,” Chaos, Solitons & Fractals, vol. 96, pp. 43-50, March 2017.
[8] Y. R. Wang, L. F. Ren, W.Li. Chen, et al., “A novel improved TOPSIS method and its application in medical science,” Journal of Central South University(Medical Science), vol. 38, no. 2, pp.196-201, February 2013.
[9] W. L. Fan, H. Ping, Z. G. Liu, et al., “Multi-attribute node importance evaluation method based on Gini-coefficient in complex power grids,” Iet Generation Transmission & Distribution, vol. 10, no. 9, pp. 2027-2034, March 2016.
[10] F. Hu, Y. H. Liu, “Multi-index algorithm of identifying important nodes in complex networks based on linear discriminate analysis,” Modern Physics Letters B, vol. 29, no. 3, pp. 1277-1286, January 2015.
[11] Affuladzie, A, Affuladzie, E, Turkson, C, et al., “A TOPSIS extension framework for re-conceptualizing sustainability measurement,” Kybernetes, vol. 45, no. 1, pp. 70-86, January 2016.
[12] H. Z. Shaheer, F. H. Daniela, “A bibliometric-based survey on AHP and TOPSIS techniques,” Expert Systems with Applications, vol. 78, pp. 158-181, February 2017.