Theory and Practice of Science and Technology
2025, VOL. 6, NO. 11, 15-20
DOI: 10.47297/taposatWSP2633-456903.20250611

Node Importance Algorithm for Directed Weighted Networks Based on Degree Difference and Standardized K-order Propagation Number

Tao Tang

Hunan University of Science and Technology, Xiangtan, Hunan, 411201, China

ABSTRACT

To overcome the shortcomings of existing key node identification techniques, such as high computational complexity, single evaluation dimension, and limited application scope, a new algorithm for key node evaluation is constructed. This algorithm addresses the need for synergistic evaluation of "node's own role characteristics" and "multi-order information propagation capability" in directed weighted networks. First, the degree difference is used to characterize the disparity between in-degree and out-degree, combined with cross strength to reflect the node's own connection characteristics, thereby quantifying the node's own role and solving the problem that "weight/number of connections alone cannot distinguish in-degree and out-degree differences". Second, the algorithm applies Min-Max normalization to weights to optimize path length, and comprehensively measures the multi-step propagation influence of nodes through multi-order K-order propagation number and structural entropy calculation. The final results show that the algorithm achieves certain improvements in accuracy, effectiveness, and identification capability.

KEYWORDS

Complex networks; Key nodes; Node degree difference; Propagation model

1 Introduction

With the deep penetration of information technology in domains such as social media, transportation, communication, and power systems, real-world complex systems (e.g., microblog public opinion networks, urban rail transit networks, logistics scheduling networks) are commonly abstracted as directed weighted complex networks. In such networks, nodes act as core units whose importance varies significantly—a small number of key nodes may, when failing, cause a sharp decline in network efficiency, interrupt information propagation, or even paralyze the entire system (e.g., shutdown of transportation hubs causing regional congestion, failure of communication core nodes causing signal disruption[1-2]).Therefore, accurately identifying key nodes in directed weighted networks holds significant theoretical and practical value for network robustness optimization, information dissemination control, and efficient resource allocation. It has become a central research direction in complex network studies.

Regarding node importance evaluation, academia has developed multiple classical methods. However, their adaptability and accuracy face challenges as network characteristics become more complex. This paper aims to construct a dual-dimension integrated node importance evaluation algorithm (DCPN) for directed weighted networks, achieving synergistic evaluation of "node's own role characteristics" and "multi-order information propagation capability". It accurately distinguishes nodes with significant in-out degree differences but similar weights or numbers of connections, efficiently quantifies the influence of nodes in global multi-order propagation, balances computational accuracy and efficiency, and adapts to directed weighted networks of different scales.

2 Related Theoretical Foundations

2.1 Network Definition

Consider a directed weighted network $G=(V,E,W)$, where: the node set $V=\{v_1,v_2,\cdots,v_n\}$ is consistent with that of an undirected network, representing individuals or units in the network, and $|V|=n$ is the total number of nodes. The edge set $E=\{(v_i,v_j)|v_i,v_j\in V,i\neq j\}$, where each edge has directionality; (v_i,v_j) denotes a directed edge from v_i to v_j, and (v_i,v_j) and (v_j,v_i) are distinct edges. A weighted directed complex network can be represented by an adjacency matrix, i. e. $W_{n\times n}=(w_{ij})$, where the element w_{ij} represents the weight of the edge from node v_i to node v_j, used to quantify attributes such as relationship strength, cost, or frequency. If (v_i,v_j) does not exist, then $w_{ij}=0$.

2.2 Cross Strength

In directed weighted networks, the strength of each node can be divided into in-strength and out-strength. To comprehensively consider both, Wang Yu[3] proposed cross strength, calculated as follows:

$$SW_i=\theta^*S_i^{in}+\left(1-\theta\right)^*S_i^{out} \tag{1}$$

where θ is the cross-weight factor satisfying$0<\theta<1$, $S_i^{in}=\sum_{j=1}^{n}w_{ji}$denotes the in-strength of node v_i and $S_i^{out}=\sum_{j=1}^{n}w_{ij}$ denotes the out-strength of node v_i. Cross-strength, through weighted fusion, allows for the adjustment of the θ parameter to emphasize the evaluation of different types of nodes, making the assessment more natural and applicable to directed networks.

2.3 Effective Distance

In undirected unweighted networks, Brockmann et al.[4]defined the effective distance between two connected nodes v_i and v_j as $d(i,j)$, calculated as:

$$d\left(i,j\right)=1-\log P_{ji} \tag{2}$$

Where P_{ji} represents the proportion of information flow from node v_i to node v_j In directed weighted networks, according to the similarity weight principle, a larger weight implies a shorter distance between nodes. The distance is usually measured by the reciprocal of the edge weight, but this may lead to excessively small path lengths. This paper uses the reciprocal of the normalized weight to measure distance. The original weight w_{ij} is normalized to the target interval [L,U] as:

$$w_{ij}'=\frac{w_{ij}'-\min(W)}{\max(W)-\min(W)}*\left(U-L\right)+L \tag{3}$$

where $\max\left(W\right)$ and $\min(W)$ are the maximum and minimum values of all edge weights in the network, respectively. The target interval is set to[1,10] in this paper.

2.4 Network Efficiency

Network efficiency[5] is calculated as the average of the reciprocals of the shortest path lengths between all pairs of nodes. It is an indicator of the overall information transmission efficiency of the network, reflecting the speed of information propagation. Higher network efficiency implies faster communication between nodes. Network efficiency is defined as:

$$I=\frac{1}{n(n-1)}\sum_{i\neq j}\frac{1}{d_{ij}} \tag{4}$$

where n is the total number of nodes in the network, and d_{ij} is the shortest path from node v_i to node v_j .

2.5 Mutual Information Method

The Mutual Information method (INF), proposed by Wang Ban et al.[6], uses the node information content $I(i)$ as its importance metric. This metric is defined as the sum of mutual information from node v_i to all nodes it points to, minus the sum of mutual information from all nodes pointing to v_i. The mutual information from v_i to v_j is based on the out-strength of v_i (S_i^{out}) and the in-strength of v_j (S_i^{out}):

$$I(i,j)=ln\frac{S_i^{out}}{S_j^{in}} \tag{5}$$

The final information content of node v_i is:

$$I(i)=\sum_{j\in V_{out}(i)}I(i,j)-\sum_{k\in V_{in}(i)}I(k,i) \tag{6}$$

where $V_{out}(i)$ is the set of nodes pointed to by v_i, and $V_{in}(i)$ is the set of nodes pointing to v_i.

2.6 Multiple Influence Matrix

The method proposed by Wang Yu[3]et al. aims to evaluate node importance in directed weighted networks. This method comprehensively considers the node's local "cross-strength" and a "multiple influence matrix, " which is constructed based on distance and the number of shortest paths, to assess node importance.

$$\mathrm{D}_i=SW_i\times\sum_{j=1,j\neq i}^{n}I_jm_{ji} \tag{7}$$

2.7 JP-degree Centrality

The JP-Degree algorithm, proposed by Zhao Gouheng et al.[7], aims to comprehensively consider the number of neighbors, edge weights, and directionality of nodes in a directed weighted network. This method evaluates node centrality using the geometric mean and arithmetic mean of the node's out-strength and in-strength:

$$C_{WD}^{a}(i)=\sqrt[3]{S_i^{out}\times S_i^{in}\times\frac{1}{2}(S_i^{out}+S_i^{in})} \tag{8}$$

2.8 Method Based on Node Degree Difference and Effective Distanc

The DND algorithm, proposed by Lan Chunjiang et al[8], aims to evaluate nodes in directed weighted networks by integrating "node degree difference importance" and "effective neighbor influence". The former quantifies the node's own in-degree/out-degree disparity and cross-strength; the latter calculates the node's reception and transfer rates based on effective neighbor distance. The calculation formula is as follows:

$$\mathrm{DND}_i = D_i + \mu \sum_{j \in \tau(i)} \{a * ND_i^{in} + (1 - a) * ND_i^{out}\} \tag{9}$$

3 Node Importance Algorithm Based on Degree Difference and Standardized K-order Propagation Number in Directed Weighted Networks

To meet the synergistic evaluation needs of "node's own role characteristics" and "multi-order information propagation capability" in directed weighted networks, this paper constructs a dual-dimension integrated node importance evaluation algorithm (DCPN). The relevant definitions required for the DCPN algorithm are elaborated below.

3.1 Degree Difference Importance

In directed weighted networks, the strength of each node can be divided into in-strength S_i^{in} and out-strength S_i^{out} ,The node degree difference describes the disparity between the in-strength and out-strength of node v_i, calculated as:

$$H_i = \frac{h_i}{h_i + |S_i^{in} - S_i^{out}|} \tag{10}$$

where $h_i = S_i^{in} + S_i^{out}$ is the total weighted strength of node v_i. Global normalization is applied to H_i to obtain the degree difference weight value v_i of node HW_i:

$$HW_i = \frac{H_i}{\sum_{l=1}^{n} H_i} \tag{11}$$

Combining the cross strength of node v_i with its degree difference weight value yields the degree difference importance D_i of node v_i:

$$D_i = HW_i \times e^{\frac{1}{n} \times SW_i} \tag{12}$$

In the equation, HW_i serves as a metric for "balance." A larger HW_i indicates that the node's in-strength S_i^{in} and out-strength S_i^{out} are approaching equality, signifying a more balanced node. Concurrently, a larger SW_i implies a higher cross-strength, meaning greater connection strength. Therefore, D_i is a comprehensive indicator reflecting both the "role balance" and "connection strength" of a node, enabling the effective differentiation between nodes that have identical weights or connection counts but differing in-degrees and out-degrees.

3.2 Propagation Capability Importance

Based on the propagation model in the K-order propagation number algorithm, the send K-order propagation number N_{out}^{K} and receive K-order propagation number N_{in}^{K} are constructed in directed weighted networks as follows:

$$N_{\mathrm{out}}^{K}(v_i) = \sum_{v_j \in V} S_{ij}, S_{ij} = \begin{cases} 0 & d'_{ij} > K \\ 1 & d'_{ij} \le K \end{cases} \tag{13}$$

$$N_{\mathrm{in}}^{K}(v_i) = \sum_{v_j \in V} S_{ji}, S_{ji} = \begin{cases} 0 & d'_{ji} > K \\ 1 & d'_{ji} \le K \end{cases} \tag{14}$$

where d'_{ij} is the shortest path length from v_i to v_j, and S_{ij} is the infection indicator function. When the shortest path length from source node v_i to v_j satisfies $d'_{ij} \le K$, $S_{ij} = 1$, indicating that node v_j is infected; otherwise, $S_{ij} = 0$. The K-order structural entropy H^K of the network is defined based on Shannon's formula:

$$H^K = -\sum_{i=1}^{n} \frac{N_i^K}{\sum_{j=1}^{n} N_j^K} \times \log\left(\frac{N_i^K}{\sum_{j=1}^{n} N_j^K}\right) \tag{15}$$

where$N_i^K = \theta \times N_{\mathrm{out}}^{K}(v_i) + (1 - \theta) \times N_{\mathrm{in}}^{K}(v_i)$ combines the sending and receiving capabilities of the node. The number of infected nodes N^K and the structural entropy H^K increase with the propagation order K. When examining infections at different orders, a weight factor c^K is introduced. The smaller the structural entropy H^K, the larger the weight factor c^K:

$$c^K = 1 - \frac{H^K - \min(H)}{\max(H) - \min(H)} \tag{16}$$

Note where max(H) and min(H) are the maximum and minimum values of all K-order structural entropies, respectively. After normalizing N_i^K, the propagation capability importance is obtained by weighted summation:

$$S_i^K = \frac{N^K - \min(N^K)}{\max(N^K) - \min(N^K)} \tag{17}$$

$$Q_i = \sum_{K=0}^{D} c^K \cdot S_i^K \tag{18}$$

where D is the network diameter, Q_i incorporates both local and global information; a larger Q_i indicates stronger multi-order propagation influence and greater importance of the node.

3.3 Propagation Capability Importance

A fusion coefficient β is introduced to combine normalized D_i and Q_i:

$$I_i = \beta \cdot \frac{D_i}{\max\left(D_i\right)} + (1-\beta) \cdot \frac{Q_i}{\max\left(Q_i\right)} \tag{19}$$

The algorithm balances the node's own role and propagation capability, with a default value of β=0.5 (0<β<1). When β >0.5, more emphasis is placed on the node's own role; when β<0.5, more emphasis is placed on propagation capability.

4 Evaluation of Algorithm Effectiveness

4.1 Validation Based on the Directed Weighted ARPA Network

To validate the effectiveness of the algorithm proposed in this paper, we select the directed weighted ARPA (Advanced Research Project Agency) network for a case analysis. We perform node importance ranking on the ARPA network shown in Figure 1. To facilitate comparison with other literature, we set θ=0.8 to place greater emphasis on the receiving perspective during the evaluation. The obtained results are then compared with four other distinct algorithms.

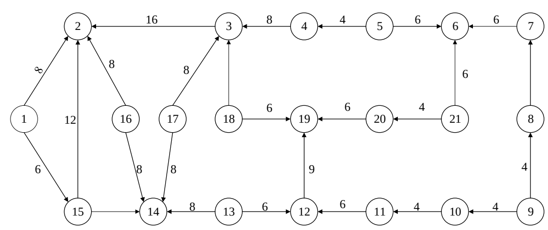

Figure 1 Directed-weighted ARPA network

Table 1 displays the node importance ranking results of the proposed algorithm alongside the INF, MIM (Multiple Influence Matrix), JPD, and DND algorithms in the directed weighted ARPA network. The top three important nodes identified by the proposed algorithm are v_2,v_3,v_{14} . After removing these nodes, the network efficiency decreased by 18.82%,18.19% and 8.33% respectively, consistent with the ranking of the DND algorithm. In contrast, the INF algorithm, which only considers first-order neighbor information, resulted in significantly different node importance rankings. The MIM algorithm provides coarse-grained ranking results. The JPD algorithm, however, fails to adequately consider the influence of neighboring nodes on the node itself. Its calculation scope is confined strictly to the node itself, which consequently leads to discrepancies in the resulting node rankings. The DND algorithm did not account for global information in multi-order propagation, causing the weight parameters of isolated nodes to approach 1, overestimating neighbor influence and ignoring actual propagation capability. Based on the above analysis, the proposed algorithm provides a more reasonable node importance ranking by comprehensively considering node in-out degrees, edge weights, and the impact of multi-order propagation.

Table 1 Node Importance Ranking of the Directed Weighted ARPA Network

Node Importance Ranking of the Directed Weighted ARPA Network											
Sort	*INF*	*MIM*	*SQD*	*DND*	*DCPN*	Sort	*INF*	*MIM*	*SQD*	*DND*	*DCPN*
1	2	2	2	2	2	12	10	4	4	1	4
2	14	14	14	3	3	13	11	20	20	4	7
3	19	3	19	14	14	14	13	16	9	11	5
4	6	19	3	15	19	15	18	17	5	5	17
5	9	6	6	12	12	16	7	1	13	13	18
6	3	12	12	19	6	17	17	13	21	21	1
7	5	15	15	6	10	18	16	18	1	1	13
8	21	11	11	17	11	19	15	5	18	18	20
9	1	7	8	16	9	20	4	21	17	17	21
10	12	8	10	18	8	21	20	9	16	16	16
11	8	10	7	13	15						

4.2 Validation Based on WSIR Model

To further evaluate the capability of the proposed DCPN algorithm in identifying key spreader nodes in directed weighted networks, this paper employs the Weighted Susceptible-Infected-Recovered (WSIR) model for validation experiments. This model is an extension of the traditional SIR (Susceptible-Infected-Recovered) epidemiological model for weighted network scenarios. Its core improvement is the association of the infection rate with network edge weights, which more closely aligns with the information propagation laws of real-world weighted directed networks.

4.2.1 WSIR Propagation Model

In the WSIR model, nodes exist in three states: Susceptible (S), Infected (I), and Recovered (R). At the initial moment, only the pre-selected Top-k nodes are in the infected state; all other nodes are susceptible. Within each time step, infected nodes infect their susceptible neighbors at a weighted infection rate p_{ij}, and simultaneously transition to the recovered state at a fixed recovery probability β (after which they no longer participate in propagation). The process terminates when no infected nodes remain in the network.Key parameters are defined as follows:

Weighted Infection Rate p_{ij}: This characterizes the infection capability of an infected node v_i towards a susceptible node v_j. It is calculated as:

$$p_{ij}=\gamma\cdot\frac{w_{ij}}{\sum_{k\in\Gamma_j^{in}}w_{kj}} \tag{20}$$

where γ is the propagation severity coefficient (set to 1 in this paper to ensure experimental generality), w_{ij} is the weight of the edge from v_i to v_j, and $\sum_{k\in\Gamma_j^{in}}w_{kj}$ is the total weight of all in-edges to node v_j. This definition ensures that a higher edge weight and less competition among in-edges for v_j lead to a higher infection probability

Standardized Infection Rate F(t): This is used to quantify the propagation coverage of the network at different time steps and serves as the core metric for evaluating node importance. It is calculated as:

$$F(t)=\frac{n_{I(t)}+n_{R(t)}}{n} \tag{21}$$

where $n_I(t)$ and $n_R(t)$ are the number of infected and recovered nodes at time step t , respectively, and n is the total number of nodes in the network. A higher $F(t)$ value indicates a stronger propagation influence from the current infection source.

4.2.2 Experimental Setup

The experiments utilized four publicly available real-world weighted directed network datasets: Songbird-Social, Sandi-Auths, USAir, and bio-DM-CX. Songbird-Social: Sourced from empirical observations of social behavior in songbird flocks, modeling directed social interactions. Sandi-Auths: Sourced from author collaborations and citation relationships in academia. Nodes represent authors, and directed edges (with weights) represent directed academic associations and their strength. USAir: Represents the 1997 US aviation infrastructure. Nodes are airports, directed edges are routes, and weights represent operational strength. bio-DM-CX: Represents a biological metabolic regulatory system. Nodes are metabolites or factors, and directed edges (with weights) represent material flow and transfer strength. Its topological characteristics are shown in Table 2:

Table 2 Statistical characteristics of the networks

Network	Number of nodes	Number of edges	Average degree	Average clustering coefficient	Maximum degree
Songbird-Social	110	1027	18	0.6429	56
Sandi-Auths	86	124	2	0.4148	12
USAir	332	2126	12	0.6252	139
bio-DM-CX	4040	76717	37	0.2675	362

4.2.3 Experimental Parameter Settings

At the initial stage, the Top-10 nodes identified by different methods (DCPN, INF, MultiInfluence, DND, JPD) were set to the infected state, with all other nodes susceptible. The WSIR model was run 5000 times for each network to simulate propagation, and the average infection rate at each time step was calculated. The recovery probability was set to 0.5. The experimental results are shown in Figure 2.

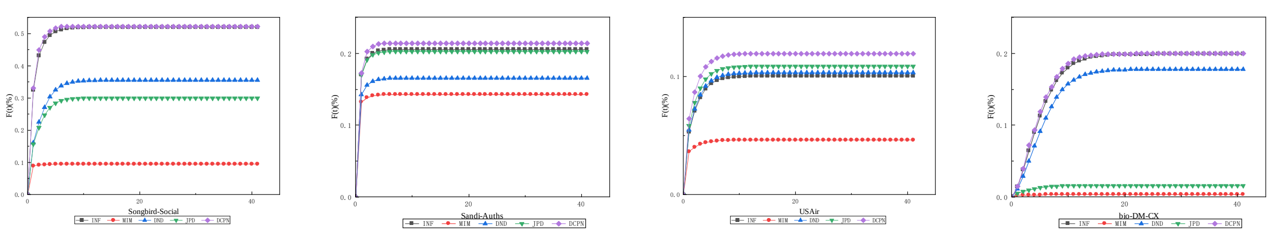

Figure 2 Variation curves of infection rates for different methods in the WSIR model

As shown in Figure 2, the infection rate of nodes identified by the DCPN method is the highest across all networks. In the large-scale bio-DM-CX network, although the infection rate of the DCPN method is close to that of other methods, it consistently remains higher. This indicates that selecting the Top-10 nodes identified by the DCPN method as infection sources in a WSIR simulation achieves the most effective propagation results, thereby validating the efficacy of the DCPN method within a propagation model.

5 Conclusions

Addressing the challenges in identifying key nodes in directed weighted networks—such as high computational complexity, single evaluation dimensions, and the difficulty in effectively distinguishing the roles of in-degree and out-degree—this paper proposes a dual-dimension evaluation algorithm, DCPN, which fuses node-intrinsic roles with multi-order propagation capability.

The core contributions of this algorithm lie in the construction of two distinct evaluation metrics:

First, at the level of node-intrinsic characteristics, this paper constructs the "Degree Difference Importance." By combining the node's cross-strength and degree difference, it effectively quantifies the node's local connection strength and role balance. This addresses the problem where traditional methods, relying solely on weights or connection counts, fail to distinguish nodes with in-degree/out-degree heterogeneity.

Second, at the level of network propagation capability, the paper constructs "Propagation Capability Importance." By applying Min-Max normalization to edge weights to optimize path calculations and integrating K-order propagation numbers with structural entropy, it achieves a comprehensive assessment of a node's multi-step propagation influence, from local to global

To validate the algorithm's effectiveness, this paper first used the ARPA network for case analysis, demonstrating that the DCPN algorithm's ranking is more reasonable than traditional algorithms like INF and DND. Subsequently, WSIR propagation dynamics simulations on four real-world directed weighted networks (including Songbird-Social and bio-DM-CX) further demonstrated that the standardized infection rate $F(t)$ curve for the Top-10 nodes identified by DCPN exhibited the fastest initial growth and consistently remained above all comparative algorithms. This fully validates the precision and superiority of the proposed algorithm in identifying key propagation nodes within a network

References

[1] Liu, YY., Slotine, JJ. & Barabási, AL. (2011). Controllability of complex networks. Nature 473, 167–173.

[2] Wang, J. W., & Rong, L. L. (2009). Cascade-based attack vulnerability on the US power grid. Safety Science, 47(10), 1332–1336.

[3] Wang, Y., Guo, J. L. (2017). Method for evaluating node importance of directed weighted networks based on multiple influence matrices. Acta Physica Sinica, 66(05), 19–30.

[4] Brockmann, D. (2014). The hidden geometry of complex, network-driven contagion phenomena. Science, 343(6172), 730.

[5] Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. Physical Review Letters, 87(19), 198701.

[6] Wang, B., Ma, R. N., Wang, G., & Chen, B. (2015). An improved mutual information method for evaluating node importance in weighted networks. Journal of Computer Applications, 35(07), 1820–1823+1828.

[7] Zhao, G. H., Jia, P., & Zhou, A. M. (2020). An improved degree centrality in directed weighted networks. Journal of Computer Applications, 40 (S1), 141–145.

[8] Lan, C. J., Sun, W. J., Yu, L. F., et al. (2025). Identification of important nodes in directed weighted networks based on degree difference and effective neighbor distance. Complex Systems and Complexity Science, n. vol.(n. iss.), 1–9.