Identifying important nodes affecting network security in complex networks

Yongshan Liu¹, Jianjun Wang¹, Haitao He¹, Guoyan Huang¹ and Weibo Shi²

Abstract
An important node identification algorithm based on an improved structural hole and K-shell decomposition algorithm is proposed to identify important nodes that affect security in complex networks. We consider the global structure of a network and propose a network security evaluation index of important nodes that is free of prior knowledge of network organization based on the degree of nodes and nearest neighborhood information. A node information control ability index is proposed according to the structural hole characteristics of nodes. An algorithm ranks the importance of nodes based on the above two indices and the nodes’ local propagation ability. The influence of nodes on network security and their own propagation ability are analyzed by experiments through the evaluation indices of network efficiency, network maximum connectivity coefficient, and Kendall coefficient. Experimental results show that the proposed algorithm can improve the accuracy of important node identification; this analysis has applications in monitoring network security.

Keywords
Network security, complex network, node importance ranking, K-shell algorithm, structural hole

Introduction
Security risk assessment of nodes is an important area of network security; hence, it is of great value to evaluate the risk and influence of important nodes in complex networks.¹ The key nodes that affect network security are found by analyzing nodes’ characteristic information.² To identify key nodes has remarkable practical value for network security. In debunking rumors, we should find the key figures and cut off the main channels that spread rumors.³,⁴ Network systems, such as communication and power systems, can be attacked through their security vulnerability. After a successful attack, some key hosts in the system cannot work normally. Because of the connection between the hosts in such a system, more hosts will crash, and the whole system will collapse.⁵,⁶ This research is also applied to technical networks (such as the Internet),⁷ transportation networks (e.g. aviation, roads, and public transportation),⁸ and biological networks (e.g. ecological, neural, genetic, and metabolic).⁹

Important nodes in complex networks usually have large information propagation ability and are easily attacked. It is important to evaluate the propagation capability and influence of nodes through the analysis of network topology. There are many methods to evaluate the importance of nodes in complex networks.

¹School of Information Science and Engineering, Yanshan University, Qinhuangdao, China
²Library, Yanshan University, Qinhuangdao, China

Corresponding author:
Jianjun Wang, School of Information Science and Engineering, Yanshan University, Qinhuangdao 066000, Hebei, China.
Email: ysu-wjj@ysu.edu.cn
Betweenness centrality,\(^9\) closeness centrality,\(^1\) and degree centrality\(^2\) are commonly used. While global metrics such as betweenness centrality and closeness centrality can identify important nodes, they cannot be applied to large or unconnected networks, because their calculation results must travel the whole network in advance. Scholars have proposed many kinds of evaluation algorithms for important nodes, in which the word “centrality” commonly arises. A node that is linked to many other nodes may be a center of information propagation. Because a degree centrality algorithm considers local but not global attributes of nodes, it is not ideal to evaluate their importance. Burt et al.\(^1\) proposed a degree centrality expansion method, which adds weights to the links of nodes according to the closeness between them. Wang et al.\(^1\) extended the degree centrality expansion method and applied it to evaluate the importance of nodes in adaptive websites. Some scholars have proposed circuit algorithms\(^5\),\(^1\),\(^8\) that rely on graph theory and related concepts of subgraphs to analyze nodes. Probabilistic models have been used to evaluate the importance of nodes, which can turn the problem into the prediction of node properties.\(^1\) On the basis of the local structure algorithm, researchers such as Battiston et al.\(^1\) evaluated nodes through a supervised framework. This is a simple way to evaluate the importance of nodes. However, its accuracy can suffer because it makes use of very limited information and cannot identify important bridging nodes. Methods have been proposed to reduce the computational complexity and use local information to evaluate key nodes.

The structural hole (SH)\(^1\) is effective and only uses local information to identify key nodes. Burt et al. proposed the SH as a classical sociological theory. This is a gap between nodes with complementary resources or information. When the nodes are connected through a third node, the gap is filled, creating important advantages for the third node. The third node, called an SH node, plays an important role in the propagation of information in a network. Therefore, even if the node is not at the center of the network and its degree is relatively small, its importance is not to be underestimated. We conclude that the more SHs a node has, the more it can control the propagation of information, and the more important the node is. However, the SH method has disadvantages like those of the degree centrality method.\(^2\)\(^,\)\(^3\) It may not be able to accurately quantify the differences between nodes and effectively identify important nodes.

The K-shell algorithm can measure the importance of a node according to its global position in a network, and the time complexity is low; hence, it has been researched and applied in recent years. De Arruda et al.\(^2\) proposed a K-shell decomposition method based on the global attributes of a network, and evaluated the importance of a node from its location. However, the K-shell value is not ideal for influence measurement. Tejedor et al.\(^3\) proposed the K-shell community (KSC) index, which divides influence into internal and external parts. Internal influence can use the degree centrality and K-shell value as a reference basis, while external influence considers the community attributes of nodes. Zareie and Sheikhahmadi\(^1\) proposed a hybrid approach using K-shell values and dispersions of node neighbors in each shell to standardize the expansion capabilities of nodes. Gao et al.\(^3\) proposed a node importance ranking algorithm based on K-shell and local information of nodes. The algorithm integrates the iteration level and degree value of a node based on the K-shell algorithm, but ignores its attributes when calculating its importance, considering only the first- and second-order neighbors, and the effect is not ideal. Zareie and Sheikhahmadi\(^3\) proposed a hierarchical algorithm to influence node ordering in complex networks (hierarchical K-shell, HKS). According to the K-shell decomposition process, two parameters \(b\) and \(f\) were proposed to determine the topological position of a node, where \(b\) determines the extent to which a node is away from the periphery and \(f\) determines its closeness to the core. The influence measurement index of a node is obtained through \(b\) and \(f\) of its second-order neighbor. Although this algorithm solves the problem of high coarse-grained partitioning by the K-shell algorithm, it only reflects the importance of a node by the position information of its second-order neighbor, which is not comprehensive enough.

This article makes the following contributions. We propose an evaluation method for important nodes that affect network security. The method has the three aspects of a node’s global location attribute, local propagation ability, and information control ability, on which we elaborate as follows. (1) We improve the K-shell algorithm. After K-shell decomposition, we further identify the importance of nodes in the same layer according to the time and degree centrality algorithm when each node is deleted, for a more accurate partition result. (2) We take the two-step reachable range of the node as the basis of its local propagation capability and propose a formula to calculate it. (3) We propose an index of node information control ability based on the number of neighbors that can form SHs. We propose a node evaluation algorithm based on the above three aspects. Comparative experiments show that our algorithm has good accuracy and effectiveness in evaluating important nodes.

### Node importance ranking algorithm

#### K-shell decomposition algorithm

The K-shell method is coarse-grained in partitioning the importance of nodes, and nodes on the same layer are considered equally important, which is not the case. Take Figure 1 as an example, where \(K_s\) is a K-shell value that denotes the node’s layer. We find all nodes
in the network of degree 1 set \( A = \{8, 10, 12, 13, 14, 15, 16, 18, 19, 20, 21, 24\} \), delete them and their associated edges with other nodes, and update the network. At this time, we find the set \( B = \{9, 11, 23\} \) of all nodes of degree 1. We delete these nodes and their connected edges with other nodes and update the network. Continuing, we find the node set \( C = \{7\} \) with degree 1 in the new network, delete node 7 and all its edges, and update the network. When there is no node with degree 1 in the network, all the deleted nodes constitute the set of nodes with \( K_s = 1 \). These nodes and the connected edges between them form the 1-shell of the network. K-shell decomposition is performed analogously. According to this decomposition process, it can be seen that the original division is too rough, the important nodes are sorted according to the \( K_s \) value to evaluate their influence magnitude, and the result is not ideal. Many nodes with the same \( K_s \) value have different influences and should not be considered equally important. For example, node 11 has the same \( K_s \) value as nodes 12 to 15, but it can activate more nodes, and is obviously more important. The nodes in sets \( A \), \( B \), and \( C \) have the same layer \( K_s = 1 \), but the deleted time of nodes in sets \( A \), \( B \), and \( C \) is different. The nodes in set \( A \) are deleted first, then the nodes in set \( B \), and finally the nodes in set \( C \). The influence of the first deleted nodes is less than that of the later deleted nodes. The influence of nodes with different deleted time in the same layer should not be the same.

**SH characteristics of network**

Burt et al. proposed SHs as a classical sociological theory. Some SH nodes can obtain higher network benefits. This plays an important role in the dissemination of information. Therefore, even if a node is not at the center of a network and its degree is relatively small, its importance should not be underestimated. Burt pointed out that a node with more SHs can better control the propagation of information; hence, it is more important.

The essence of an SH is a gap between two neighbors of a node. In Figure 2, there is no edge between nodes \( B \) and \( C \), but they are connected through the intermediate node \( H \), which confers cumulative benefits to nodes on either side; hence, node \( H \) has more information control ability. Node \( H \) has a greater impact on network information transmission than the other nodes. If a connected edge exists between nodes \( B \) and \( C \), they can communicate directly, and the information control capability of node \( H \) is weaker.

In summary, when evaluating the importance of a node, it is necessary to consider its information control capability.

**Node importance ranking algorithm**

The K-shell algorithm is coarse-grained in partitioning the importance of nodes; it cannot accurately measure the importance of nodes with the same K-shell value layer. An improved K-shell (IKS) index is proposed to solve this problem. This index considers the K-shell with a wider horizon. The algorithm has the following steps:

1. In the process of K-shell decomposition, where \( K_s \) is the K-shell value and denotes the node’s layer, the nodes on the same \( K_s \) layer are divided into sub-layers according to their time of deletion. We calculate the number of sub-layers into which the nodes on the \( K_s \) layer are divided, and the sub-layer value of each node.
2. We identify the importance of nodes in the same sub-layer according to their degree centrality index.

The IKS index of node \( i \) is

\[
S_i^{IKS} = K_s i + \frac{l_i - 1}{2 \times L_i} + \frac{d_i}{2 \times (n - 1)}
\]
where $K_{Si}$ is the K-shell value, which denotes node $i$’s layer number. For example, all nodes with degree values of 1 are deleted iteratively, and each deleted node $i$’s layer number is 1 in the K-shell algorithm. $L_i$ is the total number of sub-layers of the $K_{Si}$ layer, and $l_i$ is the sub-layer value of node $i$.

The quantities $\frac{i-1}{2 \times L}$ and $\frac{d}{2 \times (n-1)}$ are normalized to the degree values of the sub-layer and the node itself, ensuring that the sum of the two parts cannot be greater than 1. This process solves the issue that the original partition is too coarse of a granulation problem. In Figure 1, by calculating the influence of the IKS and K-shell values of each node in the network, it can be found that in a simple network with 24 nodes, K-shell decomposition can divide the nodes into three levels according to their degree of importance. IKS decomposition can divide the nodes in the network into nine levels. It can be concluded that the partitioning result of the IKS algorithm is more detailed than that of the K-shell algorithm, the influence of the nodes on the same $K_s$ layer can be further divided, and the coarse-grained partitioning of the importance of nodes in the K-shell algorithm is improved. Table 1 shows the decomposition process of the network and the number of iteration sub-layers.

Figure 3 compares two nodes, $a$ and $b$. $S_{IKS}^a$ and $S_{IKS}^b$ obviously have the same values. We cannot get the correct answer with the IKS index of the two nodes. In fact, node $a$ is more important than node $b$, because node $a$ can affect six nodes, while node $b$ can only affect three nodes. A node’s local propagation capability makes a certain contribution to its importance. So, we take the two-step reachable range of a node as the basis of its local propagation capability.

According to the SH theory, it can be known that the information control ability of a node has a certain influence on its importance. In Figure 4, the position of node 3 belongs to the SH, which plays the role of a bridge between nodes 1, 2, and 4. There are three relatively large node sets, so node 3 has an advantage in information control. A structure hole acts as a gap between two neighbors of a node, that is, other nodes can only propagate information through the node. The more SHs a node has, the stronger its information propagation ability.

The node information control ability index (ICAI) is proposed. The ICAI of node $i$ is

$$S_{ICAI}^i = \begin{cases} 0 & d_i = 1 \\ \frac{\sum_{i=1}^{m} |N_i(i)| \cdot \sum_{i=1}^{n} \sum_{i=1}^{n} |N_i(n) \cap N_i(\epsilon_i)|}{d_i \times (n-1)} & d_i > 1 \end{cases}$$

(2)
where \( len \) is the number of neighbors of node \( i \), and there are no links between neighbor nodes; \( N_d(1) \) is a set of neighbor nodes of node \( i \)'s first neighbor node; and \( d_i \) is the degree of node \( i \).

A node whose degree is 1 is at the edge of the network and does not have information control capability, so the ICAI is 0. When the degree value of a node is not 1, we obtain the set of its unconnected neighbors and the number of nodes in this set, and we calculate the sum of non-common neighbors owned by those nodes. The information control ability of node \( i \) is measured by the ratio of its sum value and degree value. Hence, the node’s information control capability is measured by the number of its neighbors that can form an SH. In Figure 4, we use equation (2) to calculate the information control abilities of nodes 1 to 5, respectively, which are \((1 + 5)/4 = 1.5, (6 + 4)/6 = 1.67, (5 + 5 + 3 + 3)/5 = 3.2, (4 + 4 + 4)/4 = 3 \), and \((5 + 2)/5 = 1.4 \). Because node 3 is in the middle of the hub position, its information control ability is strong and its ICAI is largest. Since nodes 1, 2, and 4 are directly connected to node 3, they can acquire information earlier than other nodes from node 3 and occupy the necessary path for information transmitted by their neighbor nodes, so they have stronger information control capability than node 5. The information control capability of node 4 is greater than that of nodes 1 and 2 because it can control the information acquisition and information transmission of all neighboring nodes, and the number of nodes that can be affected is large. The information control capability of node 1 is less than that of node 2, because neighbors of node 2 can only transmit and acquire information through it, and when node 1 transmits information to its neighbor nodes, they can exchange information. The information control capability of node 1 is greatly reduced. In summary, the ICAI can accurately measure the information control capabilities of network nodes.

Through the above analysis, we can conclude that the importance of a node can be evaluated by its global location attribute, local propagation capability, and information control capability. Therefore, an important node identification index for complex networks based on an improved structural hole and K-shell (ISHKS) is proposed as

\[
S_{ISHK}^u = S_{SKS}^u d_u + \alpha \sum_{w \in N_u} d_w + S_{ICAI}^u
\]  

(3)

where \( S_{SKS}^u \) is the global location attribute index of node \( u \), \( d_u \) is the degree of node \( u \), \( N_u \) is a set of node \( u \)'s neighbors (there are no links between the neighbor nodes), and \( S_{ICAI}^u \) is the ICAI of node \( u \).

In equation (3), \( S_{SKS}^u d_u \) considers the global position attribute and degree of the node; \( \alpha \sum_{w \in N_u} d_w \) considers its local propagation capability; we take the two-step reachable range of the nodes as the basis of the node’s local propagation capability; \( \alpha \in (0, 1] \) is a free parameter, whose value depends on the network; and smaller propagation probability of the network assigns less weight \( \alpha \) to local propagation paths. \( S_{ICAI}^u \) considers the information control ability of the node. The important node identification algorithm for complex networks based on an ISHKS decomposition algorithm is described in Algorithm 1.

**Experiment and analysis**

We selected real datasets from three domains for simulation comparison experiments:

1. **Zachary’s Karate Club**: An undirected and unweighted network consisting of 34 nodes and 78 edges, abstractly formed from employees and members of a karate club in the United States, and the relationships between them.
2. **Dolphin**: An undirected and unweighted network of 62 dolphins living in the New Zealand area, and their relationships, containing 62 nodes and 159 edges.
3. **Polbook**: Books about American politics published during the 2004 presidential election and sold on Amazon.com, an abstract network with 105 nodes and 441 edges.

The network structures of the three datasets are shown in Table 2.

---

**Algorithm 1. Important node identification algorithm for complex network security (ISHKS)**

Input: Graph \( G(V, E) \), where \( V \) is the set of nodes and \( E \) is the set of links; degree matrix \( D \); free adjustment parameter \( \alpha \); network size \( n \).

Output: important node sequence \( W \).

BEGIN

1: Initialize the important node index set \( S_{ISHKS} = \{ \} \), node information control ability index set \( S_{ICAI} = \{ \} \), the node’s global location attribute index set \( S_{SKS} = \{ \} \), sum of degrees of the node’s neighbors \( d_{sum} = 0 \)
2: for \( v \in V \)
3: Calculate \( S_{ICAI}^v \) and \( S_{SKS}^v \)
4: end for
5: for \( v \in V \)
6: \( d_{sum} = 0 \)
7: Find the neighbor set \( N \) of node \( v \)
8: for each node \( u \in N \)
9: \( d_{sum} = d_{sum} + D(u) \)
10: end for
11: \( S_{ISHKS} = S_{SKS} \times D(v) + \alpha \times d_{sum} + S_{ICAI}^v \)
12: end for
13: \( W = \text{Sort} \left( S_{ISHKS} \right) \)
14: Return \( W \)

END
The evaluation of a node importance ranking algorithm can be considered from the aspects of the propagation range of nodes and the influence of nodes on network robustness.

The impact of nodes on network robustness can be evaluated by the network efficiency and network maximum connectivity coefficient. The idea is to judge the importance of a node by the impact of removing a node and its edges on the network structure and connectivity.

**Network efficiency**

If the information exchange between nodes in the network is greatly affected by some nodes, then the security of the network is weak. Network efficiency $NE$ measures the efficiency of information exchange between nodes, and we use it to evaluate the accuracy of important node identification algorithms. If, when nodes and their edges are deleted, the shortest path and average path between the other nodes become longer and the propagation efficiency is reduced, then the deleted nodes are important nodes for the network. If we suppose that information is propagated based on the shortest path in the network, then the network efficiency is defined as

$$NE = \frac{1}{N(N-1)} \sum_{i\neq j \in G} \frac{1}{d_{ij}}$$

where $G$ is the set of nodes in the network, and $d_{ij}$ is the shortest path length between nodes $i$ and $j$. We evaluate the accuracy of a node importance ranking algorithm by the relative decline ratio of network efficiency

$$NU = 1 - \frac{ne}{NE}$$

where $ne$ is the network efficiency of $G$ after deleting one or more nodes, and $NE$ is the network efficiency of $G$ before deleting any node. The larger the $NU$ value, the more important the deleted nodes, and the weaker the network security.

**Maximum connectivity coefficient**

If the connectivity of a network is greatly affected by some nodes, then the network security is weak. To evaluate the accuracy of important node identification algorithms, we consider the impact of deleting these important nodes on the maximum connectivity coefficient. The decline ratio of maximum connectivity coefficient $NC$ is defined as

$$NC = \frac{RN}{N}$$

where $RN$ is the number of connected nodes in the network after deleting some nodes, and $N$ is the number of nodes in the network. The smaller the $NC$ value, the more important the deleted nodes, and the weaker the network security.

**Kendall coefficient**

To evaluate the accuracy of important node identification algorithms, we compare the ranking results of important nodes obtained by the algorithms with the real influence of nodes in the network. The higher the consistency, the more reliable the algorithm.

The Kendall coefficient is used to express the correlation between two sets. Suppose sets $X$ and $Y$ each have $n$ elements. $X_i$ and $Y_i$ denote the $i$th values of $X$ and $Y$, respectively. For any pair of variables $(X_i, Y_i)$ and $(X_j, Y_j)$, if $X_i > X_j$ and $Y_i > Y_j$, then the pair of variables is consistent, otherwise they are inconsistent.

Suppose $X$ is a set of ranked important nodes obtained by the algorithms to be evaluated, and $Y$ is a set of ranked important nodes obtained by the real influence of nodes in the network. The Kendall coefficient $\tau$ is used to evaluate the results of the algorithms and is defined as

$$\tau = \frac{N_c - N_d}{\frac{1}{2}n(n-1)}$$

where $N_c$ is the number of consistent node pairs in $X$ and $Y$, $N_d$ is the number of inconsistent node pairs in $X$ and $Y$, $n$ is the number of nodes in the network, and $\frac{1}{2}n(n-1)$ is the total number of node pairs. The value of the Kendall coefficient $\tau$ is between $-1$ and $+1$. A value of $+1$ means that the prediction result is completely consistent with the real situation, and $-1$ means it is completely opposite to the real situation. Therefore, the closer the Kendall coefficient is to $+1$, the more realistic and reliable the result.

| Network         | Node | Edge | Average degree | Network diameter | Average path length |
|-----------------|------|------|----------------|------------------|--------------------|
| Karate Club     | 34   | 78   | 4.588          | 5                | 2.408              |
| Dolphin Network | 62   | 159  | 5.129          | 8                | 3.357              |
| Polbook Network | 105  | 441  | 4.2            | 7                | 3.079              |
The local propagation ability of nodes is easily affected by different structures. In the ISHKS algorithm, we adjust $\alpha \in (0, 1]$ to make our algorithm adapt to different networks. To find a suitable value of $\alpha$, we set $\alpha$ to 0.2, 0.4, 0.6, 0.8, and 1, and obtained a suitable value by experiments. Such a value can maximize the network efficiency and maximum connectivity coefficient. Figure 5(a)–(c) shows the network efficiency of the Dolphin, Polbook, and Karate Club networks, respectively, based on different $\alpha$ values, and Figure 5(d)–(f) shows the corresponding maximum connectivity coefficients. Obviously, when $\alpha = 0.2$, the decline ratio of network efficiency $NU$ and decline ratio of maximum connectivity coefficient $NC$ are optimal in different networks. In summary, $\alpha = 0.2$ is a suitable parameter value.

We compare the ISHKS algorithm with the popular K-shell and HKS algorithms through experiments on different real networks. A certain number of important nodes obtained by the algorithms are deleted from a network, and the decline ratios of the network efficiency $NU$ and maximum connectivity coefficient $NC$ are obtained. We take these ratios as the basis to judge whether the important node results of an algorithm are reliable. Figure 6(a)–(c) shows the network efficiency decline curves of the ISHKS, K-shell, and HKS algorithms on the Dolphin, Polbook, and Karate Club networks, respectively. With the increase of deleted nodes, the ratios ($NU$) of all networks increase gradually. According to the calculation formula of network efficiency decline ratio $NU$, the closer a curve is to the top, the faster the network efficiency will decline, and the greater the impact these deleted nodes will have on the network. We can see that the curve of ISHKS is always above the other curves in the process of node deletion. So, we conclude that ISHKS has the highest accuracy, followed by HKS, and K-shell performs the worst.

After deleting more nodes, the $NU$ value of each algorithm almost does not change, and the three curves tend to coincide. We know that when the curves coincide, the relatively important nodes in the network have been deleted, the network efficiency is almost at its lowest value, and there will be no obvious change. Figure 6(d)–(f) shows the maximum connectivity coefficient decline curves of ISHKS, K-shell, and HKS on the Dolphin, Polbook, and Karate Club networks, respectively. With the increase of deleted nodes, the ratios ($NC$) of all networks decline gradually. According to the calculation formula of the maximum connectivity coefficient decline ratio $NC$, the closer a curve is to the bottom, the faster the maximum connectivity coefficient will decline, and the greater the impact of these deleted nodes on the network. In Figure 6(d)–(f), we can see that the curve of ISHKS is always below the others in the process of node deletion. So, we conclude that the accuracy of ISHKS is best, followed by HKS.
and then K-shell. After deleting more nodes, the NC value of each algorithm almost does not change, and the three curves tend to coincide. We know that at this point, the relatively important nodes in the network have been deleted, the maximum connectivity coefficient is almost at its lowest value, and there will be no obvious change.

The Kendall coefficient is used to evaluate the accuracy of ISHKS, K-shell, and HKS and to analyze the relationship between node local propagation ability and node activation probability.

The experimental results in Figure 7 show that the Kendall coefficients of the ISHKS algorithm are closest to +1 on the three networks. Hence, ISHKS obtains more accurate ranking results of important nodes than the other two algorithms. The Kendall coefficients of ISHKS increase with the node activation probability on the three networks. Because we take the two-step reachable range of the nodes as the basis of a node’s local propagation capability, and the possibility of two-step reachable node activation increases with the activation probability, Kendall coefficients also increase,

---

**Figure 6.** Comparison of ISHKS algorithm with K-shell and HKS algorithms on different networks: (a) the NU of Dolphin network, (b) the NU of Polbook network, (c) the NU of the Karate Club network, (d) the NC of Dolphin network, (e) the NC of Polbook network, and (f) the NC of the Karate Club network.

**Figure 7.** Kendall coefficients of three algorithms on different networks: (a) Dolphin network, (b) Polbook network, and (c) Karate Club network.
as does the algorithm’s accuracy. We also find that when the activation probability of the nodes is smaller and $\alpha = 0.2$, the accuracy of ISHKS is higher. The Kendall coefficient is larger when the node’s activation probability is larger and $\alpha = 0.8$. As the probability of nodes’ activation increases, they can influence more nodes through their neighbors; hence, the local propagation capability of nodes increases. In short, if the activation probability of nodes in the network is large, then $\alpha$ is relatively large.

Conclusion

We studied an evaluation method for important nodes that affect security in complex networks. The IKS index of a node was proposed based on a K-shell algorithm to obtain its global location attribute. We identified the importance of nodes in the same layer according to a time and degree centrality algorithm as each node was deleted, whose results were more accurate than those of K-shell. A node ICAI was proposed according to the SH characteristics of nodes while considering nodes’ neighbors that can form SHs. An important node identification algorithm (ISHKS) for complex networks was proposed based on the above two indices and the nodes’ local propagation ability. The Kendall coefficient, maximum connectivity coefficient, and network efficiency were used as indicators to evaluate the algorithm. In experiments, the ISHKS algorithm had better accuracy and effectiveness than two popular algorithms, and could be applied to various complex networks.

Acknowledgements

The authors gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by the National Natural Science Foundation of China (No. 61972334 and No. 61772449) and National Science Foundation of Hebei Province, China (No. F2017203019).

Data Availability

Some or all data, models, or code generated or used during the study are available from the corresponding author by request.

ORCID iD

Jianjun Wang https://orcid.org/0000-0002-4463-3200

References

1. Zhang Q, Wu X and Liu J. Pinning synchronization of discrete-time complex networks with different time-varying delays. J Syst Sci Complex 2019; 32(6): 1560–1571.
2. Quick D and Choo KKR. Pervasive social networking forensics: intelligence and evidence from mobile device extracts. J Netw Comput Appl 2017; 86: 24–33.
3. Li Y, Deng Y, Xiao Y, et al. Attack and defense strategies in complex networks based on game theory. J Syst Sci Complex 2019; 32(6): 1630–1640.
4. Sotoodeh H and Falahrad M. Relative degree structural hole centrality, CRD–SH: a new centrality measure in complex networks. J Syst Sci Complex 2019; 32(5): 1306–1323.
5. Tran VH, Cheong SA and Bui ND. Complex network analysis of the robustness of the Hanoi, Vietnam bus network. J Syst Sci Complex 2019; 32(5): 1251–1263.
6. Li Q, Guo J, Sun C, et al. Finite-time synchronization for a class of dynamical complex networks with nonidentical nodes and uncertain disturbance. J Syst Sci Complex 2019; 32(3): 818–834.
7. Lü L, Zhou T, Zhang QM, et al. The H-index of a network node and its relation to degree and coreness. Nat Commun 2016; 7(1): 10168.
8. Kitsak M, Gallos LK, Havlín S, et al. Identification of influential spreaders in complex networks. Nat Phys 2010; 7(11): 888–893.
9. Wang W, Tang M, Eugene Stanley H, et al. Unification of theoretical approaches for epidemic spreading on complex networks. Rep Prog Phys 2017; 80(3): 036603.
10. Sheikhahmadi A and Nemathakhsh MA. Identification of multi-spreaders users in social networks for viral marketing. J Inf Sci 2017; 43(3): 412–423.
11. Nasirian F, Pajouh FM and Balasundaram B. Detecting a most closeness-central clique in complex networks. Eur J Oper Res 2020; 283(2): 461–475.
12. Salavati C, Abdollahpouri A and Manbari Z. Ranking nodes in complex networks based on local structure and improving closeness centrality. Neurocomputing 2019; 336: 36–45.
13. Burt RS, Kilduff M and Tasselli S. Social network analysis: foundations and frontiers on advantage. Annu Rev Psychol 2013; 64(1): 527–547.
14. Wang F, Zhang B and Chai S. Deep auto-encoded clustering algorithm for community detection in complex networks. Chinese J Electron 2019; 28(3): 489–496.
15. Dereich S and Mörters P. Emergence of condensation in Kingman’s model of selection and mutation. Acta Appl Math 2013; 127(1): 17–26.
16. Hellbing D. Globally networked risks and how to respond. Nature 2013; 497(7447): 51–59.
17. Wang ZY and Liu GP. Performance improvement of NCSs under complex network via concurrent paths. J Syst Sci Complex 2019; 32(2): 453–478.
18. Battiston S, Farmer JD, Flache A, et al. Complexity theory and financial regulation. *Science* 2016; 351(6275): 818–819.
19. Lazega E. Reviewed work: *structural holes: the social structure of competition* by Ronald S. Burt. *Rev Fr Sociol* 1995; 36(4): 779–781.
20. Zhang ZK, Liu C, Zhan X, et al. Dynamics of information diffusion and its applications on complex networks. *Phys Rep* 2016; 651: 1–34.
21. Zhang RX, Jin Z and Li SP. Epidemic spreading with time delay on complex networks. *Acta Math Appl Sin* 2016; 32(2): 319–326.
22. De Arruda GF, Rodrigues FA and Moreno Y. Fundamentals of spreading processes in single and multilayer complex networks. *Phys Rep* 2018; 756: 1–59.
23. Tejedor A, Longjas A, Foufoula-Georgiou E, et al. Diffusion dynamics and optimal coupling in multiplex networks with directed layers. *Phys Rev X* 2018; 8(3): 031071.
24. Zareie A and Sheikhamadi A. A hierarchical approach for influential node ranking in complex social networks. *Expert Syst Appl* 2018; 93: 200–211.
25. Gao L, Wang W, Shu P, et al. Promoting information spreading by using contact memory. *Europhys Lett* 2017; 118(1): 18001.
26. Zareie A and Sheikhamadi A. EHC: extended H-index centrality measure for identification of users’ spreading influence in complex networks. *Physica A* 2019; 514: 141–155.