

Identification of key recovering node for spatial networks

Zijian Yan(严子健), Yongxiang Xia(夏永祥), Lijun Guo(郭丽君), Lingzhe Zhu(祝令哲), Yuanyuan Liang(梁圆圆), and Haicheng Tu(涂海程)

Citation: Chin. Phys. B, 2023, 32 (**6**): 068901. DOI: 10.1088/1674-1056/acb75f
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

AG-GATCN: A novel method for predicting essential proteins

Peishi Yang(杨培实), Pengli Lu(卢鹏丽), and Teng Zhang(张腾)
Chin. Phys. B, 2023, 32 (**5**): 058902. DOI: 10.1088/1674-1056/acb9f9

AIGCrank: A new adaptive algorithm for identifying a set of influential spreaders in complex networks based on gravity centrality

Ping-Le Yang(杨平乐), Lai-Jun Zhao(赵来军), Chen Dong(董晨),Gui-Qiong Xu(徐桂琼), and Li-Xin Zhou(周立欣)
Chin. Phys. B, 2023, 32 (**5**): 058901. DOI: 10.1088/1674-1056/ac8e56

Analysis of cut vertex in the control of complex networks

Jie Zhou(周洁), Cheng Yuan(袁诚), Zu-Yu Qian(钱祖燏), Bing-Hong Wang(汪秉宏), and Sen Nie(聂森)
Chin. Phys. B, 2023, 32 (**2**): 028902. DOI: 10.1088/1674-1056/aca208

Vertex centrality of complex networks based on joint nonnegative matrix factorization and graph embedding

Pengli Lu(卢鹏丽) and Wei Chen(陈玮)
Chin. Phys. B, 2023, 32 (**1**): 018903. DOI: 10.1088/1674-1056/ac6867

Analysis of identification methods of key nodes in transportation network

Qiang Lai(赖强) and Hong-Hao Zhang(张宏昊)
Chin. Phys. B, 2022, 31 (**6**): 068905. DOI: 10.1088/1674-1056/ac4a6c

---

Identification of key recovering node for spatial networks

Zijian Yan(严子健)[1], Yongxiang Xia(夏永祥)[1,†], Lijun Guo(郭丽君)[2], Lingzhe Zhu(祝令哲)[1], Yuanyuan Liang(梁圆圆)[1], and Haicheng Tu(涂海程)[1]

[1] *School of Communication Engineering, Hangzhou Dianzi University, Hangzhou* 310018, *China*
[2] *China Intelligent Transportation Systems Association, Beijing* 100160, *China*

(Received 12 October 2022; revised manuscript received 19 January 2023; accepted manuscript online 31 January 2023)

Many networks in the real world have spatial attributes, such as location of nodes and length of edges, called spatial networks. When these networks are subject to some random or deliberate attacks, some nodes in the network fail, which causes a decline in the network performance. In order to make the network run normally, some of the failed nodes must be recovered. In the case of limited recovery resources, an effective key node identification method can find the key recovering node in the failed nodes, by which the network performance can be recovered most of the failed nodes. We propose two key recovering node identification methods for spatial networks, which are the Euclidean-distance recovery method and the route-length recovery method. Simulations on homogeneous and heterogeneous spatial networks show that the proposed methods can significantly recover the network performance.

Keywords: complex networks, spatial networks, congestion, key recovering node

PACS: 89.75.–k, 89.75.Fb, 89.75.Hc, 89.40.–a **DOI:** 10.1088/1674-1056/acb75f

1. Introduction

Many complex systems in the real world can be abstracted into complex networks, such as transportation networks,[1] power grids,[2] Internet,[3] and social networks.[4] Recently, researchers have studied various aspects of complex networks, including network topology,[5–8] robustness,[9–14] network propagation characteristics,[15–18] and node importance.

Resilience is used to measure the ability of a system to absorb, adapt, and recover from real or potentially destructive events.[9,19] For complex networks, the whole resilience process includes two stages: robustness to failure propagation and recovery after failure occurrence. Recently, the focus of research has gradually shifted from study of network robustness to research of the whole process of network resilience.[20] In the study of resilience, it is important to distinguish three networks. The first is the original network, i.e., the network before the failure occurs. The second is the damaged network, which is the survived network after the failure propagation. The last is the recovered network, which is the network after the recovery process. Network recovering process is an important stage to study network resilience, and it is related to the damaged network and recovered network. In the damaged network, we call all the nodes which fail during the failure propagation the failed nodes. As these failed nodes do not work anymore, the network performance will be reduced. Then, the key recovering node needs to be found quickly and effectively to recover network performance. In the case of limited recovery resources, i.e., assuming that only one node is allowed to be recovered, the key recovering node identification method determines which node is recovered first to achieve the best recovery effect. Even if the recovery resources are sufficient, the recovery of nodes has a sequential order, and the key recovering node identification method can determine which node to recover first. In summary, an effective key node identification method can recover the damaged network performance to the maximum extent.

There has been a lot of research on key node identification methods.[21–31] Researchers often use node centrality methods to measure the importance of nodes, such as degree centrality, betweenness centrality, and closeness centrality.[21] Based on these methods, some researchers have proposed new methods by combining different centrality metrics or improving them. In Ref. [23], Brandes *et al.* proposed a new key node identification method, which combines the degree centrality, betweenness centrality and closeness centrality of nodes, indicating that the importance of nodes is not determined only by a single indicator. In Ref. [24], Liu *et al.* improved the betweenness centrality indicator, and further combined with the closeness centrality indicator to comprehensively evaluate the importance of nodes. In Ref. [25], Tian *et al.* added the degree centrality of the node to the degree centrality of its neighbor nodes as the key node evaluation indicator, and the results show that it is better than traditional centrality indicators. In Ref. [26], Lin *et al.* proposed a multi-attribute node importance decision method based on entropy.

In the above-mentioned research on key node identification methods, most of them are applied in cascading failure scenarios (i.e., the stage of the network is damaged), and the failure of nodes can cause a decline in the performance of the initial network. However, in the damaged network, prioritizing which node to recover will result in recovery of network performance to the maximum extent. It is generally believed that if the failure of a node will make the network performance decrease mostly, then recovering that node increases the network

[†]Corresponding author. E-mail: xiayx@hdu.edu.cn
© 2023 Chinese Physical Society and IOP Publishing Ltd

performance mostly. However, this is not the case in practice. This is because the initial network is not the same as the damaged network, and it is not suitable to apply the key node identification methods in the former to the latter. Therefore, it is necessary to design a more effective key recovering node identification method for damaged networks. It is also worth noting that most studies on key nodes only consider the topological properties of the network, without considering the spatial attributes of the network. In fact, many complex networks have spatial attributes, such as the location of nodes and the length of edges, which are called spatial networks. We should take spatial attributes into account in the design of the key recovering node identification method for spatial networks.

In this paper, we propose two key recovering node identification methods for spatial networks, namely the Euclidean-distance (ED) recovery method and the route-length (RL) recovery method, which focus on the recovering process of damaged networks. The ED method considers the location of nodes (i.e., coordinates) in the spatial network. By calculating the Euclidean distance between failed nodes and the congestion node (i.e., the node which is easiest to become congested in the damaged network), it finds the failed node closest to the congestion node as the key recovering node. The RL method considers not only the location information of failed nodes and the congestion node, but also the routing information between nodes. By calculating the path length from failed nodes to the congestion node, it finds the failed node with the minimum path length as the key recovering node. In both the proposed methods, the spatial properties are important factors in the calculation. In the ED method, we need the location of nodes to calculate the Euclidean distance between failed nodes and the congestion node. In the RL method, we need the location of nodes to calculate the length of edges and the path length. By using these spatial factors, the two proposed methods can significantly improve the network recovery performance, and their recovery effects are significantly better than the traditional key node identification methods.

The rest of this paper is organized as follows. In Section 2 we introduce two spatial network models. In Section 3, we describe the load transmission model, including the routing strategy and network congestion. In Section 4, we present five traditional key node identification methods, and propose two new key recovering node identification methods and their performance indicators. The simulation results are presented and discussed in Section 5. Section 6 summarizes the conclusions of this paper.

2. Network models

Many complex networks in the real world have spatial attributes, such as the location of nodes and the length of edges, called spatial networks. In this paper, we consider two spatial network models: random geometric graph (RGG)[32] and local-area energy-efficient (LAEE)[33] evolution model. The former represents homogeneous spatial network, and the latter is a heterogeneous spatial network with scale-free property. To test the effectiveness of the proposed two methods, we perform simulations on homogeneous and heterogeneous spatial networks, respectively, to show the universality of the proposed methods.

2.1. Random geometric graph (RGG)

RGG is the simplest spatial network model. In this model, N nodes are randomly distributed in the square region $[0,1]\times[0,1]$. Each node is assigned random coordinate (x_i, y_i) $(i = 1,2,\ldots,N)$ and a uniform connection radius r. As shown in Fig. 1, each node randomly connects nodes within its connection radius r until its degree reaches a predefined upper bound or until all nodes within the r region have been connected.

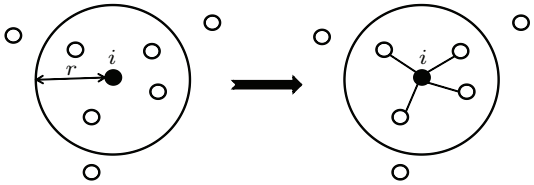


Fig. 1. The connection process of node i in RGG.

2.2. Local-area energy-efficient (LAEE) evolution model

Jiang *et al.*[33] provide an evolution model to generate a spatial network with scale-free property, called the LAEE evolution model. We use this model to build heterogeneous spatial networks.

The LAEE model generation process is mainly divided into two stages.

Stage 1: All nodes are distributed randomly in the region S, and each node has the same connection radius r. Scattered nodes are these nodes not connected to the network. The potential neighbor nodes of the node are those scattered nodes within its connection radius.

Stage 2: The network will evolve according to the following rules:

(i) Choose the node closest to (0,0) as the sink. The network topology evolution starts with m_0 nodes closest to the sink. Connect them with e_0 random links.

(ii) At each time step, a scattered node is connected to the network, as shown in Fig. 2. First, we find the node that is connected in the network and has the most scattered neighbor nodes, denoted as v. Then randomly select one of its scattered potential neighbors, denoted by s.

(iii) Randomly choose m nodes, which are already connected in the network and are potential neighbors of node s, If the number of potential neighbor nodes of node s is less than m, all these nodes will be connected to node s. Otherwise, connecting these nodes to node s with the probability Π_i,

$$\Pi_i = \Pi_i'\,(i \in \text{local area})\frac{f(E_j)\cdot k_i}{\sum_{\text{local-area}} f(E_j)\cdot k_j - q\cdot k_{\max}}, \quad (1)$$

where local area is the set of potential neighbor nodes of node s within its transmission range, $k_{\max}$ is the default upper bound of the predefined node degree, q is the number of nodes that have the degree of $k_{\max}$, and $f(E)$ is a function. For simplicity, $f(E_i) = 1$ is set here.

(iv) Repeat (ii) and (iii) until all nodes are added to the network.

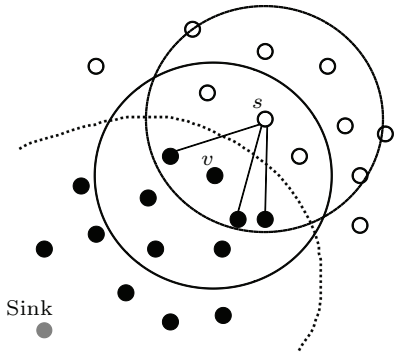


Fig. 2. The topology evolution process of LAEE. Here s represents a scattered node to be connected to the network.

3. Load transportation

3.1. Routing strategy

In spatial networks, each node has a spatial location, so the length of the edge between node i and node j can be calculated by Euclidean metric, defined as

$$d(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \tag{2}$$

where (x_i, y_i) and (x_j, y_j) represent the coordinates of nodes i and j, respectively.

For a routing strategy, a path between node s and node t can be denoted by

$$P(s \to t) : s \equiv x(0), x(1) \ldots, x(n) \equiv t, \tag{3}$$

where $x(0)$, $x(1)$,..., $x(n)$ represent nodes in the path. The length of a path is the sum of the lengths of all the edges along the path, denoted as

$$L(s \to t) = \sum_{i=0}^{n-1} d\left(x(i)x(i+1)\right), \tag{4}$$

where $d(x(i)x(i+1))$ is the length of the edge between nodes $x(i)$ and $x(i+1)$. Among these paths, the path with the smallest length is called the shortest length path, and its length is called the shortest path length. The routing strategy using the shortest length path is called the shortest path length routing,[34] which is considered in this paper.

3.2. Network congestion

We assume that the network generates R units of load at each time step. Their sources and destinations are randomly assigned. For a given pair of source and destination, the load is transmitted between them according to the shortest path length routing strategy. When the load reaches its destination node, it is automatically deleted from the network. The maximum load that each node can handle is called the capacity of node. We refer to the Motter–Lai model[35] and define that the capacity of the node is proportional to its initial load. The initial load is represented by the betweenness of the node, defined as

$$B(i) = \sum_{s \neq t} \frac{g_{st}^i}{g_{st}}, \tag{5}$$

where g_{st} is the number of shortest paths from node s to node t, and g_{st}^i is the number of shortest paths from node s to node t that pass through node i. The capacity of the node is defined as

$$C_i = (1+\alpha) \cdot B(i), \tag{6}$$

where the constant α ($\alpha \geq 0$) is the tolerance parameter. For simplicity, here we set $\alpha = 1$. If the load on a node exceeds its capacity, the node will be congested.

When R is small, all loads can be processed in time, and the network is in a state called free flow. If R gradually increases, some nodes need to process loads greater than their capacity, which will result in these loads not being processed in time. In this case, network congestion occurs. In the initial network, the node capacity is proportional to the initial betweenness of the node, and all nodes in the network will be congested at the same time. However, in the damaged network, the node's betweenness changes, and the node capacity is not proportional to the node betweenness anymore. Therefore, there must be a node congested first, which is called the congestion node in the damaged network.

The R_c is defined as the maximum value of R that makes the network work without congestion. The value of R_c can reflect the maximum capability of the network processing load, called the maximum throughput of the network. However, changes in the network topology can cause R_c to change. For example, in the damaged network after some nodes in the initial network fail due to an attack, the value of R_c decreases; when a node is recovered from the failed nodes (called recovered network), the value of R_c increases.

4. Methods to identify the key recovering node

4.1. Key recovering node

In fact, many networks in the real world are subject to random or deliberate attacks, which cause some nodes in the network to fail. In this paper, we assume that some randomly selected nodes fail. The resulting network is called the damaged network. In this way, the maximum throughput of the damaged network is less than that of the initial network. Therefore, our goal is to recover a failed node to improve the network throughput to the maximum extent. Obviously, recovering different nodes may have different recovery effects. In

the case of limited recovering resources, the question we are interested in is which one should be chosen as the recovering node in order to recover the network throughput mostly. This node is called the key recovering node.

4.2. Traditional methods to identify the key recovering node

There have been many studies on identifying the key node in complex networks. In this part, we introduce five traditional methods to identify the key node.

4.2.1. Degree centrality (DC)

The degree of node i is the number of nodes connected to it, denoted as k_i. DC of node i is defined as

$$\mathrm{DC}(i) = \frac{k_i}{N-1}, \tag{7}$$

where N is the number of nodes in the network. The greater the node degree is, the more important the node is, which is the simplest method to measure the importance of a node. In the scenario of recovery, the failed node with the largest DC will be recovered.

4.2.2. Betweenness centrality (BC)

BC can characterizes the ability of a node to transmit load along the path determined by the routing strategy. Therefore, the larger the betweenness of a node, the greater its impact on the load transmission, which means that the node is more important. BC of node i is defined as Eq. (5). By using this method, the node with the largest BC in failed nodes will be recovered.

4.2.3. Closeness centrality (CC)

CC represents the topological distance between the node and other nodes. The larger the CC, the more central the node is in the network. CC of node i defined as

$$\mathrm{CC}(i) = \frac{N}{\sum_{j=1}^{N} d_{ij}}, \tag{8}$$

where d_{ij} is the number of hops from node i to node j. By using this method, the node with the largest CC in failed nodes will be recovered.

4.2.4. The node importance based on degree and the importance of lines (DIL)

Liu *et al.*[36] proposed an effective ranking method based on degree value and the importance of lines (i.e., edges). Firstly, the importance of an edge is measured by the attributes of the nodes connected to it. Then, the node's contribution to the importance of the edge is calculated. Finally, the importance of nodes is ranked by considering the nodes' degree and the nodes' contribution to the importance of the edges.

Suppose that the network $G = (V,E)$ is an undirected and unweighted network with $N = |V|$ nodes and $M = |E|$ edges. $V = [v_1, v_2, \ldots, v_N]$, $E = [e_1, e_2, \ldots, e_M]$. The edge between node v_m and node v_n is e_{mn}. Its importance is defined as[36]

$$I_{\mathrm{emn}} = \frac{U}{\lambda}, \tag{9}$$

where $U = (k_m - p - 1) \cdot (k_n - p - 1)$ reflects the connectivity ability of edge e_{mn}, k_m is the degree of node v_m, k_n is the degree of node v_n, p is the number of triangle, and one edge of the triangle is e_{mn}; λ is alternative index of edge e_{mn}, which is defined as $\lambda = p/2 + 1$.

In order to get a more reasonable result, they introduce $W_{v_i v_j}$

$$W_{v_i v_j} = I_{eij} \cdot \frac{k_i - 1}{k_i + k_j - 2}, \tag{10}$$

where $W_{v_i v_j}$ stands for the contribution that node v_i makes to the importance of edge e_{ij}, k_i is the degree of node v_i, k_j is the degree of node v_j, node v_j is the neighborhood node of node v_i. Finally, the importance of node v_i can be defined as

$$L_{vi} = k_i + \sum_{v_j \in \Gamma_i} W_{v_i v_j}, \tag{11}$$

where Γ_i is the set of neighbors of node v_i. The larger the L_{vi} is, the more important the node v_i is.

By using this method, the node with the largest L_{vi} in failed nodes is recovered.

4.2.5. Improved structural hole method (ISH)

A structural hole (SH) can be understood as a gap between two unconnected nodes. When the two nodes are connected by the third node, the gap is filled and an important advantage is created for bridging nodes. A bridging node can access different information flows and gain more benefits from its neighbors. To measure the structural hole in the network, Burt introduces the network constraint index,[37] the higher the network constraint index is, the smaller the structural hole becomes. Conversely, the lower the network constraint index is, the more important the node's location is. The network constraint index of node i is defined as

$$\mathrm{NC}(i) = \sum_{j \in \Gamma(i)} \left(p_{ij} + \sum_{b \in \Gamma(i)} p_{ib} p_{bj} \right)^2, \tag{12}$$

where $i \neq b \neq j$, $\Gamma(i)$ represents the set of neighbor nodes of node i, $p_{ij} = z_{ij} / \sum_j z_{ij}$ is the ratio of the network time and energy of node i to its neighbor node j, and z_{ij} is the connection strength between node i and node j. Usually, in an undirected network, $z_{ij} = 1$, $p_{ij} = z_{ij} / \sum_j z_{ij} = 1/k_i$.

However, the SH method cannot quantify the node differences in the network well. The definition of p_{ij} does not take into account the difference in the strength of controllability between different neighbor nodes. Based on this, Yu *et al.*[38] proposed the improved structural hole (ISH) method in which the time or energy that node i invests in its neighbor

nodes should vary according to their location in the network. The specific steps are as follows:

(i) The edge's weight is defined as $w_{ij} = k_i + k_j$.

(ii) The node's weight is defined as $w(i) = \sum_{j\in\Gamma(i)} w_{ij}$, which is the sum of the weights of all edges connected to node i.

(iii) The relative importance is defined as $p(i,j) = w_{ij}/w(i)$, indicating the relative importance of the neighbor node j to node i. Here $p(i,j)$ is the proportion of network time and energy of node i to its neighbor node j. If the degree of neighbor node j is large, node i will be more inclined to invest more energy in node j.

(iv) Node importance $\mathrm{NC}'(i)$ is defined as

$$\mathrm{NC}'(i) = \sum_{j\in\Gamma(i)} \left(p(i,j) + \sum_{b\in\Gamma(i)} p(i,b)p(b,j) \right)^2. \quad (13)$$

Here $p(i,j)$ replaces p_{ij} in Eq. (12).

The smaller the $\mathrm{NC}'(i)$ is, the more important the node i is. By using this method, the failed node with the smallest value of $\mathrm{NC}'(i)$ will be recovered.

The above five traditional methods are used as comparison methods in this paper to verify the effectiveness of the proposed two new methods.

4.3. Proposed methods to identify the key recovering node

As defined in Subsection 3.2, the congestion node is the node which congests first in the damaged network. Therefore, we try to find one node among all failed nodes and recover it to help bypass the load on the congestion node. To do that, the chosen key recovering node should be close to the congestion node. That is to say, the spatial property should be the fatal factor to find the key recovering node. The five traditional methods mentioned above neither focus on the recovering process of the network, nor consider the spatial information (i.e., the location of nodes and the length of edges). Based on the above considerations, this paper proposes two key recovering node identification methods for the recovering process of the damaged network, which consider the spatial relationship between the congestion node and failed nodes.

4.3.1. Euclidean distance (ED) recovery method

In spatial networks, each node has its own location (i.e., coordinate). The physical distance between any two nodes can be represented by the Euclidean distance between them. The key point of ED recovery method is to find the failed node with the shortest Euclidean distance to the congestion node as the key recovering node. The method is defined as

$$d_E^m = \min_i \left(\sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \right), \quad (14)$$

where (x_i, y_i) and (x_m, y_m) are the coordinates of the failed node i and the congestion node m. That is to say, the key recovering node should be the failed node which has the shortest Euclidean distance to the congestion node. That is why we denote it as the ED method.

4.3.2. Route length (RL) recovery method

In spatial networks, the length of each edge is measured by the Euclidean distance between two nodes connecting the edge. The shortest length path between any pair of nodes is defined as the path that minimizes the sum of the lengths of all edges from the source node to the destination node, and its length is called the shortest path length. In this way, we can calculate the shortest path length from a failed node i to the congestion node m, denoted as d_R^{im}. In the RL recovery method, the key recovering node is the failed node which minimizes d_R^{im} among all the failed nodes. In this method, finding the key recovering node not only considers the location of failed nodes and the congestion node, but also needs the routing information.

4.4. Performance indicators for identification methods

In this study, the key recovering node is the node that can make the network performance improve mostly. Here, we use two indicators to measure how good a method can find the key recovering node.

4.4.1. Indicator 1: The maximum throughput of the recovered network R_{c2}

The main purpose of recovering is to improve the network performance of the damaged network. The maximum throughput of the network can directly reflect the transmission performance of the network. An effective key node identification method can improve the maximum throughput of the damaged network mostly. Therefore, we use the maximum throughput of the recovered network, denoted as R_{c2}, to evaluate the recovery effect of identification methods. That is, the larger the R_{c2}, the better the recovery effect of the identification method.

4.4.2. Indicator 2: The maximum throughput growing rate G_r

The recovery effect of different methods can also be measured by the relative growing rate of the network maximum throughput after recovering. We denote the maximum throughput of the damaged network as R_{c1}, then the maximum throughput growing rate can be defined as

$$G_r = \frac{|R_{c2} - R_{c1}|}{R_{c1}}. \quad (15)$$

5. Simulations

5.1. Parameter setting

In this study, in order to test the effectiveness of our proposed methods, we make simulations on the RGG and LAEE

networks, respectively, in which the former is a homogeneous spatial network and the latter is a heterogeneous spatial network. In the above two network models, we assume that all nodes are randomly distributed in a 1×1 square region, the connection radius of each node is 0.2 (i.e., $r = 0.2$). In the following simulations, we randomly select 20 nodes in the initial network to fail to obtain the damaged network, and we assume that the recovery resources are limited, that is, only one node will be recovered in the damaged network.

5.2. Results and analysis

Figure 3 shows the comparison of recovery effect of different identification methods in the RGG network. In each subgraph, we find that the performance of the two proposed methods overlap. This is because in homogeneous spatial networks, the key recovering nodes found by these two methods are almost the same. In addition, we find that the maximum throughput of the recovered network by using five traditional methods are almost identical to that of the damaged network. That is to say, in homogeneous spatial networks, the recovery by using five traditional methods is almost useless. As the network size N increases, the recovery effects of our proposed methods remain the best, much better than the other five traditional key node identification methods.

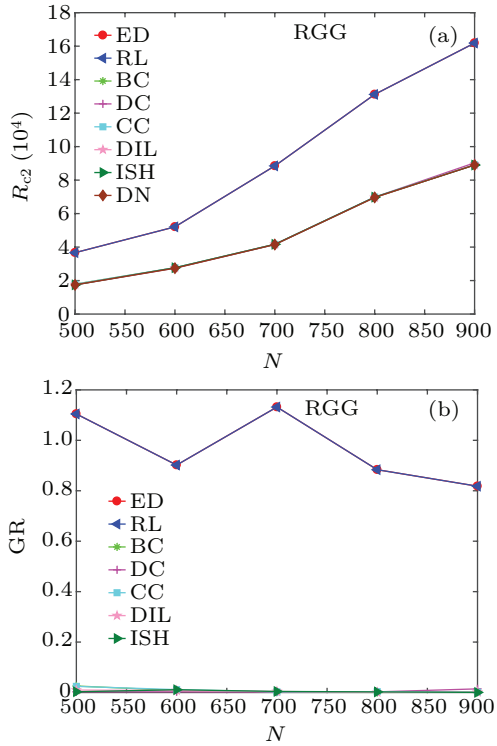


Fig. 3. The comparison of recovery effects of different identification methods in the RGG. (a) The maximum throughput of the recovered network. (b) The maximum throughput growing rate (GR) of the recovered network. DN shows the performance of the damaged network (i.e., without any recovery). The results are the average of 50 realizations.

Figure 4 shows the comparison of recovery effect of different identification methods in the LAEE network. In Fig. 4(a), the maximum throughputs of the recovered network by using five traditional methods are greater than the maximum throughput of the damaged network, and in Fig. 4(b), the maximum throughput growing rates of the recovered network by using five traditional methods are all greater than 0. That is, in heterogeneous spatial networks, the recovery effects of five traditional methods are effective to a certain extent. The performances of five traditional methods are not the same, meaning that they select different failed nodes as the key recovering node. This fact indicates that the structural characteristics of LAEE network are much complicated than RGG network. Even though, the recovery effects of ED and RL are still much better than those five traditional methods, and the recovery effect of RL is better than that of ED.

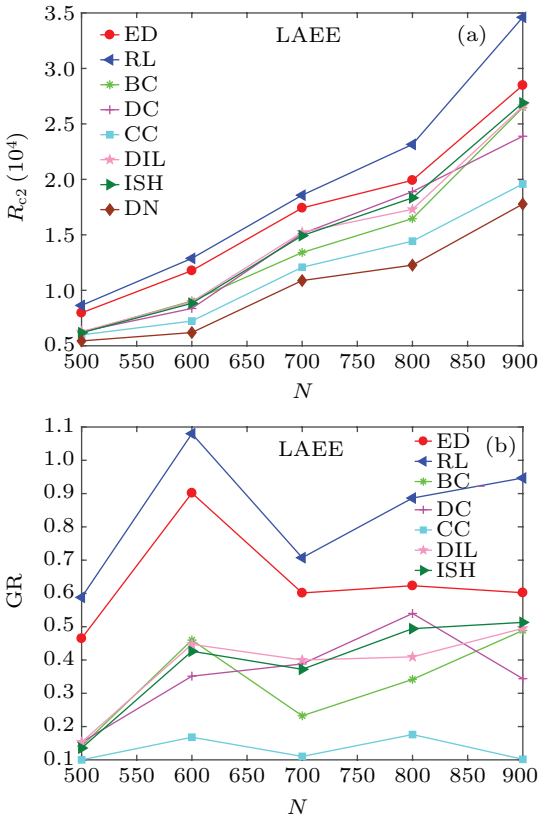


Fig. 4. The comparison of recovery effect of different identification methods in the LAEE. (a) The maximum throughput of the recovered network. (b) The maximum throughput growing rate of the recovered network. The results are the average of 50 realizations.

The recovery effects of five traditional methods are quite different in Figs. 3 and 4. This is because LAEE is a heterogeneous network, and the degree and betweenness of its nodes are significantly different, which leads to different key recovery nodes found by these methods, so that their recovery effects are different.

The above simulation results verify the effectiveness of the proposed methods in spatial networks. ED finds the key recovering node only based on the location information of all failed nodes and the congestion node, and does not need to obtain the routing information between nodes. In contrast, RL

not only needs to obtain the location information of all failed nodes and the congestion node, but also needs to obtain routing information between nodes. In other words, RL requires more information than ED to achieve higher recovering performance. Therefore, in practical applications, if sufficient information can be obtained, the RL will be a good choice, as it can improve the network performance mostly. ED may become the solution when the routing information is unavailable or the calculation of it is too time-consuming.

6. Conclusion

In this paper, we propose two key recovering node identification methods, namely ED and RL, which not only consider the recovering process of the network, but also consider the spatial attributes of spatial networks. The simulations on homogeneous and heterogeneous spatial networks show that the recovery effects of ED and RL are much better than the traditional key node identification methods. The recovery effect of RL is better than ED in heterogeneous spatial networks. However, the former not only needs to obtain the location information of failed nodes and the congestion node, but also needs to obtain the routing information between nodes, while the latter only needs the node location information. Therefore, in practical applications, we can choose ED or RL according to the available information. This study can provide a simple and effective tool for the recovering process of spatial networks.

Acknowledgements

Project supported by Zhejiang Provincial Natural Science Foundation of China (Grant No. LQ23F030012), and the Fundamental Research Funds for the Provincial Universities of Zhejiang (Grant No. GK229909299001-018).

References

[1] Feng X, He S W and Li Y B 2019 *Transportmetrica A: Transport Sci.* **15** 1825
[2] Pagani G A and Aiello M 2013 *Physica A* **392** 2688
[3] Gan C Q, Yang X F, Liu W P, Zhu Q Y, Jin J and He L 2014 *Commun. Nonlinear Sci. Numer. Simul.* **19** 2785
[4] Freeman L C 1978 *Social Networks* **1** 215
[5] Ganesh A, Massoulié L and Towsley D 2005 *IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, March 13–17, 2005, Miami, USA, p. 1455
[6] Lin J Y and Ban Y F 2013 *Transport Rev.* **33** 658
[7] Boss M, Elsinger H, Summer M and Thurner 4 S 2004 *Quant. Finance* **4** 677
[8] Fang P J, Zhang D M and He M H 2015 *Chin. Phys. Lett.* **32** 88901
[9] Dong G G, Wang F, Shekhtman L M, Danziger M M, Fan J F, Du R J, Liu J G, Tian L X, Stanley H E and Havlin S 2021 *Proc. Natl. Acad. Sci. USA* **118** e1922831118
[10] Xia Y X, Wang C, Shen H L and Song H N 2020 *Physica A* **559** 125071
[11] Liang Y Y, Xia Y X and Yang X H 2022 *Physica A* **591** 126800
[12] Tu H C, Xia Y X, Iu H H and Chen X 2018 *IEEE Trans. Circuits Syst. II: Express Briefs* **66** 126
[13] Wu J, Liu Y and Ruan Q 2022 *Nonlinear Analysis: Hybrid Systems* **46** 101216
[14] Han W T, Yi P and Ma H L 2019 *Acta Phys. Sin.* **68** 186401 (in Chinese)
[15] Li S B, Wu J J, Gao Z Y, Lin Y and Fu B B 2011 *Acta Phys. Sin.* **60** 050701 (in Chinese)
[16] Cowen L, Ideker T, Raphael B and Sharan R 2017 *Nat. Rev. Genet.* **18** 551
[17] Yang P, Li X L, Wu M, Kwoh C K and Ng S K 2011 *PloS One* **6** e21502
[18] Liu Y, Sanhedrai H, Dong G, Shekhtman L M, Wang F, Buldyrev S V and Havlin S 2021 *Natl. Sci. Rev.* **8** nwaa229
[19] Righi A W, Saurin T A and Wachs P 2015 *Reliability Engineering and System Safety* **141** 142
[20] Liu T, Bai G, Tao J, Zhang Y A, Fang Y and Xu B 2022 *Reliability Engineering and System Safety* **226** 108663
[21] Liu J G, Ren Z M, Guo Q and Wang B H 2013 *Acta Phys. Sin.* **62** 178901 (in Chinese)
[22] Goh K, Oh E, Kahng B and Kim D 2003 *Phys. Rev. E* **67** 017101
[23] Brandes U, Borgatti S P and Freeman L C 2016 *Social Networks* **44** 153
[24] Liu J X, Chen S D and Wang Y G 2012 *International Conference on Machine Learning and Cybernetics*, July 15–17, 2012, Xi'an, China, p. 920
[25] Tian H, Nie B and Zhang W J 2010 *International Conference on Machine Vision and Human-machine Interface*, April 24–25, 2010, Kaifeng, China, p. 737
[26] Lin D W 2018 *J. Algorithms Comput. Technol.* **12** 62
[27] Yang Y Z, Yu L, Wang X, Zhou Z L, Chen Y and Kou T 2019 *Physica A* **526** 121118
[28] Fan W L, Liu Z G and Hu P 2013 *Phys. Scr.* **88** 065201
[29] Shan X W, Zhang X and Tse C K 2022 *IEEE Trans. Circuits Syst. II: Express Briefs* **69** 4859
[30] Yang Y Z, Hu M and Huang T Y 2020 *Chin. Phys. B* **29** 088903
[31] Yan X L, Cui Y P and Ni S J 2020 *Chin. Phys. B* **29** 048902
[32] Dall J and Christensen M 2002 *Phys. Rev. E* **66** 016121
[33] Jiang L R, Jin X Y, Xia Y X, Ouyang B, Wu D P and Chen X 2014 *Int. J. Distrib. Sensor Networks* **10** 764698
[34] Lin H, Xia Y X and Jiang L R 2022 *Acta Phys. Sin.* **71** 068901 (in Chinese)
[35] Motter A E and Lai Y C 2002 *Phys. Rev. E* **66** 065102
[36] Liu J, Xiong Q Y, Shi W R, Shi X and Wang K 2016 *Physica A* **452** 209
[37] Burt R S 2004 *Am. J. Sociol.* **110** 349
[38] Yu H, Cao X, Liu Z and Li Y J 2017 *Physica A* **486** 318