

Article

Identifying Influential Edges by Node Influence Distribution and Dissimilarity Strategy

Yanjie Xu, Tao Ren *  and Shixiang Sun

Software College, Northeastern University, Shenyang 110169, China; 1710462@stu.neu.edu.cn (Y.X.); 1810476@stu.neu.edu.cn (S.S.)

* Correspondence: rent@swc.neu.edu.cn or chinarentao@163.com

Abstract: Identifying influential edges in a complex network is a fundamental topic with a variety of applications. Considering the topological structure of networks, we propose an edge ranking algorithm DID (Dissimilarity Influence Distribution), which is based on node influence distribution and dissimilarity strategy. The effectiveness of the proposed method is evaluated by the network robustness R and the dynamic size of the giant component and compared with well-known existing metrics such as Edge Betweenness index, Degree Product index, Diffusion Intensity and Topological Overlap index in nine real networks and twelve BA networks. Experimental results show the superiority of DID in identifying influential edges. In addition, it is verified through experimental results that the effectiveness of Degree Product and Diffusion Intensity algorithm combined with node dissimilarity strategy has been effectively improved.

Keywords: complex network; influence distribution; dissimilarity strategy; edge ranking



Citation: Xu, Y.; Ren, T.; Sun, S. Identifying Influential Edges by Node Influence Distribution and Dissimilarity Strategy. *Mathematics* **2021**, *9*, 2531. <https://doi.org/10.3390/math9202531>

Academic Editor: Nikita Frolov

Received: 17 September 2021

Accepted: 6 October 2021

Published: 9 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A complex network is a multidisciplinary topic in many domains including informatics, psychology, management, sociology, biology and engineering [1]. In reality, networks arise in a multitude of domains and are useful in solving numerous problems of human communities, such as the detection of bot accounts on Twitter [2], the discovery of vulnerabilities in electrical grids [3], identification of potentially harmful interactions between drugs [4], health care programs to predict the spread of epidemic diseases [5], improvement of routes in the development of road networks [6] and so on. The question of how to find influential nodes and edges is an important issue. Many methods are used to rank the nodes in networks. Degree [7], H-index [8] and k-shell [9] are based on nodes' neighbors. Closeness centrality [10] and betweenness centrality [11] are based on the path. Quasi-Laplacian centrality [12], Local Gravity Model [13] and AWLM [14] are based on semi-local structural information. In comparison, influential edges also play a significant role in complex network study. Influential edges analysis will be beneficial for guiding or controlling the network from a global perspective, moving the epidemic tipping point through topologically targeted social distancing [15] and so on.

In fact, for different problems, the importance of the edge in the network contains different meanings. In terms of the transmission of infectious diseases, the edge in the network represents the path of the transmission of infectious diseases, and its importance depends on the ability of this edge to spread diseases in the network. The stronger the transmission ability, the higher the edge importance. In electric transportation networks, edges in the network represent circuit connections. Its importance depends on the impact of the circuit failure on the network connectivity; the greater the impact on the network connectivity, the higher the edge importance. Therefore, the importance of the edge in the network under different problems is different. In this paper, we use network robustness R [16] as the target measure defining the importance of an edge for the impact on the network connectivity.

It is arduous that the number of edges in the network is more than the number of nodes. However, there are still great signs of progress madding in the research of identifying influential edges in complex networks. The earliest studies have reported in 1973. Granovetter [17] proposed that weakly connected edges may be more important than strongly connected edges, which has captured the attention of many researchers. Then, the research on edge strength has gradually emerged. Radicchi et al. [18] extended the clustering coefficient of nodes to edges and considered that the edges with lower clustering coefficient bridges communities generally. Gilbert and Karahalios [19] considered the attribute information and interaction strength of two nodes based on user characteristics and interaction behavior. En-Yu Yu [20] considered not only the degrees of nodes and cliques (local characteristics) but also the betweenness centrality (global characteristics) in order to rank important edges. Kossinets [21], Goyal [22] and Saito [23] proposed algorithms by learning a node behavior sequence and by calculating the influence probability.

In addition, Girvan and Newman [24] proposed edge betweenness that is based on the centrality of betweenness. It can accurately identify the important edges in the network but consumes huge computing resources. Holme et al. [25] multiplied the degrees of the two nodes as the centrality value of the edge. Liu et al. [26] measured the influence of an edge by counting the number of node first-order neighbors at the ends of the edge that are not connected to another node. These two methods are very fast, but the accuracy is very low. Onnela et al. [27] proposed a topological overlap method to judge the proportion of common neighbors in the total number of neighbors to measure the importance of the edge. It can improve the accuracy, but it was still poor at identifying edges that have a significant impact on network connectivity. The vital edge cannot be accurately identified; thus, the target removal edge has little effect on the network connectivity. An edge ranking algorithm that can accurately identify important edges is urgent needed.

Maintaining global network connectivity is the basic function of edges. In fact, the importance of an edge is related to the influence of the nodes at its two ends. However, the most important nodes tend to have many edges, which also causes these edges to be replaceable. In the paper, the scale of the gaint component is its nodes number. If edges that are highly replaceable are chosen to be removed, it will have a little impact on the scale of the giant component. On the contrary, the scale of the giant component will decrease sharply by removing the irreplaceable edge. For example, in the power grid, if the most irreplaceable edge is destroyed, it will cause a large-scale blackout. If such an edge is protected in advance, the impact will not be significant for other replaceable edges are damaged. Therefore, the rule of node influence distribution and the irreplaceability of edges should be considered comprehensively.

In order to improve the accuracy for identifying influential edges in complex network, we propose an edge ranking algorithm considering both local information and global information. Firstly, a node influence distribution model is employed for measuring the effect of the node on the edge. Subsequently, edge irreplaceability is revealed via the node dissimilarity strategy. Afterward, the edge ranking algorithm is proposed by combining the node influence distribution model (ID) and the node dissimilarity strategy (DIS). The purpose of the proposed DID algorithm is to accurately detect edges that can exert a strong influences over complex networks. Empirical results show that DID performs best in comparison with the four methods on nine real networks and twelve BA networks. In addition, DIS also can improve other methods that consider local information of the network.

The structure of the paper is as follows: In Section 2, ID and DIS are proposed first. Next, our method DID is proposed, and an analysis of DID is represented. The network data description and numerical results based on various classic methods applying to real networks and BA networks are shown in Section 3, respectively. Moreover, the experimental results are discussed in Section 4. Finally, conclusions are made in Section 5.

2. Algorithm

For different problems, the edge importance in the network is different. For the spreading problems, people take the number of infected nodes per unit time as the evaluation index of the network edge importance under the same transmission probability. On the contrary, for network connectivity problems, researchers measure edge importance by calculating the change of the scale of the most connected component in the network. The purpose of this paper is to find edges that have a significant impact on network connectivity.

In this paper, we introduce the node influence distribution model and node dissimilarity strategy, which are the basis of our proposed edge ranking algorithm. The proposed algorithm can identify influential edges widely. It works on unweighted and undirected networks. The proposed algorithm consists of the following three steps: distributes node influence, measures edge irreplaceability and proposes an edge ranking algorithm. Table 1 summarizes the symbols and notations used in the paper.

Table 1. Used symbols and variables.

Notation	Description
$\Gamma(i)$	Neighbor set of node i
CC_i	Closeness centrality of node i
N	Number of nodes in the network
M	Number of edges in the network
d_{ij}	Distance from node i to node j
$\Gamma(i) \cap \Gamma(j)$	The common neighbor of node i and j
k_i	Degree of node i
σ_{st}	Number of the shortest path between node s and t
$\sigma_{st}(e_{ij})$	Number of the shortest path between node s and t that goes through edge e_{ij}
$n_{i \rightarrow j}$	Neighbor of node j that is not connected with node i and not node i
$\langle k \rangle$	Average degree of the network
$\langle d \rangle$	Average distance of the network
D	Network diameter
C	Clustering coefficient of the network
r	Assortative coefficient of the network
G_i	Number of nodes in the largest connected component after removing i edges removed

2.1. Node Influence Distribution Model

It is known that the edge connected with the greater influence node also has greater influence on the network. However, it is difficult to distinguish the importance of these edges connecting the same influential node for many important nodes in reality with many edges. For example, nodes F and G are the most influential nodes in Figure 1. However, removing the connecting edge between nodes F and G has less impact on the overall connectivity of the network than by removing the connecting edge between nodes H and I. The influence of a node cannot directly reflect the influence of the edges connected to the node. When a node has many edges, the influence will be diluted accordingly. To solve these problems, node influence should be allocated to edges according to the actual rules.

Considering the global characteristics, closeness centrality [10] is selected to evaluate the influence of nodes. The influence of node i allocated to an edge e_{ij} is calculated by node i influences and the proportion of node j influence relative to the influence of all node i neighbors. The ID value is the product of the influence values allocated to the edge by the nodes at both ends of the edge. Thus, the edge influence is obtained as follows.

$$ID_{ij} = (CC_i \times \frac{CC_j}{\sum_{t \in \Gamma(i)} CC_t}) \times (CC_j \times \frac{CC_i}{\sum_{z \in \Gamma(j)} CC_z}) = \frac{(CC_i \times CC_j)^2}{\sum_{t \in \Gamma(i)} CC_t \times \sum_{z \in \Gamma(j)} CC_z} \quad (1)$$

CC_i is defined as the following.

$$CC_i = \frac{N}{\sum_{i \neq j} d_{ij}} \quad (2)$$

The node influence distribution model considers not only the information of the edge itself but also the path, which can more comprehensively reflect the importance of the edge.

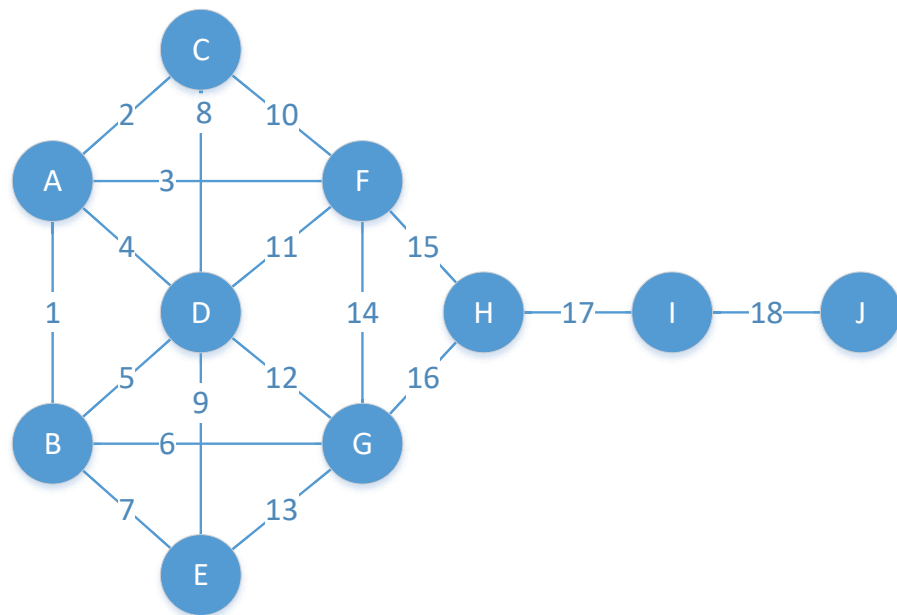


Figure 1. Kite network [28].

2.2. Strategy Based on Node Dissimilarity

One of the greatest challenges for edge ranking is that the number of edges is much larger than the nodes. Therefore, the cost of calculation is unacceptable for the dissimilarity between edges. The more common the neighbors between the two nodes with the same edge, the more backup paths there are corresponding to the edge. For example, in Figure 1, after the edge between node F and G is removed, there are two paths between them. The common neighbor plays the role of replacing the directly connected edge between the two nodes. Thus, the dissimilarity of nodes should be considered when identifying important edges to be more practical and efficient.

In order to measure the nodes similarity, Salton [29] is selected to explore the similarity influence in two nodes on the edge. The cost of calculating the edge by this method is $O(m)$. By calculating the similarity between nodes, the dissimilarity in nodes at both ends of the network is judged, and the irreplaceability of the edge is evaluated by the dissimilarity. The node dissimilarity is calculated as follows.

$$DIS_{ij} = 1 - \frac{1 + |\Gamma(i) \cap \Gamma(j)|}{\sqrt{k_i k_j}} \quad (3)$$

It should be noted that, in this paper, node dissimilarity is only used for pair of nodes with connected edges. Therefore, when the neighbors of two nodes are identical except for the degree of the two nodes being same and equal to the number of common neighbors plus one (for an edge between them), their similarity is one, and so the dissimilarity of the two nodes is zero.

2.3. Dissimilarity Influence Distribution Algorithm

Considering node influence distribution and edge irreplaceability, a new edge ranking algorithm is obtained by combining DIS with the ID model, named dissimilarity influence distribution algorithm (DID). DID is calculated as follows.

$$DID_{ij} = ID_{ij} \times DIS_{ij} = \left(\frac{(CC_i \times CC_j)^2}{\sum_{t \in \Gamma(i)} CC_t \times \sum_{z \in \Gamma(j)} CC_z} \right) \times \left(1 - \frac{1 + \Gamma(i) \cap \Gamma(j)}{\sqrt{k_i k_j}} \right) \quad (4)$$

The kite network (Figure 1) is used to describe the DID calculation process. The closeness centrality of each node in the kite network is calculated, which is shown in Table 2.

Table 2. The closeness centrality of each node in kite network.

Node	Closeness Centrality
A	0.5294118
B	0.5294118
C	0.5
D	0.6
E	0.5
F	0.6428571
G	0.6428571
H	0.6
I	0.4285714
J	0.03103448

The closeness centrality of A is 0.5294117647058824, which is the same as B. The neighbors of node A are B, C, D and F. The influence of edge 1 from node A can be obtained as follows.

$$CC_i \times \frac{CC_j}{\sum_{t \in \Gamma(A)} CC_t} = CC_A \times \frac{CC_B}{CC_B + CC_C + CC_D + CC_F} = 0.123$$

The influence of edge 1 from node B also is 0.123. The ID value of Edge 1 can be obtained by $0.123 \times 0.123 = 0.0152$.

Then, the dissimilarity of nodes A and B is calculated as the irreplaceability of edge 1. Node A is connected with nodes B, C, D and F, and node B is connected with nodes A, D, E and G. The common neighbor between node pair A B is node D, the number of the common neighbor is one and so the dissimilarity is 0.5. Therefore, the DID of edge 1 is calculated by $0.0152 \times 0.5 = 0.0076$. Correspondingly, the DID values of other edges can be obtained, as shown in Table 3.

Edge 17 has the highest score calculated by DID. Edge 17 is the only way to connect the left and right modules, which plays the role of the bridge. Although edges 4 and 5 connect the more important nodes in the network, their importance is decreased due to the presence of many neighbors relative to these nodes. The DID method can better reflect the real edge ranking sequence in the network.

Table 3. DID of each edge in kite network.

Edge	DID
1	0.007607
2	0.001166
3	0.003282
4	0.001076
5	0.000986
6	0.002363
7	0.00103
8	0.002224
9	0.00294
10	0.001585
11	0.002088
12	0.00273
13	0.002287
14	0.003826
15	0.007306
16	0.009036
17	0.012536
18	0.00664

3. Experiment

In this section, all experiments comprise targeted edge removal. Firstly, we explain four algorithms used in comparison with DID. Then, we describe the data sets used in our experiments. Next, the evaluation criterion network robustness R is described. The results are explained at the end.

3.1. Compared Algorithms

The performance of the proposed algorithm is compared with the following four algorithms:

(1) Edge Betweenness (EB [24]): EB considers the global information of the network and measures the edge importance by judging the proportion of an edge on the shortest path between any two nodes. It can be calculated as follows.

$$EB_{ij} = \sum_{s \neq t} \frac{\sigma_{st}(e_{ij})}{\sigma_{st}} \quad (5)$$

(2) Degree Product (DP [25]): The centrality of the edge can be obtained by multiplying the degree value of nodes at both ends of an edge, as follows.

$$DP_{ij} = k_i \times k_j \quad (6)$$

(3) Diffusion Intensity (DI [26]): The centrality of the edge can be obtained by counting the number of neighbors of one end node that is not connected to the other end node, as follows.

$$DI_{ij} = \frac{n_{i \rightarrow j} + n_{j \rightarrow i}}{2} \quad (7)$$

(4) Topological Overlap (TO [27]): The centrality of the edge can be obtained by calculating the ratio of common neighbors to unconnected neighbors, as follows.

$$TO_{ij} = \frac{\Gamma(i) \cap \Gamma(j)}{(k_i - 1) + (k_j - 1) - (\Gamma(i) \cap \Gamma(j))} \quad (8)$$

3.2. Data Set

In this paper, nine real networks from disparate fields including four social networks (Dolphin [30], polblogs [31], Sex [32] and Facebook [33]; three communication networks

(Email [34], As-733 [35] and PG [36]); and two collaboration networks (Jazz [37] and CA-CondMat [3]) are used to test the performance of DID and DIS combined with several classic methods. Dolphin is a social network of 62 dolphins. Polblogs is a social network in the political blogosphere of the United States. Sex is a bipartite network in which nodes are females (sex sellers) and males (sex buyers), and links between them are established when males write posts indicating sexual encounters with females. Facebook describes social circles from Facebook. Email describes email interchanges between users including faculty, researchers, technicians, managers, administrators and graduate students of the Rovira i Virgili University. As-733 contains the daily instances of autonomous systems from 8 November 1997 to 2 January 2000. PG is a snapshot of the Gnutella peer-to-peer file sharing network from August 2002. Jazz is a collaboration network of jazz musicians. Ca-CondMat is a collaboration network of Arxiv Condensed Matter category. Table 4 summarizes the key properties of the selected real sets.

The BA networks are used to test the performance of DID and nodes, and the average degrees include (500, 3), (500, 6), (500, 9), (500, 12), (5000, 3), (5000, 6), (5000, 9), (5000, 12), (50,000, 3), (50,000, 6), (50,000, 9) and (50,000, 12), respectively.

Table 4. The basic topological features of 9 real networks.

Networks	N	M	$\langle k \rangle$	$\langle d \rangle$	D	C	r
Dolphin	62	159	2.5645	3.3570	8	0.2590	−0.0436
Jazz	198	2742	13.8485	2.2350	6	0.6175	0.0202
Email	1133	5451	9.6222	3.6060	8	0.2540	0.0782
AS-733	6474	12,572	1.9419	3.705	9	0.2522	−0.1818
polblogs	1222	16,714	13.6776	2.7375	8	0.3203	−0.2213
PG	6299	20,776	3.2983	4.6430	9	0.0108	0.0355
Sex	15,810	38,540	2.4377	5.7846	17	0	−0.1145
CA-CondMat	23,133	93,239	4.0392	5.3522	15	0.6334	−0.1340
Facebook	26,954	497,878	18.4714	3.6925	8	0.2358	0.1421

3.3. Evaluation Criterion

The network robustness R is used as an evaluation criterion to compare the performance of DID with four algorithms on the considered data sets. The calculation process is as follows: delete the connected edges in the network one by one and calculate the size of the most connected subgraph of the normalized network until the network is empty. R could be obtained as follows:

$$R = \frac{1}{M} \sum_{i=1}^M \frac{G_i}{N} \quad (9)$$

where i is the number of edges removed.

3.4. Experiment of DID Performance

The proposed DID is compared with four classic algorithms in nine real networks. The results are shown in Table 5. The lower the network robustness R , the better performance of the algorithm.

Table 5 shows that the network robustness R obtained by the proposed DID generally outperforms the competitive methods as marked in bold. From Table 5, DID is the best performing algorithm on seven data sets (namely, Dolphin, Jazz, Email, polblogs, PG, Sex and Facebook) whereas TO is the best performing algorithm on the as-733 and CA-CondMat data set. Compared with TO, which considers the semi local information, DID comprehensively considers the global and local information of the network. By combining the relationship between nodes and edges, it can better reflect the importance of network edges.

Table 5. The R of DID and classic algorithms in real networks.

Networks	EB	DP	DI	TO	DID
Dolphin	0.511	0.6745	0.6165	0.4709	0.4407
Jazz	0.666	0.9136	0.8051	0.6398	0.6024
Email	0.6253	0.809	0.7622	0.5533	0.4721
As-733	0.5262	0.564	0.5675	0.4492	0.4934
polblogs	0.5956	0.9362	0.8537	0.4809	0.4379
PG	0.5764	0.7348	0.7377	0.6002	0.4504
Sex	0.5208	0.6416	0.6395	0.5925	0.3895
CA-CondMat	0.4395	0.6278	0.5313	0.4125	0.4143
Facebook	0.6979	0.9284	0.8824	0.728	0.5156

We further study the efficiency of the algorithm by observing the ratio of the remaining giant component relative to the original network after removing a certain proportion of edges. By observing the proportion dynamic change of the giant component relative to that of the original network, we can better represent the destruction of the algorithm. The results are as shown in Figure 2.

Figure 2 exhibits that DID has excellent performance in finding the key edges in the real network. With the important edges removed, the connectivity of the network is greatly damaged. Even in the as-733 network and Ca-Condmat network where the performance of DID is not optimal, the damage of DID is stronger than the other four methods when deleting the top 50% edges. Moreover, in the Facebook network, DID is the best in the entire process. In other networks, after more than half of the edges are deleted, DID is still more destructive than other methods. Dynamic network experiment proves the effectiveness of DID.

Next, we compared the proposed DID with four classic algorithms in 12 BA networks. The results are as shown in Table 6 where the best values are marked in bold. Figure 3 shows the dynamic damage of DID and four comparison algorithms relative to network connectivity. From Table 6 and Figure 3, DID has an excellent performance in finding the vital edges in the BA network.

Table 6. The R of DID and classic algorithms in the BA networks.

Networks	EB	DP	DI	TO	DID
BA-500-3	0.7325	0.7251	0.7262	0.7103	0.6399
BA-500-6	0.8651	0.844	0.8306	0.7771	0.6877
BA-500-9	0.886	0.8661	0.8551	0.8161	0.725
BA-500-12	0.9066	0.8728	0.8634	0.8384	0.7621
BA-5000-3	0.7346	0.7257	0.7227	0.7352	0.6493
BA-5000-6	0.8636	0.8445	0.8341	0.7982	0.7264
BA-5000-9	0.8962	0.8702	0.8575	0.8298	0.7597
BA-5000-12	0.9069	0.8788	0.8653	0.8693	0.7775
BA-50000-3	0.7358	0.723	0.7222	0.7379	0.6545
BA-50000-6	0.8578	0.8445	0.8333	0.8214	0.7283
BA-50000-9	0.8958	0.8715	0.8576	0.8484	0.775
BA-50000-12	0.9154	0.8804	0.8655	0.866	0.8131

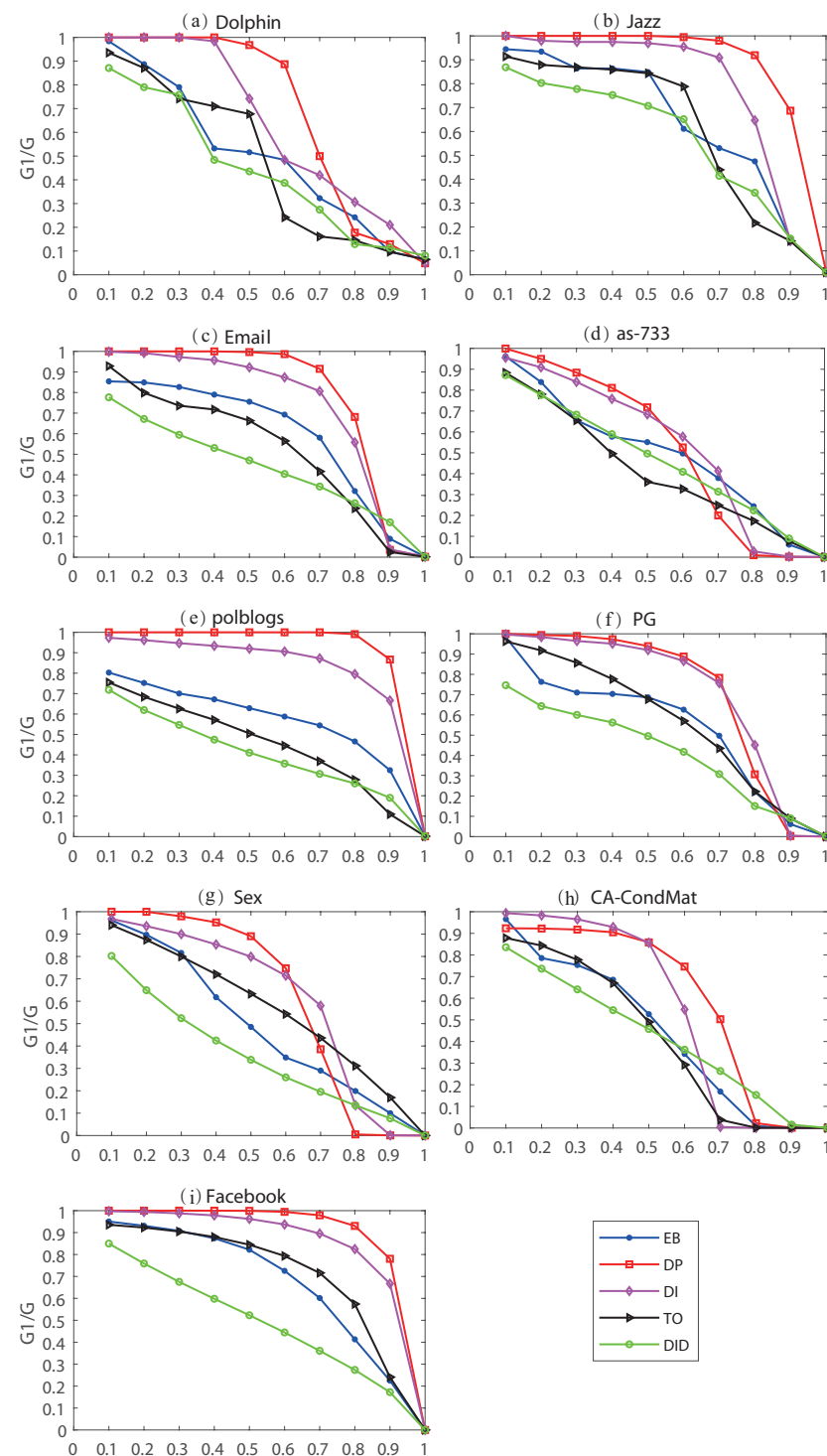


Figure 2. The dynamic damage of DID and 4 comparison algorithms to the real networks connectivity (a–i).

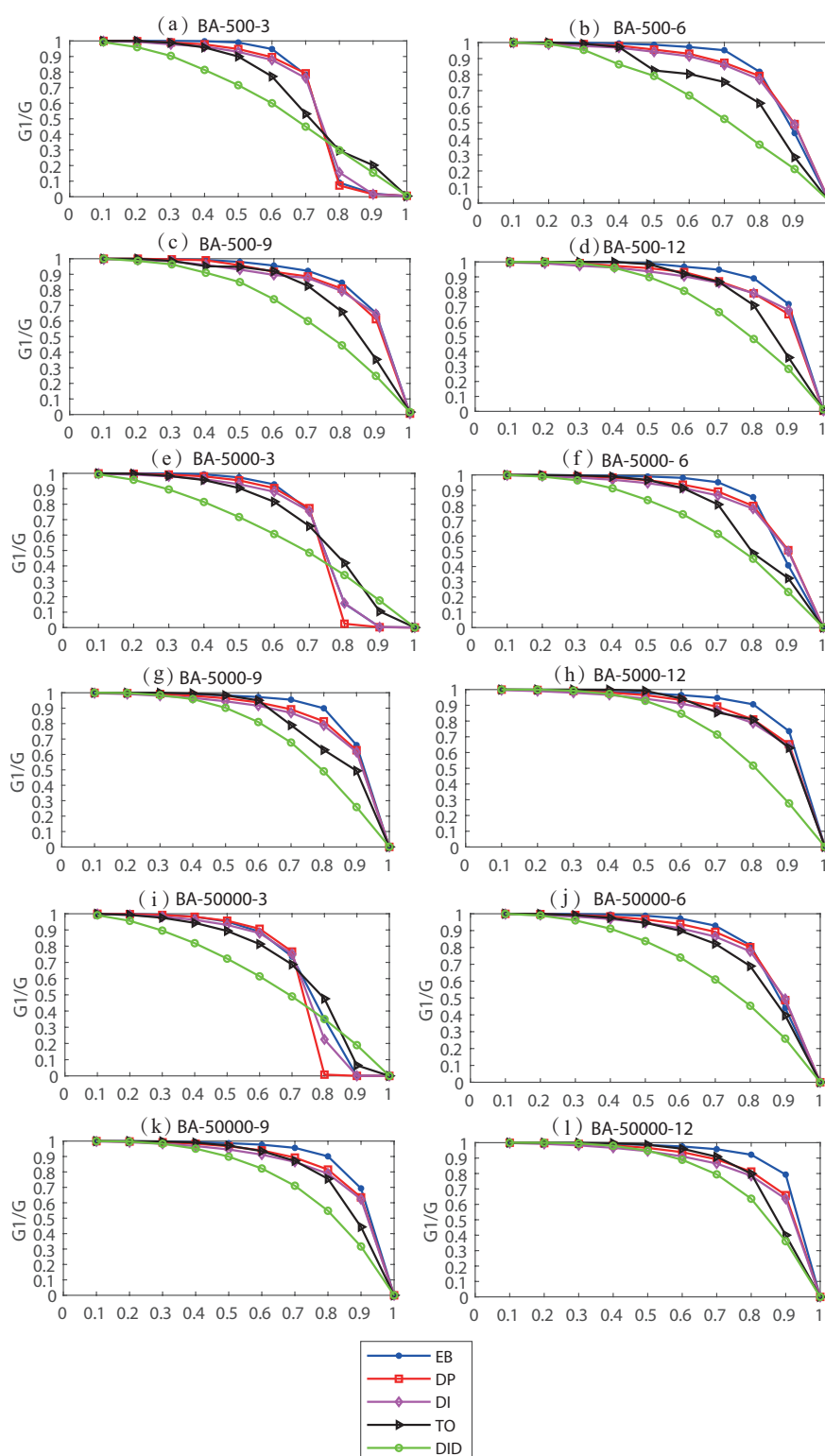


Figure 3. The dynamic damage of DID and 4 comparison algorithms to BA networks connectivity (a–l).

3.5. Experiment of DIS

In order to verify the proposed DIS, we compared the combination methods (named DEP, DDP, DDI and DTO) with the original methods in nine real networks. The network robustness R is used as the evaluation index. The experimental results are shown in Table 7

where the best values are marked in bold. The lower the R value, the better performance of the algorithm.

Table 7. The R of original methods and combination methods.

Networks	EB	DEB	DP	DDP	DI	DDI	TO	DTO
Dolphin	0.511	0.4912	0.6745	0.6494	0.6165	0.564	0.4709	0.4664
Jazz	0.666	0.6514	0.9136	0.8758	0.8051	0.7264	0.6398	0.07328
Email	0.6253	0.6129	0.809	0.7907	0.7622	0.7376	0.5533	0.555
As-733	0.5262	0.5357	0.564	0.5622	0.5675	0.5634	0.4492	0.449
polblogs	0.5956	0.5778	0.9362	0.9356	0.8537	0.8449	0.4809	0.4749
PG	0.5764	0.613	0.7348	0.7347	0.7377	0.7315	0.6002	0.6002
Sex	0.5208	0.5446	0.6416	0.6416	0.6395	0.6379	0.5925	0.5925
CA-CondMat	0.4395	0.4449	0.6278	0.5725	0.5313	0.5058	0.4125	0.5223
Facebook	0.6979	0.7097	0.9284	0.9242	0.8824	0.8943	0.728	0.721

As Table 7 shows, when compared with the original methods, it is obvious that DEB and DTO are not suitable, but DDP and DDI generally outperformed DP and DI. The performance of DID is the best when compared with these eight algorithms based on Tables 5 and 7.

4. Discussion

Identifying influential nodes and edges is a hot topic with a variety of applications in different fields, such as informatics, psychology, management, sociology, biology, engineering and so on. Degree product [24], as the simplest measure, considers that the importance of the edge is related to the importance of the nodes at its two ends. If nodes have a large number of neighbors, the edge will be crucial. However, it is likely that some bridge nodes are ignored, which connect different components but have a few neighbors. Edge Betweenness [20] considers an edge as important if most nodes' shortest path proceeds through it but has high computational complexity. To overcome this shortcoming, Diffusion Intensity [25] and Topological Overlap [26] are proposed, which can identify vital edges by the semi local information. These two measures seem to be more suitable for identifying the influential edges having a significant impact on network connectivity; however, they ignore the edge irreplaceability, which is more realistic. Thus, the DID algorithm is proposed in this paper, which firstly computes the influence that nodes distribute to the edge and then considers edge irreplaceability by computing node dissimilarity. The proposed DID algorithm is capable of identifying vital edges. Experimental results on nine real networks and twelve different BA networks show the feasibility and efficiency of DID.

Firstly, the experiments comparing network robustness R exhibit the superiority of DID, for the vital edges identified by the algorithm are observed to easily exert strong influence than compared to the competitors. Although inferior to TO on the as-733 and CA-CondMat data set, it is also quite competitive. As shown in Figure 2, by observing the ratio of the remaining giant component to the original network after removing a certain proportion of edges, the networks are more easily broken up by DID. The results on BA networks also verify the feasibility and efficiency of DID. Secondly, we apply network robustness R to evaluate the DIS effect on the classical algorithm (as shown in Table 7). In general, DIS can only improve the performance of the method in some real networks, some of which may even be reduced. Generally speaking, for some identifying influential edge methods that consider local information of nodes, this strategy can effectively enhance the performance of the algorithm at the cost of $O(m)$ time when ranking edges for target removal edges. It is worth noting that the DID is still the best.

5. Conclusions

In this paper, by considering node influence distribution and edge irreplaceability, we proposed an edge ranking method named DID and compared it with four classic methods

in nine real networks and twelve BA networks by network robustness R and the proportion of dynamic change relative to the giant component. The results show that DID performs well in identifying influential edges that have a significant impact on network connectivity. This will help us in some real-life applications such as controlling the spreading of rumors and targeted attacks on networks and so on.

In addition, we combined DIS with four classic methods (such as DEB, DDP, DDI and DTO). The results show that DIS can effectively improve the performance of DP and DI algorithms, which are based on local information of the network. The reason is that the node dissimilarity strategy is more realistic, and these algorithms combined with this strategy are more comprehensive for considering the topology of structure. However, for TO and EB, which consider the network's semi-local or global information, their accuracies are reduced when identifying the edge that is more important compared to network connectivity. Therefore, for different methods, the question of whether to choose this strategy based on node dissimilarity should be answered and improved upon.

Author Contributions: Y.X. designed the algorithm and wrote the original draft; T.R. revised the manuscript; T.R. and S.S. checked the manuscript and made some modifications. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fundamental Research Funds for the Central Universities under Grant No. N181706001, N2017009, N2017008, N182608003 and N181703005, National Natural Science Foundation of China under Grant No.61902057 and Joint Fund of Science and Technology Department of Liaoning Province and State Key Laboratory of Robotics China under Grant No.2020-KF-12-11.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data sets are available at <http://snap.stanford.edu/data> and <https://www.neusncp.com/api/datasets>.

Acknowledgments: We would like to thank the anonymous reviewers for their careful reading and useful comments that helped us to improve the final version of this paper.

Conflicts of Interest: The authors declare that there are no conflict of interest regarding the publication of this paper.

References

1. Newman, M.E.J. *Networks*; Oxford University Press: Oxford, UK, 2018.
2. Davis, C.A.; Varol, O.; Ferrara, E.; Flammini, A.; Menczer, F. BotOrNot: A System to Evaluate Social Bots. In Proceedings of the 25th International World Wide Web Conference Companion, Republic and Canton of Geneva, CHE, Montreal, QC, Canada, 11–15 April 2016; pp. 273–274.
3. Pagani, G.A.; Aiello, M. The Power Grid as a complex network: A survey. *Phys. A Stat. Mech. Its Appl.* **2013**, *11*, 2688–2700. [CrossRef]
4. Guimerà, R.; Sales-Pardo, M. A network inference method for large-scale unsupervised identification of novel drug-drug interactions. *PLoS Comput. Biol.* **2013**, *9*, e1003374. [CrossRef] [PubMed]
5. Kamath, P. S.; Wiesner, R.H.; Malinchoc, M.; Kremers, W.; Therneau, T.M.; Kosberg, C.L.; D'Amico, G.; Dickson, E.R.; Kim, W.R. A model to predict survival in patients with end-stage liver disease. *Hepatology* **2001**, *33*, 464–470. [CrossRef] [PubMed]
6. Linyuan, L.; Tao, Z. A model to predict survival in patients with end-stage liver disease. *Phys. A Stat. Mech. Its Appl.* **2011**, *390*, 1150–1170.
7. Bonacich, P. Factoring and weighting approaches to status scores and clique identification. *J. Math. Sociol.* **1972**, *2*, 113–120. [CrossRef]
8. Lu, L.; Tao, Z.; Zhang, Q.; Stanley, H. The H-index of a network node and its relation to degree and coreness. *Nat. Commun.* **2016**, *7*, 10168. [CrossRef]
9. Kitsak, M.; Gallos, L.K.; Havlin, S.; Liljeros, F.; Muchnik, L.; Stanley, H.E.; Makse, H.A. Identification of influential spreaders in complex networks. *Nat. Phys.* **2010**, *6*, 888–893. [CrossRef]
10. Freeman, L.C. Centrality in social networks conceptual clarification. *Soc. Netw.* **1979**, *1*, 215–239. [CrossRef]
11. Freeman, L.C. A set of measures of centrality based on betweenness. *Sociometry* **1977**, *40*, 35–41. [CrossRef]
12. Ma, Y.; Cao, Z.L.; Qi, X.Q. Quasi-Laplacian centrality: A new vertex centrality measurement based on Quasi-Laplacian energy of networks. *Phys. A Stat. Mech. Its Appl.* **2019**, *527*, 121130. [CrossRef]

13. Li, Z.; Ren, T.; Ma, X.; Liu, S.; Zhang, Y.; Zhou, T. Identifying influential spreaders by gravity model. *Sci. Rep.* **2019**, *9*, 8387. [\[CrossRef\]](#)
14. Li, Z.; Ren, T.; Xu, Y.; Chang, B.; Chen, D.; Sun, S. Identifying Influential Spreaders Based on Adaptive Weighted Link Model. *IEEE Access* **2020**, *8*, 66068–66073. [\[CrossRef\]](#)
15. Ansari, S.; Anvari, M.; Pfeffer, O.; Molkenthin, N.; Moosavi, M.R.; Hellmann, F.; Kurths, J. Moving the epidemic tipping point through topologically targeted social distancing. *Eur. Phys. J. Spec. Top.* **2021**. [\[CrossRef\]](#)
16. Schneider, C.M.; Moreira, A.A.; Andrade, J.S., Jr.; Havlin, S.; Herrmann, H.J. Mitigation of malicious attacks on networks. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3838–3841. [\[CrossRef\]](#)
17. Granovetter, M.S. The Strength of Weak Ties. *Am. J. Sociol.* **1973**, *78*, 1360–1380. [\[CrossRef\]](#)
18. Radicchi, F.; Castellano, C.; Cecconi, F.; Loreto, V.; Parisi, D. Defining and identifying communities in networks. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 2658–2663. [\[CrossRef\]](#)
19. Gilbert, E.; Karahalios, K. Predicting Tie Strength with Social Media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 09), Boston, MA, USA, 4–9 April 2009; pp. 211–220.
20. Yu, E.Y.; Chen, D.B.; Zhao, J.Y. Identifying critical edges in complex networks. *Sci. Rep.* **2018**, *8*, 14469. [\[CrossRef\]](#)
21. Kossinets, G.; Kleinberg, J.; Watts, D. The Structure of Information Pathways in a Social Communication Network. In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 08), Las Vegas, NV, USA, 24–27 August 2008; pp. 435–443.
22. Goyal, A.; Bonchi, F.; Lakshmanan, L.V. Learning Influence Probabilities in Social Networks. In Proceedings of the Third ACM International Conference on Web Search and Data Mining (WSDM 10), New York, New York, USA, 4–6 February 2010; pp. 241–250.
23. Saito, K.; Kimura, M.; Ohara, K.; Motoda, H. Behavioral Analyses of Information Diffusion Models by Observed Data of Social Network. In *Advances in Social Computing' SBP 2010*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2010; pp. 149–158.
24. Girvan, M.; Newman, M.E.J. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821–7826. [\[CrossRef\]](#)
25. Holme, P.; Kim, B.J.; Yoon, C.N.; Han, S.K. Attack vulnerability of complex networks. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* **2002**, *65*, 056109. [\[CrossRef\]](#)
26. Liu, Y.; Tang, M.; Zhou, T.; Do, Y. Improving the accuracy of the k-shell method by removing redundant links: From a perspective of spreading dynamics. *Sci. Rep.* **2015**, *5*, 13172. [\[CrossRef\]](#)
27. Onnela, J.P.; Saramäki, J.; Hyvönen, J.; Szabó, G.; Lazer, D.; Kaski, K.; Kertész, J.; Barabási, A.L. Structure and tie strengths in mobile communication networks. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7332–7336. [\[CrossRef\]](#)
28. Krackhardt, D. Assessing the Political Landscape: Structure, Cognition, and Power in Organizations. *Adm. Sci. Q.* **1990**, *35*, 342–369. [\[CrossRef\]](#)
29. Salton, G.; McGill, M.J. *Introduction to Modern Information Retrieval*; McGraw-Hill, Inc.: New York, NY, USA, 1986.
30. Lusseau, D.; Schneider, K.; Boisseau, O.J.; Haase, P.; Slooten, E.; Dawson, S.M. The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. *Behav. Ecol. Sociobiol.* **2003**, *54*, 396–405. [\[CrossRef\]](#)
31. Adibi, J.; Chalupsky, H.; Grobelenik, M.; Mladenovic, D.; Milic-Frayling, N. KDD-2004 workshop report link analysis and group detection (LinkKDD-2004). *SIGKDD Explor. Newsl.* **2004**, *6*, 136–139. [\[CrossRef\]](#)
32. Rocha, L.E.; Liljeros, F.; Holme, P. Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS Comput. Biol.* **2011**, *7*, e1001109. [\[CrossRef\]](#)
33. McAuley, J.J.; Leskovec, J. Learning to discover social circles in ego networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS'12), Lake Tahoe, NV, USA, 3–6 December 2012; Curran Associates Inc.: Red Hook, NY, USA, 2012; Volume 1, pp. 539–547.
34. Guimerà, R.; Danon, L.; Díaz-Guilera, A.; Giralt, F.; Arenas, A. Self-similar community structure in a network of human interactions. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* **2003**, *68*, 065103. [\[CrossRef\]](#)
35. Leskovec, J.; Kleinberg, J.; Faloutsos, C. Graphs over time: Densification laws, shrinking diameters and possible explanations. In Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining (KDD 05), New York, NY, USA, 21–24 August 2005; pp. 177–187.
36. Leskovec, J.; Kleinberg, J.; Faloutsos, C. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data* **2007**, *1*, 1–40. [\[CrossRef\]](#)
37. Gleiser, P.M.; Danon, L. Community Structure in Jazz. *Adv. Complex Syst.* **2003**, *6*, 565–573. [\[CrossRef\]](#)