Reliability Engineering and System Safety 174 (2018) 41–52

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

# Measuring the vulnerability of community structure in complex networks

Daijun Wei [a,b,1], Xiaoge Zhang [b,1], Sankaran Mahadevan [b,*]

[a] School of Science, Hubei University for Nationalities, Enshi 445000, China

[b] Department of Civil and Environmental Engineering, School of Engineering, Vanderbilt University, Nashville, TN 37235, USA

ARTICLE INFO

*Keywords:*
Complex networks
Community
Vulnerability
Connectivity

ABSTRACT

This paper develops a quantitative method to measure the vulnerability of community structure with emphasis on both internal and external connectivity characteristics of the community. In particular, the number of links between communities and the strength of links connecting two communities are considered as external factors, while the connection density, the degree of gateway nodes, as well as the strength of links within each community are treated as internal factors. A non-linear weighted function is used to combine the internal factors with external factors. Then the developed method is used to illustrate the vulnerability analysis of community structure of a power transmission grid, a karate club network, and an air transportation network. The results reveal that the proposed measure is effective in differentiating the vulnerability level of community structure in a variety of networks.



## 1. Introduction

Community structure is a widely acknowledged characteristic in a large body of research on complex networks, as highlighted in recent studies [1–3]. Typically, a community is a subnetwork consisting of a group of nodes with a higher connection density of edges within the group than between the groups [2]. In general, there are two fundamental issues related to the community study in complex networks. The first one is how to discover community structure in networks, and the other one is what properties these communities have in common. Numerous approaches have been developed to detect the community structure in a variety of networks. For example, since it is common for two communities to share some nodes in complex networks, Orman et al. developed a novel method to detect overlapping communities [4]. Rocco et al. analyzed the effects of multi-state links on community detection [5]. To account for the dynamic behavior of each node in evolving networks, Orman et al. developed an innovative community detection method, in which the evolution of topology, nodal attributes, and community structure over time were considered [6]. In [7], they also characterized the role of each node by studying the evolution of its neighborhood based on the assumption that the neighborhood changes reveal the importance of the node in the entire network. The aforementioned approaches have been studied for several real-world networks, including scientific collaboration networks, and a network of Jazz listeners extracted from LastFM [6,7]. Meanwhile, several algorithms have been developed for discovering community structure in networks [8,9]. For example, a hierarchical agglomerative algorithm (HAA) for detecting community structure in unweighted network was proposed by Newman et al. [2,10] and the algorithm was further extended for weighted networks [11,12], which will be introduced in Sections 2.3 and 2.4, respectively.

With respect to the second issue, the vulnerability evaluation on community structure in complex networks has received increasing attention. For example, Zio et al. identified three elements related to the vulnerability of infrastructure systems, namely the degree of loss caused by a hazard, the degree of exposure to the hazards, and the degree of ability to recover to a stable state, and they demonstrated the analysis procedures for Critical Infrastructures (CIs) [13]. Holme et al. defined vulnerability as the reduction of network functionality due to selected removal of certain vertices or links, and investigated which attack strategy (based on measures such as degrees of nodes and betweenness centrality) is most effective [14]. Torres-Vera et al. measured the vulnerability of a pipeline system as the amount of the damage suffered by a structure due to a seismic activity [15].

Recently, a qualitative metric to measure the vulnerability of community structure was developed in Ref. [16], in which the vulnerability is quantified in terms of the connectivity between any two communities. The vulnerability between communities $x$ and $y$, denoted as $v_{xy}$, is a metric to provide a measure of the degree to which two communities are disconnected. This is a qualitative metric; however, it offers some insights regarding the relative strength of each community. The developed community structure vulnerability measure has been illus-

* Corresponding author.
*E-mail address:* sankaran.mahadevan@vanderbilt.edu (S. Mahadevan).

[1] D. Wei and X. Zhang contributed equally to this work.

https://doi.org/10.1016/j.ress.2018.02.001
Received 10 May 2017; Received in revised form 29 January 2018; Accepted 1 February 2018

trated in several networks, such as telephone network, power network, and terrorist network [16]. Considering that vulnerability denotes the ability of an individual or group to withstand and cope with the impact of natural or man-made hazards [17], connectivity is certainly an important consideration in measuring a community's vulnerability. Therefore, the proposed method in this paper also focuses on connectivity. A community that has more connections to other communities is more capable of coping with the impact of hazardous events, due to the opportunity to draw assistance and resources from neighboring communities. Similarly, within a community, the more the number of connections, the stronger is its ability to cope with the hazard, due to the increased ability to get in contact with other community members for receiving support. One study worthy of mention is Carrington et al., which considered the inter and intra links among communities to measure group centrality scores [18]; they stated that the number of connections alone cannot fully characterize the vulnerability of community structure. The quality or strength of the connections, as well as the structure of the community, are also important considerations. Actually, the vulnerability metric developed in Ref. [16] is a special case of the group centrality score developed in Ref. [18].

To illustrate the above points, consider a ground invasion scenario where the objective of the invading force is to defeat a community. The community is less vulnerable to the attack if it is well connected and able to draw support from other neighboring communities. The more the number of connections, the less vulnerable the community – this is the implication of the metric in Ref. [16]. Therefore, the attacking force might try to isolate the targeted community from its neighbors by severing or blocking its transportation and communication links to the neighbors. However, different links might have different strengths and thus might require different amounts of resources in order to be destroyed or blocked. Thus, in addition to the number of connections, the strength of the connections is also a significant consideration in affecting a community's vulnerability.

When a community is being attacked, its structure and topology will also determine its ability to withstand the attack. If the community has higher connection density and connection strengths, when a specific zone is attacked, it can receive more help from other zones, which in turn enhances its ability to withstand the attack. Another factor is the connectivity of the gateway nodes (i.e., nodes that connect the community to other communities) to other nodes within the community. Such nodes play an important role in distributing food, water, fuel, energy, and information received from other communities. The adequate operation of resource distribution within the community is vital in strengthening its ability to cope with the attack.

Considering the above factors, the number of external connections to neighboring communities alone cannot differentiate the vulnerability of different communities. For example, the telephone network in Belgium is divided into 7 communities in Ref. [16]. Five of the seven communities are different from each other in topology and structure, but all of them have the same number of connections to their neighboring communities. Based on the number of external connections alone, all five communities would be judged to have the same level of vulnerability. Such a measure has a low resolution in differentiating community vulnerability, which limits its usefulness in practical applications.

The goal of this paper is to develop a generalized measure that has the ability to distinguish the degree of vulnerability among separate communities. The proposed measure fuses the information regarding connectivity characteristics of the community, both internal and external. Our major idea is described as follows: vulnerability of community structure is not only related to the outer connectivity of a community, but also its inner structure. Two communities may have different vulnerability even if they have the same number of edges connecting with other communities. Besides the number of external connections, the degree of the gateway nodes, the connection density within the community, the strengths of edges connecting with other communities, and the strengths of edges within the community are also important considerations when measuring the vulnerability of community structure. Thus, in this paper, we consider 5 factors - 2 external factors (number of links between communities, and strengths of the links connecting two communities), and 3 internal factors (connection density, degrees of gateway nodes, and strengths of the links within each community). All five factors are combined into one composite qualitative metric in this paper. Five parameters are used to indicate the weights related to the factors we considered. As a result, the qualitative vulnerability measure in Ref. [16] becomes a special case of the vulnerability measure proposed in this paper. The proposed method is illustrated in three networks, namely a power transmission grid [19], a karate club network [20], and an air transport network [21]. The results show the effectiveness of the proposed method in differentiating the vulnerability values of separate individual communities.

The rest of the paper is organized as follows: the basic concepts are introduced in Section 2. In Section 3, we propose a generalized metric to measure the vulnerability of community structure. In Section 4, we illustrate the proposed vulnerability metric with three different networks. In Section 5, we provide concluding remarks.

## 2. Preliminaries

In this section, basic concepts such as the degree of a node, connection density, and strength of edges in complex networks, are introduced. In addition, HAA [2,10], a classical method for detecting community structure in unweighted networks and its extension to detect community in weighted networks, are described.

### 2.1. Degree and connection density in unweighted networks

Consider a complex network $G(E, N)$, where $E = (1, 2, \cdots, m)$ is the set of edges and $N = (1, 2, \cdots, n)$ is the set of nodes, where $m$ and $n$ are the numbers of edges and nodes in network $G$, respectively.

**Definition 2.1.** The degree of a node $i$ in an unweighted network, denoted as $k_i$, is defined as [22]:

$$k_i = \sum_{j=1}^{N} x_{ij} \tag{1}$$

where $x_{ij}$ represents the connection between node $i$ and node $j$. $x_{ij} = 1$ if node $i$ is connected to node $j$, and $x_{ij} = 0$ otherwise.

It can be observed that the degree of a node is the number of links connecting itself with other nodes. The average degree of all the nodes is $\overline{d} = \frac{1}{n}\sum_{i=1}^{n} k_i$, where $n$ is the number of nodes in network $G$.

The network density describes the fraction of the potential connections in a network that are actual connections [23]. By contrast, a potential connection is a connection that could potentially exist between two nodes (regardless of whether or not it) [24].

**Definition 2.2.** The network density, denoted as $\rho$, is defined as,

$$\rho = \frac{m}{C_n^2} \tag{2}$$

where $m$ and $n$ are the number of edges and nodes in the network, respectively. $C_n^2$ is the number of links if any two nodes are connected in the network.

From this definition, we have $0 \le \rho \le 1$. $\rho = 0$ means there is no connection in the network, while $\rho = 1$ reveals that the network is fully connected. Network density is a significant indicator on the connectivity of a network.

### 2.2. Strengths of edges in weighted networks

Generally, a weighted network can be modeled as: $G(E, N, W)$, where $W \in R^+$ is the weight of an edge. If $W \equiv 1$ for all the links, then the weighted network degenerates to an unweighted network. As mentioned before, the vulnerability of a community is defined as its capability to withstand and cope with man-made or natural hazards. For the sake of simplicity, we assume that the strength value of a link in a community is proportional to its weight. This assumption holds for many applications. For example, in ground traffic networks, link weights are often used to represent the amount of traffic going through them. Typically, a link that carries more traffic has more lanes, and the quality of the link is usually designed to be higher than other roads due to its significant role in transporting passengers and goods. In this case, the attackers need to spend more resources to make this link lose its functionality. From this point of view, the link weight reflects the capability of the link to withstand the attack.

### *2.3. HAA algorithm*

A number of measures have been proposed for discovering community structure in networks, for instance, the hub-dominance measure [25], the within-community degree [26], and Newman's modularity measure [27]. In this paper, Newman's modularity is used as a quantitative measure to determine the community structure. Accompanying the metrics for community detection, several algorithms have been proposed for detecting community structure in complex networks, such as the Kernighan–Lin algorithm [28], spectral partitioning [29] and others [30,31]. Based on greedy optimization, Newman et al. developed the concept of modularity and a hierarchical agglomerative algorithm (HAA) to detect community [2]. The HAA is faster than many competing algorithms [10].

**Definition 2.3.** For a network with $l$ communities, the modularity, denoted as $Q$, is defined as [2,10],

$$Q = \sum_{k=1}^{l} \left( \frac{g_k}{m} - \left( \frac{d_k}{2m} \right)^2 \right) \tag{3}$$

where $l$ is the number of communities, $g_k$ is the number of edges in community $k$, $d_k$ represents the total degree of all the nodes in community $k$, and $m$ represents the total number of edges in the network.

The value of $Q$ is a good indicator of the presence of community structure. $Q = 0$ indicates that all the nodes are in one single community (the community structure does not exist in the network), while $Q = 1$ indicates the presence of a strong community structure and $Q > 0$ indicates there are some kind of community structure in network. Newman suggest "in practice, values for such networks typically fall in the range from about 0.3 to0.7" for the value of $Q$ in Ref. [27]. In fact, an error range of $Q$ is permitted due to variability in the estimated values. For example, for the social network of a community of 62 bottlenose dolphins living in Doubtful Sound, New Zealand discussed in Ref. [27], the modularity value Q was estimated to be $0.38 \pm 0.08$.

To detect the community structure, the main idea in HAA is to find the changes in $Q$ in three steps. First, every node in a complex network is considered as a single community. Second, a new community is obtained by combining communities $i$ and $j$, and the values of modularity change ($\Delta Q_{ij}$) are calculated. Third, according to the highest $\Delta Q_{ij}$, the two communities $i$ and $j$ are combined into a single community. The second and third steps are repeated until $\Delta Q_{ij} \leq 0$.

### *2.4. Extended HAA for weighted networks*

Several researchers studied whether an appropriate algorithm exists for weighted networks after the community detection algorithm was proposed. Newman et al. proposed a community detection algorithm for a weighted network based on HAA [12]. The core idea of the extended algorithm is still based on modularization. Similar to Eq. (3), the modularity, denoted as $Q_w$ for a weighted network, is given as [11],

$$Q_w = \sum_{k=1}^{l} \left( \frac{g_k'}{m'} - \left( \frac{d_k'}{2m'} \right)^2 \right) \tag{4}$$

where $l$ is the number of communities, $m'$ is the total weights of all the links in the entire network, $g_k'$ is the sum of weights of links within community $k$, and $d_k'$ represents the total weights of all the nodes in community $k$. In reality, when the value of $Q_w$ is more than 0.3, it usually indicates a good division [11].

To determine the community structure in weighted networks, it is important to note that not all weights on network edges can be appropriately used as inputs. Weights do not always indicate close connections or similarity between vertices. For example, Barrat et al. have revealed that high-volume routes are not necessarily close or similar between airports in an air transport network [32]. According to Newman, "traffic between Los Angeles and Tokyo is very high, but this does not mean that Los Angeles and Tokyo are similar places, or that they are close to one another" [11].

## 3. Proposed method

In this section, the qualitative metric of community vulnerability defined in Ref. [16] is first introduced. The limitation of the qualitative metric is illustrated using a numerical example. Then a generalized measure of community vulnerability is developed.

### *3.1. Classical community vulnerability qualitative measure*

We review two classical vulnerability qualitative metrics used in Ref. [16].

**Definition 3.1.** The vulnerability and the relative vulnerability of community $x$, denoted as $v_x$ and $R_x$ respectively, are defined as

$$v_x = \frac{1}{|V_x|} \tag{5}$$

and

$$R_x = \frac{v_x}{v}; \quad v = \min_y(v_y) \tag{6}$$

where $V_x$ is a set of communities, which are connected with the community $x$. $|V_x|$ is the number of links between community $x$ and other communities.

In the above definition, the community vulnerability measure $v_x$ is inversely proportional to the number of edges connecting with other communities. Even though $v_x$ and $R_x$ have quantitative values, their practical use is in qualitatively comparing the relative vulnerability of community structure between two communities. The actual values may not be directly useful, but we could infer that the vulnerability of a particular community structure is either high or low. To illustrate $v_x$ and $R_x$, an example network with 9 nodes and 14 edges is considered in Ref. [16]. Using HAA, the communities in the network are detected, which are shown in Fig. 1. The modularity ($Q$) of the communities is 0.29 (using Eq. (3)), and since it is quite close to the lower bound of $Q$ in Ref. [27], it is assumed to indicate that the network has a community structure. This assumption is further reinforced by the fact that the network connection density within each community is higher than between the communities.

In Fig. 1, there are three communities: $A$, $B$ and $C$. Community A includes nodes 1 and 2, community B includes nodes 3, 4 and 5, and community C includes nodes 6, 7, 8 and 9. Communities $A$ and $B$ have two links each connecting to community $C$. Community $C$ has four edges connecting to other communities. Thus, we have $|V_C| = 2|V_A| = 2|V_B|$. According to Eq. (5), the vulnerability of each community is: $v_A = v_B = 0.5$ and $v_C = 0.25$. Using Eq. (6), the relative vulnerability of each community is : $R_A = R_B = 2$ and $R_C = 1$. The results show that the community vulnerability of $A$ and $B$ is twice as community $C$.

In the above definition, the community vulnerability measure only depends on the number of edges between the communities. Thus, $v_A = v_B = 2v_C$ and $R_A = R_B = 2R_C$ because $|V_C| = 2|V_A| = 2|V_B|$ in Fig. 1.

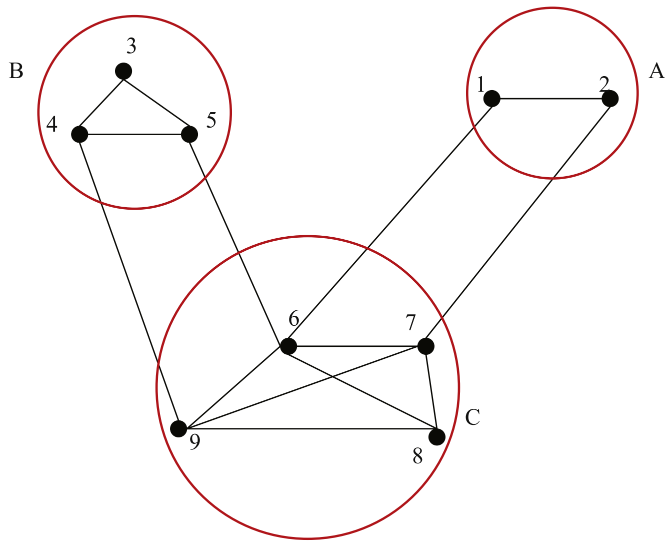


**Fig. 1.** A partition of communities for an unweighted network with 9 nodes [16].

However, as can be noted from Fig. 1, there are only two nodes in community $A$. Whereas, the topology of community $B$ is a triangle, and there are three nodes 3, 4 and 5, and three links among the nodes. Clearly, communities $A$ and $B$ are different with respect to their topologies. As mentioned earlier, the topology of a community significantly impacts its ability to withstand the hazardous event. For example, in community $A$, node 1 can only receive help from node 2, and vice versa. While in community $B$, each node can get help from two other nodes. From this perspective, community $A$ and $B$ have different abilities to cope with the hazard. Thus, they should have different vulnerability values.

### 3.2. Proposed community vulnerability measure

In this section, we propose a new measure to characterize the community vulnerability. In our metric, we consider five factors, i.e., two external factors (number and strengths of edges connecting to other communities), and three internal factor (connection density, average strength of edges within community, and average degree of the gateway nodes within community). As a result, we develop a generalized vulnerability measure for community $x$.

**Definition 3.2.** The vulnerability of community $x$ is defined as,

$$v_x' = \frac{1}{|V_x|^{\alpha}(\overline{s}_x)^{\beta}} \cdot \frac{1}{(\rho_{x'})^{\gamma}(\overline{s}_{x'})^{\delta}(\overline{d}_{x''})^{\lambda}} \tag{7}$$

where $\alpha,\ \beta,\ \gamma,\ \delta,\ \lambda \geq 0$, $\overline{s}_x$ is the average strength of edges between the communities, $\rho_{x'}$ is the connection density of community $x$, $\overline{s}_{x'}$ is the average strength of links within community $x$, and $\overline{d}_{x''}$ represents the average degree of gateway nodes in community $x$.

In practice, to keep all the parameters in the same scale, we normalize all of their values of $|V_x|$, $\overline{s}_x$, $\rho_{x'}$, $\overline{s}_{x'}$ and $\overline{d}_{x''}$. The parameters $\alpha$, $\beta$, $\gamma$, $\delta$ and $\lambda$ represent the relative weights of the five factors, namely the number of externally connected edges, the average strength of externally connected edges, the connection density within the community, the average strength of links within the community and the average degree of the gateway nodes below, respectively. We show some special cases of the developed metric.

- When $\alpha = 1$, and $\beta = \gamma = \delta = \lambda = 0$, the proposed measure degenerates to the same form as the vulnerability measure defined in Ref. [16].
- When $\alpha = \beta = \gamma = \delta = \lambda$, it denotes that the five factors are equally weighted.

**Table 1**
The community vulnerability for Fig. 1 using the proposed method.

| Communities | $\|V_x\|$ | $\overline{s}_x$ | $\rho_{x'}$ | $\overline{s}_{x'}$ | $\overline{d}_{x''}$ | $v_x'$ | $R_x'$ | $R_x$ [16] |
|---|---|---|---|---|---|---|---|---|
| $A$ | 1/2 | 1 | 1 | 1 | 1/3 | 6.00 | 6.00 | 2.00 |
| $B$ | 1/2 | 1 | 1 | 1 | 2/3 | 3.00 | 3.00 | 2.00 |
| $C$ | 1 | 1 | 1 | 1 | 1 | 1.00 | 1.00 | 1.00 |

- When $\gamma = \delta = \lambda = 0$, it reveals that only the connections between the communities (external factors) are considered.
- When $\alpha = \beta = 0$, it means that we only account for the internal factors.
- When $\alpha = \beta = \gamma = \delta = \lambda = 0$, it means the community is an isolated node. This case exists from a theoretical point of view. We can define the vulnerability of the community is infinity since it is an isolated node.

According to the proposed measure, the community vulnerability is inversely proportional to the five factors: the number of connected edges, the average strength of connected links, the connection density within the community, the average strength of links within the community, and the average degree of gateway nodes. Similarly, to compare the vulnerability of communities quantitatively, we have:

**Definition 3.3.** The relative vulnerability of a community $x$ is defined as,

$$R_x' = \frac{v_x'}{v};\quad v = \min_y(v_y') \tag{8}$$

### 3.3. An illustrative example

In this section, an illustrative example is given to demonstrate the difference between the proposed method and the measure developed in Ref. [16]. This example originates from the network shown in Fig. 1. Two different cases are considered in this section. On the one hand, we consider a community division that has the same topologies and structure with the unweighted network as shown in Fig. 1, and the weight associated with each edge is considered when evaluating the community structure vulnerability. On the other hand, the extended HAA is used to discover community structure in the weighted network, then we perform community structure vulnerability analysis.

In Fig. 1, the weights along all the edges are one (i. e. unweighted network). Thus, we have $(\overline{s}_x)^{\beta} = (\overline{s}_{x'})^{\delta} = 1$ for any value of $\beta$ and $\delta$. Suppose all the five factors are equally important, then we have $\alpha = \beta = \gamma = \delta = \lambda = 1$. According to Eqs. (7) and (8), the vulnerabilities of the three communities are shown in Table 1. For the sake of comparison, the relative community vulnerability of the network in Ref. [16] is also shown at the last column of Table 1.

From Table 1, it can be observed that the community $C$ has the lowest vulnerability, which is the same as that in Ref. [16]. However, the vulnerabilities of community $A$ and $B$ are different; community $A$ is more vulnerable than community $B$, as expected. Whereas, based on Eqs. (5) and (6), both $A$ and $B$ are estimated to have the same vulnerability.

To illustrate the difference between unweighted networks and weighted networks, we consider a network as shown in Fig. 2, and it has the same community structure with Fig. 1, but the weight associated with each link is different. As indicated in Fig. 2, the number along each link represents the strength of that particular link, which is proportional to its thickness. From Fig. 2, the links among the communities have different strength values. For example, the links 5 – 6 and 4 – 9 connect community $B$ with $C$, and the links 1 – 6 and 2 – 7 connect communities $A$ and $C$. Even though the number of links among these communities is the same, each link that connects the communities has different strength in coping with the attack. According to Eq. (7), we calculate the vulner-

**Table 2**
The community vulnerability for the network in Fig. 2 using the proposed method.

| Communities | $\bar{s}_x$ | $\bar{s}_{x'}$ | $v'_x$ | $R'_x$ |
|---|---|---|---|---|
| $A$ | 1.00 | 0.2142 | 28.0112 | 8.2963 |
| $B$ | 0.8889 | 1.00 | 3.375 | 1.1384 |
| $C$ | 0.9444 | 0.3571 | 2.9647 | 1.00 |

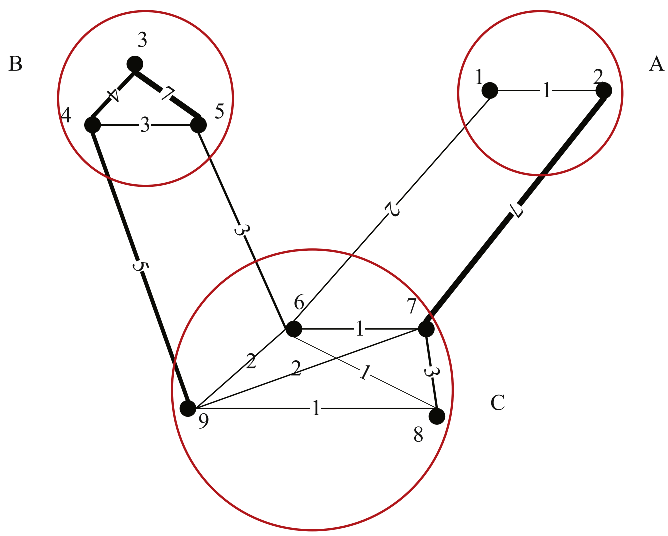


**Fig. 2.** Community structure illustration in a weighted network.

ability for each community shown in Fig. 2. Compared with Fig. 1, the values of $|V_x|$, $\rho_{x'}$, and $\bar{d}_{x''}$ remain the same. For $\bar{s}_x$ and $\bar{s}_{x'}$, we have,

$$\bar{s}_B = \frac{5+3}{2} = 4,\ \bar{s}_{B'} = \frac{7+3+4}{3} = 4.6667$$

Suppose $\alpha = \beta = \gamma = \delta = \lambda = 1$, the vulnerability and relative vulnerability of each community in the network are shown in Table 2. As can be observed from Table 2, community $A$ is much more vulnerable than communities $B$ and $C$. The results indicate that the vulnerability of weighted network is different from the unweighted network even if their topologies are the same.

On the other hand, if the weight associated with each link is considered, the communities discovered in the network will change accordingly. Fortunately, an extended HAA algorithm has been developed to discover communities in weighted networks. The extended HAA algorithm divides the weighted network shown in Fig. 2 into two communities ($Q_w = 0.322$). The two communities are denoted as communities $D$ and $E$, and they are shown in Fig. 3. Nodes 3, 4, 5 and 9 are grouped into community $E$, while the nodes 1, 2, 6, 7 and 8 are grouped into community $D$.

The community structure shown in Fig. 3 differs from the communities discovered in the unweighted network (see Fig. 1). With respect to the modularity value $Q$, the unweighted network has a value of 0.29, while the modularity of weighted network is 0.32. By comparing the modularity value in the two different cases, we can see that the weight associated with each edge strongly affects the community structure. Suppose $\alpha = \beta = \gamma = \delta = \lambda = 1$, according to Eqs. (5) and (6), the vulnerability and relative vulnerability of communities $D$ and $E$ in this network are shown in Table 3.

From Table 3, it can be observed that community $D$ is more vulnerable than community $E$ when several other connectivity factors are considered, e.g., the connection density, the degree of gateway nodes, and the strength of links within each community, by which we have $R'_D > R'_E$. However, as indicated at the last row of Table 3, the measure developed in Ref. [16] cannot differentiate the vulnerability of the

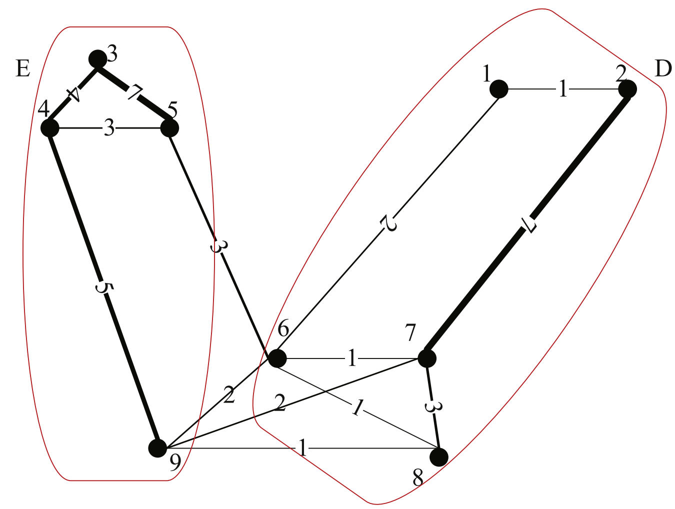


**Fig. 3.** A partition of communities for a weighted network shown in Fig. 2 by the extended HAA.

two communities $D$ and $E$. Such comparison demonstrates that the developed vulnerability measure in this paper has a better resolution in distinguishing the vulnerability levels related to communities than the one in Ref. [16].

### 3.4. Sensitivity analysis

As can be observed from Eq. (7), the weights (i.e., $\alpha$, $\beta$, $\gamma$, $\delta$, and $\lambda$) play an important role in the quantification of community structure vulnerability. The variability of the weight associated with each decision criterion results in the variation of community structure vulnerability evaluation. However, how to determine the weight associated with each decision criterion is an open issue. In this paper, we employ variance-based global sensitivity analysis to quantify the impact of the weight variability on the community vulnerability quantification. Typically, global sensitivity analysis is used to quantify the influence of stochastic model inputs on the output variability of a physical or mathematical model [33]. Among the various sensitivity measures, Sobol' indices based on variance decomposition have received a lot of attention [34]. The Sobol' indices is defined as [35]

$$S_i^I = \frac{V_{x_i}\left(\mathbb{E}_{x_{-i}}(y|x_i)\right)}{V(y)},\ S_i^T = \frac{\mathbb{E}_{x_{-i}}\left(V_{x_i}(y|x_{-i})\right)}{V(y)} = 1 - \frac{V_{x_{-i}}\left(\mathbb{E}_{x_i}(y|x_{-i})\right)}{V(y)}. \tag{9}$$

where $y$ is the system output, $\boldsymbol{x} = (x_1, \cdots, x_n)$ are $n$ independent input variables, $\boldsymbol{x}_{-i}$ means all the model inputs other than $x_i$, $S_i^I$ is referred to as first-order Sobol' index quantifying the contribution of $x_i$ alone to the variance of $y$, $V(y)$ measures the variance when all the input variables change, and $V_{x_i}\left(\mathbb{E}_{x_{-i}}(y|x_i)\right)$ quantifies the variance when $x_i$ is fixed and all the other variables $x_{-i}$ change. Different from $S_i^I$, $S_i^T$ is referred to as total effect index. As implied by the name, $S_i^T$ measures the contribution to the output variance of $y$ caused by the variability of each input variable considering both its individual effect and its interactions with all other variables.

With respect to the community structure vulnerability assessment, we generate many different realizations to represent the possible set of weights with their values uniformly distributed in the range [0, 5] that can be associated with every decision criterion, then we calculate the community vulnerability for all the randomly generated weight combinations, and quantify how the combinations of different weights impact the community vulnerability measure with the first-order Sobol' index.

**Table 3**
The community structure vulnerability for the weighted network shown in Fig. 3.

| Community | $\lvert V_x \rvert$ | $\bar{s}_x$ | $\rho_{x'}$ | $\bar{s}_{x'}$ | $\bar{d}_{x''}$ | $v'_x$ | $R'_x$ | $R_x$ (method of Ref. [16]) |
|---|---|---|---|---|---|---|---|---|
| $D$ | 1 | 1 | 0.9 | 0.5263 | 1 | 2.3457 | 1.3194 | 1.00 |
| $E$ | 1 | 1 | 1 | 1 | 0.5625 | 1.7778 | 1.00 | 1.00 |

**Table 4**
Community sets in Italian 380KV power grid.

| Communities | Nodes |
|---|---|
| 1 | 1,2,3,4,5,6,7,8,9,11,12,13,19,20 |
| 2 | 30,31,32,33,34,35,36,37,38,39,42,60 |
| 3 | 43,44,45,46,47,49,50,51,52,54,55,56,57 |
| 4 | 10,15,16,17,18,21,22,23,24,25,26,27,28,29,58,59 |
| 5 | 40,41,48,53,61,62,63,64,65,68,66,67, 69,70,71 |
| 6 | 14,72,73,74,75,76,77,78,79,81,82 |
| 7 | 83,84,85,86,100,101 |
| 8 | 102,110,111,112,113,114,115,116,117,118,120 |
| 9 | 80,87,88,89,90,91,92,93,94,95,96,97,98,99,103 104,105,106,107,108,109,119 |
| 10 | 121,122,123,124,125,126,127 |

## 4. Applications to real-world networks

In this section, three real-world networks are used to illustrate the effectiveness of the proposed method.

### 4.1. Italian 380KV power grid

The Italian 380KV power transmission grid has been frequently used to study network vulnerability assessment techniques [19,36,37]. In the Italian 380KV power grid, nodes represent power sources and substations, and links represent the transmission lines. Using the HAA algorithm introduced in Section 2, and we divide the network into 10 communities ($Q = 0.76$), as shown in Table 4 and Fig. 4 [16].

From Table 4, it can be seen that community 9 has 22 nodes, which is the largest community among all the communities. Community 7 only has 6 nodes, and it is the smallest community. Suppose $\alpha = \beta = \gamma = \delta = \lambda = 1$, according to Eqs. (7) and (8), the community structure vulnerability and their relative values of the Italian 380KV power grid are reported in Table 5. At the last column of Table 5, the relative vulnerability of the communities obtained by Eqs. (5) and (6) in Ref. [16] are also reported for the sake of comparing the results yielded by the two different methods.

From Table 5, we can notice that the most vulnerable community is community 10, which is consistent with the result in Ref. [16]. From Fig. 4, we observe that community 10 has a series connection, in which only one link is connected with other communities. The least vulnerable community is community 7. Although community 10 has more nodes than community 7, community 10 is still the most vulnerable community while community 7 is the least vulnerable community. Such fact implies that there is no correlation between vulnerability and the number of nodes in each community, while the topology of communities plays an important role in affecting the community vulnerability.

In Ref. [16], the ten communities are ranked in five levels according to their relative vulnerability values, i.e., $R_{10} > R_1 = R_3 = R_7 = R_8 > R_2 = R_4 = R_6 > R_9 > R_5$. However, the ten communities have unique vulnerability values in the updated relative community vulnerability measure, i.e., $R_{10} > R_9 > R_1 > R_8 > R_3 > R_4 > R_2 > R_5 > R_6 > R_7$, which indicates that the proposed method has a better resolution in differentiating the vulnerability of separate communities. Specifically, communities (1, 3, 7, 8) and (2, 4, 6) have different vulnerability values now.

Next, following the method introduced in Section 3.4, we perform global sensitivity analysis for the community structure vulnerability measure with respect to the variability of the weights associated with the five attributes under consideration. Table 6 reports the sensitivity analysis results for the vulnerability measure related to each community structure with respect to the five weight variables $\alpha$, $\beta$, $\gamma$, $\delta$ and $\lambda$ in the Italian 380KV power grid. As can be noticed, with respect to the first-order Sobol's index, the vulnerability of community 1 is most sensitive to the variability of the weight variable $\alpha$, followed by the weight variables $\gamma$ and $\lambda$. Both the first-order and total effect sensitivity indices ($S^I_\beta, S^T_\beta, S^I_\delta$ and $S^T_\delta$) of the vulnerability measure to the weight variables $\beta$ and $\delta$ are almost zero because $\bar{s}_x$ and $\bar{s}_{x'}$ are one. Thus, no matter how the weight variables $\beta$ and $\delta$ change, the values of $(\bar{s})^\beta_x * (\bar{s}_{x'})^\delta$ remain the same, which results in that the effect of the variability of these two weight variables on the vulnerability measure of the first community is negligible.

Another observation regarding the first-order sensitivity index is that the smaller the attribute value, the more sensitive the community vulnerability measure is to the weight associated with that attribute. This is because the vulnerability measure defined in Eq. (7) is the product of the reciprocals of the normalized values of the five factors under consideration; hence the factor with lower value has a larger impact. Such patterns can also be observed from the sensitivity analysis results of all the other communities. For example, the vulnerability measure of the second community is sensitive to the variability of weight variable $\gamma$ the most, followed by the weight variables $\alpha$ and $\lambda$. Such order ($S^I_\gamma > S^I_\alpha > S^I_\lambda > S^I_\beta = S^I_\delta$) is consistent with the ranking of the values of the five factors ($\rho_{x'} < \lvert V_x \rvert < \bar{d}_{x''} < \bar{s}_x = \bar{s}_{x'}$), see Table 5 for details.

Since the sum of the first-order sensitivity index over the five variables for each community is far less than 1, it reveals that there is an important contribution due to the presence of interactions among the five variables. Such fact is also reflected in the total effect sensitivity index of each variable, as shown in the third, seventh, and eleventh column of Table 6. Due to the presence of strong interactions among the considered variables, large differences arise between the first-order index and the total effect index. An interesting observation is that the ranking of the total effect indices is consistent with the ranking of the first-order indices in this problem. For example, regarding community 1, the ranking of the five variables with respect to the first-order sensitivity index is: $S^I_\gamma > S^I_\alpha > S^I_\lambda > S^I_\beta = S^I_\delta$. Obviously, the order is in accordance with the ranking of total effect index: $S^T_\gamma > S^T_\alpha > S^T_\lambda > S^T_\beta = S^T_\delta$. Since both the first-order and total effect indices can be computed in such a straightforward manner and provide the same insight as gained by the attribute values, we do not present the sensitivity analysis results for the subsequent examples.

### 4.2. Karate club network

In this subsection, a social network, named the karate club network, is used to verify the performance of the developed community vulnerability measure. In the karate club network, there are 34 nodes and 156 edges, and they represent the friendships among 34 members of a karate club at a US university [20], and its specific structure is illustrated in Fig. 5.

In the karate club network, nodes represent the members of the club, including the administrators and instructors, and links denote the interactions beyond the normal activities of the club (karate classes and club

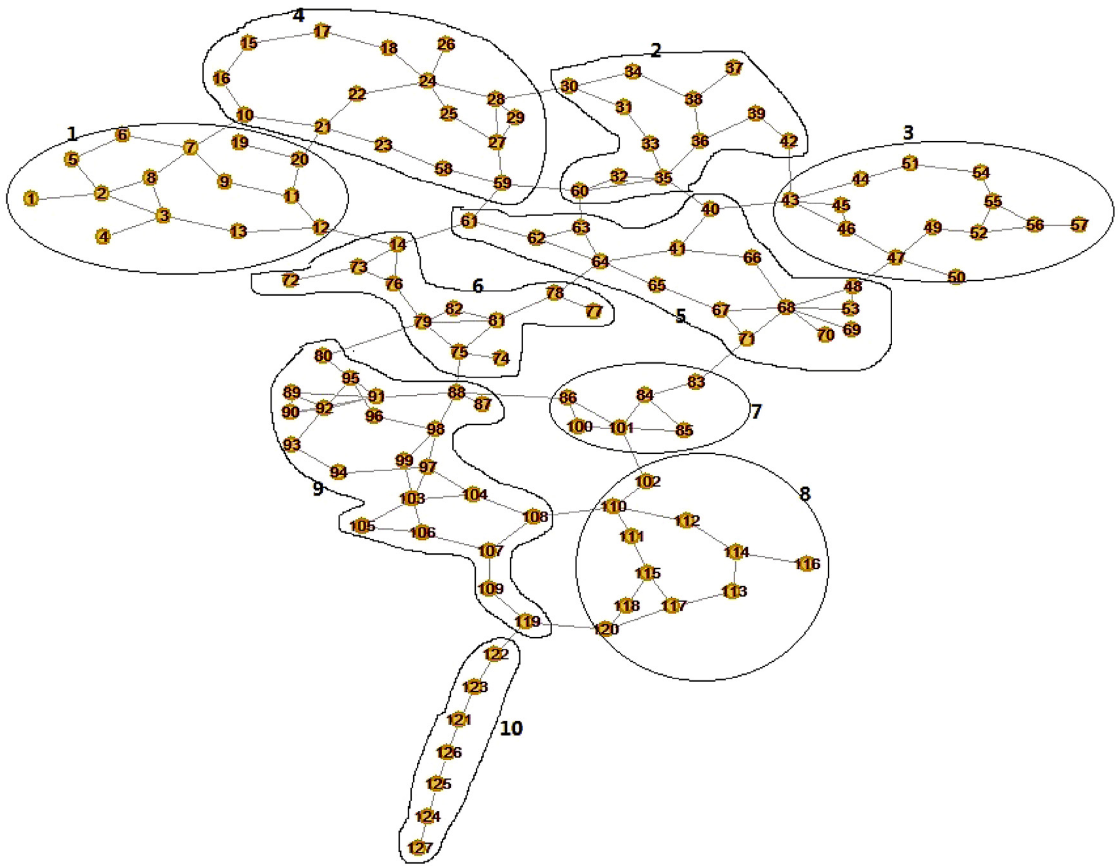

**Fig. 4.** Communities in Italian 380KV power grid [38].

**Table 5**
Italian 380KV power grid: community vulnerability measures.

| Community | $\|V_x\|$ | $\bar{s}_x$ | $\rho_{x'}$ | $\bar{s}_{x'}$ | $\bar{d}_{x''}$ | $\upsilon'_x$ | $R'_x$ | $R_x$ [16] |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.375 | 1.00 | 0.3767 | 1.00 | 0.7777 | 9.1020 | 2.6547 | 2.67 |
| 2 | 0.625 | 1.00 | 0.4221 | 1.00 | 0.75 | 5.0539 | 1.4741 | 1.60 |
| 3 | 0.375 | 1.00 | 0.3846 | 1.00 | 1.00 | 6.9333 | 2.0222 | 2.67 |
| 4 | 0.625 | 1.00 | 0.3392 | 1.00 | 0.8333 | 5.6606 | 1.6510 | 1.60 |
| 5 | 1 | 1.00 | 0.3673 | 1.00 | 0.6667 | 4.0843 | 1.1912 | 1.00 |
| 6 | 0.625 | 1.00 | 0.5065 | 1.00 | 0.9167 | 3.4459 | 1.0050 | 1.60 |
| 7 | 0.375 | 1.00 | 1.00 | 1.00 | 0.7777 | 3.4286 | 1.00 | 2.67 |
| 8 | 0.375 | 1.00 | 0.4675 | 1.00 | 0.6667 | 8.5554 | 2.4953 | 2.67 |
| 9 | 0.625 | 1.00 | 0.2876 | 1.00 | 0.5833 | 9.5387 | 2.7821 | 1.33 |
| 10 | 0.125 | 1.00 | 0.6122 | 1.00 | 0.3333 | 39.2052 | 11.4347 | 8.00 |

**Table 6**
Sensitivity analysis with respect to the weight variables $\alpha$, $\beta$, $\gamma$, $\delta$ and $\lambda$ related to the five decision factors in the Italian 380KV power grid.

| Community | $S^I_\alpha$ | $S^T_\alpha$ | $S^I_\beta$ | $S^T_\beta$ | $S^I_\gamma$ | $S^T_\gamma$ | $S^I_\delta$ | $S^T_\delta$ | $S^I_\lambda$ | $S^T_\lambda$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.2572 | 0.7166 | 0.0002 | 0.0000 | 0.2574 | 0.7332 | 0.0002 | 0.0000 | 0.0222 | 0.1380 |
| 2 | 0.1674 | 0.3966 | 0.0002 | 0.0000 | 0.4784 | 0.7516 | 0.0002 | 0.0000 | 0.0662 | 0.1860 |
| 3 | 0.3001 | 0.6882 | 0.0002 | 0.0000 | 0.2902 | 0.6820 | 0.0002 | 0.0000 | 0.0002 | 0.0000 |
| 4 | 0.1226 | 0.3714 | 0.0002 | 0.0000 | 0.4830 | 0.8408 | 0.0002 | 0.0000 | 0.0206 | 0.0822 |
| 5 | 0.0002 | 0.0000 | 0.0002 | 0.0000 | 0.6306 | 0.8661 | 0.0002 | 0.0000 | 0.1349 | 0.3348 |
| 6 | 0.2444 | 0.4757 | 0.0002 | 0.0000 | 0.4648 | 0.7099 | 0.0002 | 0.0000 | 0.0095 | 0.0241 |
| 7 | 0.8363 | 0.9279 | 0.0002 | 0.0000 | 0.0002 | 0.0000 | 0.0002 | 0.0000 | 0.0728 | 0.1733 |
| 8 | 0.2686 | 0.7893 | 0.0002 | 0.0000 | 0.1778 | 0.5952 | 0.0002 | 0.0000 | 0.0588 | 0.2831 |
| 9 | 0.0771 | 0.3785 | 0.0002 | 0.0000 | 0.4200 | 0.8155 | 0.0002 | 0.0000 | 0.1002 | 0.4120 |
| 10 | 0.2177 | 0.9111 | 0.0002 | 0.0000 | 0.0237 | 0.3095 | 0.0002 | 0.0000 | 0.0909 | 0.6649 |

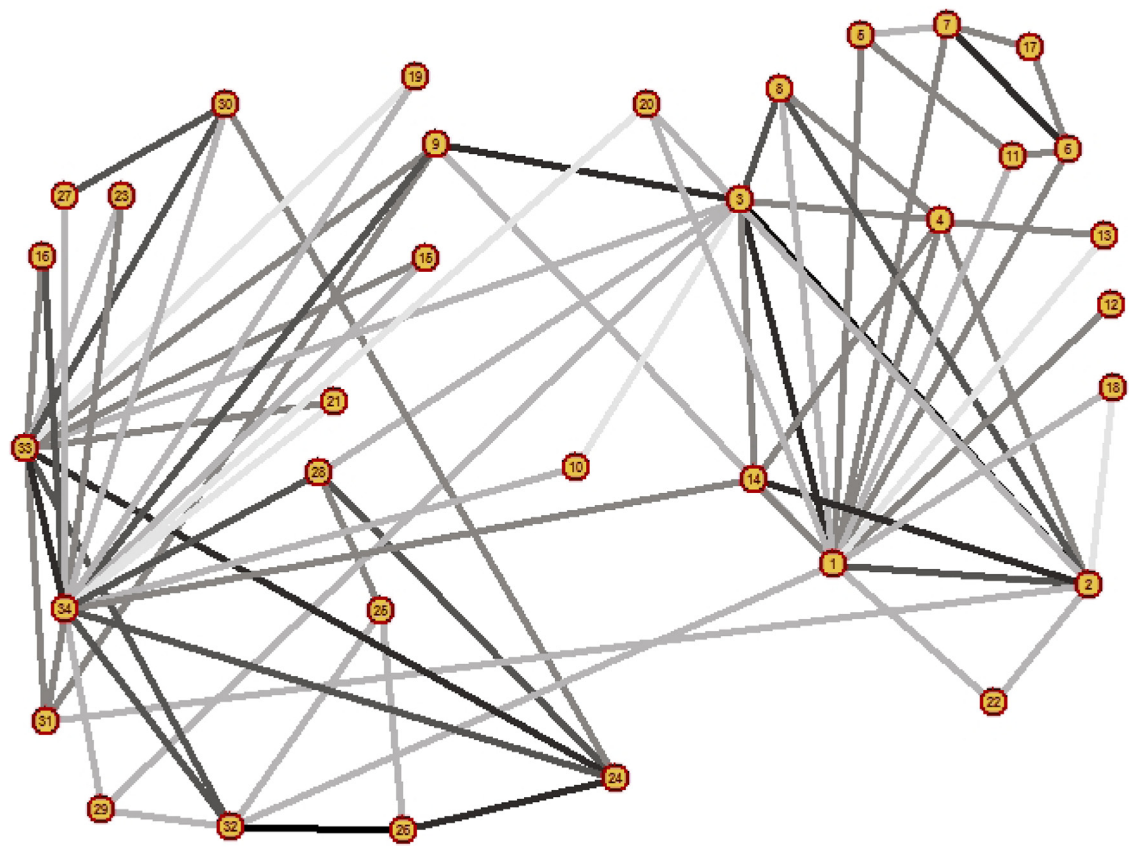

**Fig. 5.** The karate club network. Nodes represent the administrator and the instructor of the karate club network, links are interaction between two individuals, and the weights of the edges denote the strengths of ties between individuals. The darker the line, the stronger the strength.

**Table 7**
Karate club network: community vulnerability measures .

| Communities | $\|V_x\|$ | $\overline{s}_x$ | $\rho_{x'}$ | $\overline{s}_{x'}$ | $\overline{d}_{x''}$ | $v'_x$ | $R'_x$ | $R_x$[16] |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.00 | 1.00 | 1.00 | 1.00 | 1.0294 | 0.9714 | 1.00 | 1.00 |
| 2 | 1.00 | 1.00 | 1.0083 | 1.00 | 1.00 | 0.9918 | 1.0210 | 1.00 |
| $1'$ | 1.00 | 1.00 | 1.1159 | 1.00 | 1.0294 | 0.8705 | 1.00 | 1.00 |
| $2'$ | 1.00 | 1.00 | 1.00 | 1.0476 | 1.00 | 0.9545 | 1.0966 | 1.00 |

meetings) between two individuals, and the link weight is the strength of ties between the individuals, which is estimated using a variety of measure by Zachary, see Ref. [20] for details. In the karate club network, the maximum strength value is 7, which exists between node 26 and 32, while the minimum strength value is 1, which is between node 1 and 13. The community structure of karate club network is already well-known, and it is divided into two communities because of the disagreement that emerged between the instructors and the administrators [39].

The karate club network will degenerate to an unweighted network if the strength of links between any two individuals is ignored. Using HAA, the unweigthed network is divided into two communities ($Q = 0.38$), and the number of nodes in each community is 17. As illustrated in Fig. 6, the two communities are represented by different shapes. The division found by HAA, denoted as community 1 and community 2, almost perfectly matches the real karate club community except that one node (node 10) is grouped into the wrong community. According to Eqs. (7) and (8), the vulnerability of each community is calculated, and the results are demonstrated in the first and second rows of Table 7.

For the weighted karate club network, the analysis result by the extended HAA algorithm is shown in Fig. 7 ($Q_w = 0.4345$). As can be observed, the weighted network is also divided into two communities, which are denoted as community $1'$ and community $2'$, respectively. According to Eqs. (7) and (8), the vulnerabilities of communities $1'$ and $2'$ are calculated, which are shown in the third and fourth rows of Table 7. It can be seen that there are always two communities no matter whether we treat the karate club network as an unweighted or weighted graph, and the number of links between two communities is always the same. As a result, the vulnerabilities of the two communities are always equal to each other according to the method developed in Ref. [16], and their values are shown in the last column of Table 7.

Figs. 6 and 7 indicate that the weights of edges have important role in determining the community structure. The relationships among individuals are better characterized in the weighted network because some individuals are much closer friends than others. For example, the link between node 32 and 26 has the largest weight, which indicates they have the most intimate relationship. According to the method of modularity, every node is considered as a single community in the first place. Then the two communities $i$ and $j$ are grouped into the same community based on their similarities. From Fig. 7, nodes 26 and 32 are combined into a single community in the first step while node 26 is combined with node 25 at the first step in Fig. 6.

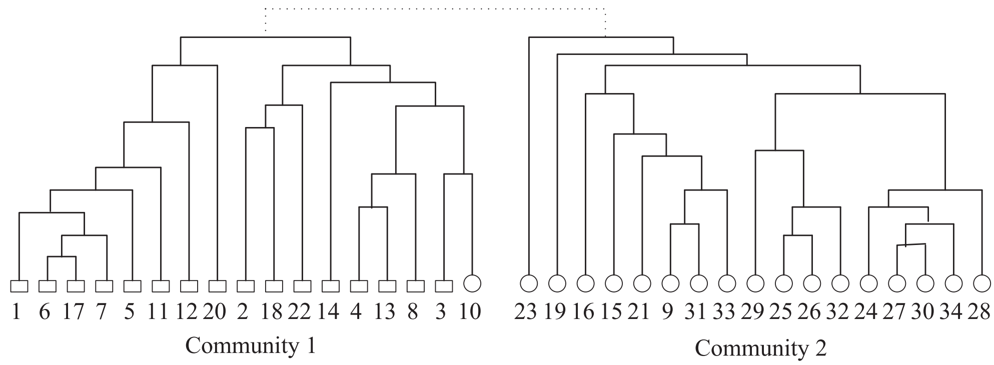


**Fig. 6.** Two communities are found by HAA in the unweighted karate club network. The shapes of the vertices represent the community that the nodes are grouped into.

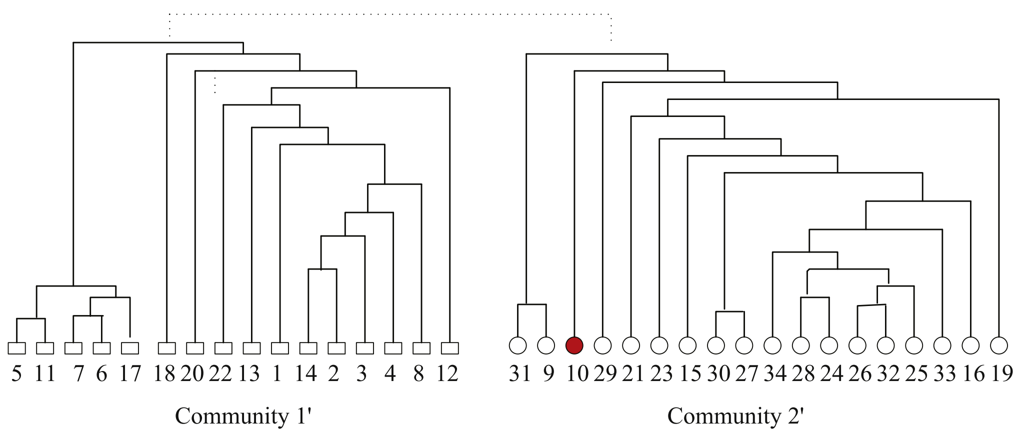


**Fig. 7.** Two communities found in the weighted karate club network by the weighted HAA algorithm: The shapes of the vertices represent the community that the nodes are grouped into. Compared with Fig. 6, node 10 is grouped into the correct community.

From Table 7, we have $R_2{}' > R_1{}'$, $R_{2'}{}' > R_{1'}{}'$, $R_2 = R_1$ and $R_{2'} = R_{1'}$. The vulnerability of communities can be distinguished by the proposed method no matter whether we consider it as an unweighted karate club network or a weighted karate club network. Whereas, the method in Ref. [16] obtains the same vulnerability values for this network.

### *4.3. Air transportation network*

The vulnerability of air transportation network is a topic of increasing interest [40,41]. The community vulnerability in the air transportation network measures the capability of the airlines in withstanding the disturbances and maintaining the nominal operations in the presence of unexpected events, such as extreme weather. The higher the vulnerability, the more fragile the air transport network. For the USAir97 network with 332 nodes and 2126 edges [21], nodes and edges represent US airports and direct flights among the airports, respectively. In this section, we consider the vulnerability of air transportation networks from two different cases to illustrate effectiveness of proposed method.

#### *4.3.1. Community detection in the unweighted USAir97 network*

We first consider the unweighted USAir97 network. Using HAA, the USAir97 network is divided into 7 communities and $Q = 0.32$, as shown in Fig. 8. Community 4 has the most number of nodes (150 nodes), and community 5 has the least number of nodes (2 nodes). Suppose $\alpha = \beta = \gamma = \delta = \lambda = 1$. Using Eqs. (7) and (8), the relative vulnerability of each community is calculated and the result is shown at the seventh column of Table 8.

From Table 8, we have $R_4 < R_2 < R_5 < R_3 < R_6 < R_1 < R_7$. The most vulnerable community is community 7, and the least vulnerable community is community 4. The vulnerability value of community 7 is almost 671 times that of community 4 from a comparative perspective. In practice, a precise numerical value such as 671 is not directly meaningful, except to imply that community 4 has a much stronger capability to withstand an attack than community 7. With respect to external factors, there is big disparities between community 7 and community 4. In community 7, only node 329 is connected with community 2 (i.e., $|V_7| = 1$), while $|V_4| = 414$. In practical world, node 329 represents ”Guam Intl airport”, and it is connected with ”Saipan Intl airport”, ”Rota Intl airport” and ”Koror airport”. Community 7 is on the ”periphery” of the air transport network, and the community vulnerable is the most biggest. In community 4, external factors, the value of $|V_4|$ is 414, which is much larger than for other communities except community 2; and its average strength of external links ranks the third among all the communities.

Consider internal factors, the value of $\overline{d}_{4'}$ is the maximum, which is almost ten times as that of community 7. $\overline{s}_{4'}$ is also the maximum. Thus, community 4 has the lowest vulnerability. In fact, there are many airports with good capacity to withstand an attack in community 4. Many large airports are present in community 4, such as node 118 (“Chicago O'Hare Intl”), and it has 40 links that are connected to other communities. In complex networks, these nodes are named as “hub” nodes since they have higher node degree. Until 1998, Chicago O'Hare was the busiest airport in the world in terms of number of passengers (https://en.wikipedia.org/wiki/O%27Hare_International_Airport). In other words, it can handle the disruption caused by the unexpected events by diverting the impacted aircraft to other airports within the community or to other airports outside the community. For internal factor, “Chicago O'Hare Intl” has 99 links for transporting resources to other airports within community 4. In one word, “Chicago O'Hare Intl” has a better absorption capacity from the outside world and transport capacity within community.

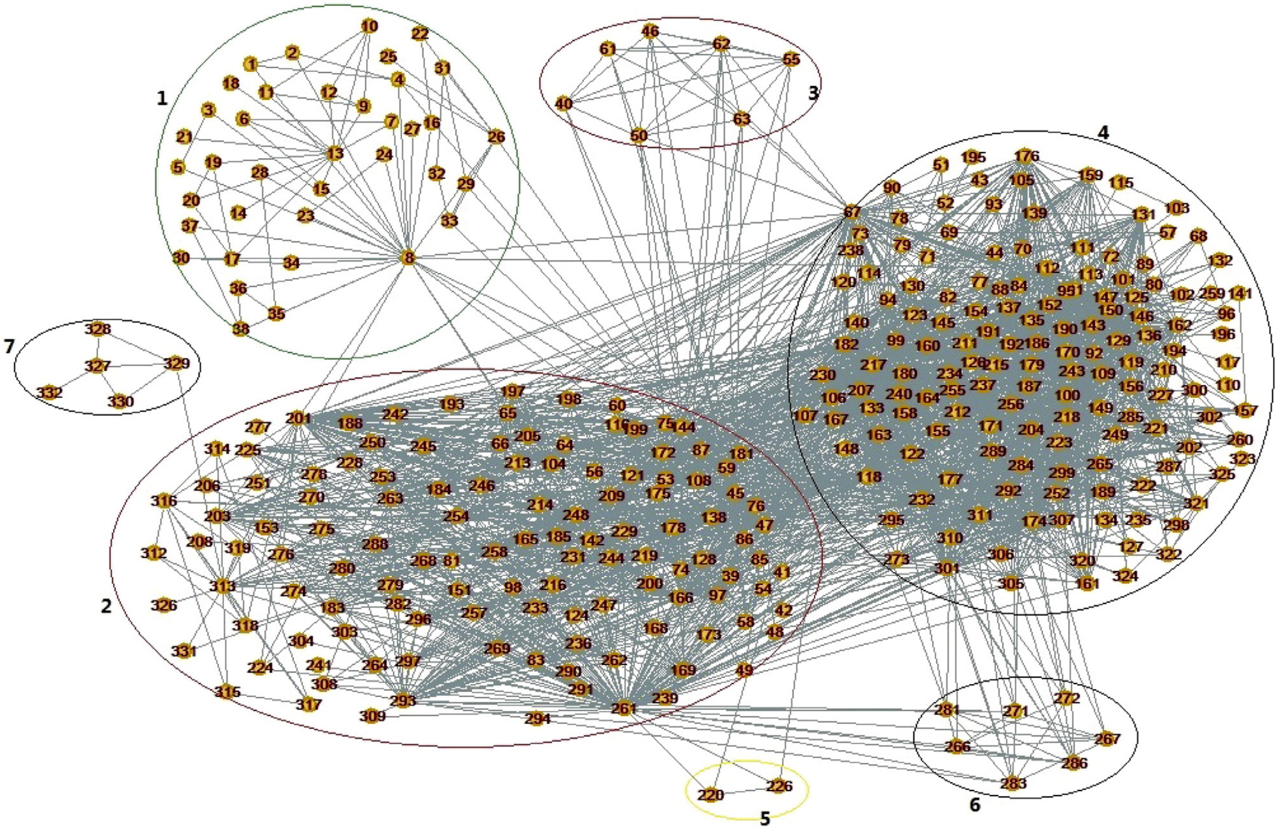


**Fig. 8.** Communities in USair97 network.

**Table 8**
Community vulnerability for Fig. 8.

| Community | $\lvert V_x \rvert$ | $\bar{s}_x$ | $\rho_{x'}$ | $\bar{s}_{x'}$ | $\bar{d}_{x''}$ | c$R'_x$(unweighted) | c $R'_x$(weighted) | c $R_x$(method of Ref [16].) |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.0309 | 0.8276 | 0.0896 | 0.6447 | 0.2368 | 140.189 | 113.7035 | 32.3887 |
| 2 | 1 | 0.4351 | 0.0704 | 0.9915 | 0.5324 | 2.4503 | 2.4581 | 1.00 |
| 3 | 0.0475 | 0.1906 | 0.9048 | 0.2860 | 0.1737 | 12.3006 | 97.6475 | 21.0526 |
| 4 | 0.9834 | 0.4327 | 0.0934 | 1.00 | 1.00 | 1.00 | 1.00 | 1.0168 |
| 5 | 0.0095 | 0.1093 | 1.00 | 0.1083 | 0.032 | 3.1335 | 302.0906 | 11048.27 |
| 6 | 0.0736 | 0.1295 | 0.619 | 0.3469 | 0.1189 | 16.9538 | 2.7707 | 163.2985 |
| 7 | 0.0024 | 1.00 | 0.6 | 0.5804 | 0.096 | 671.3124 | 3.4665 | 500.556 |

Next we consider the weighted USAir97 network. The weight of each edge corresponds to the number of seats available on the scheduled flights (unit: million/year) [21]. We consider how the weight of the edge affects the community vulnerability because the value of strength of both external and internal connectivity of the community may get changed. A simple way is to assume that the community stays the same as that of the unweighted USAir97 network for comparison purposes. According to Eqs. (7) and (8), their relative vulnerabilities are calculated and shown in the last column of Table 8. As can be observed from the last two columns of Table 8, community 4 has the lowest vulnerability because it has a better capacity to withstand an attack, considering both external and internal factors. Note that the vulnerability of community 5 is higher than that of community 7 the strength of the links is taken into consideration.

If we utilize the metric developed in Ref [16]., the results are indicated at the last column of Table 8. Obviously, the vulnerability of each community has changed. But the most vulnerable community remains the same for the two measures ($R_x$ and $R'_x$). If we only consider the number of links between communities according to method in Ref. [16]. The vulnerability $R_x$ is shown in the last column of Table 8, where $x = (1, 2, 3, 4, 5, 6, 7)$. As can be seen from Table 8, the community vulnerability values have changed.

#### 4.3.2. Community detection in weighted USAir97 network

As mentioned earlier, if the weight associated with each link is considered, the communities discovered in the network will change accordingly. In this section, we utilize the extended HAA to identify the communities in the weighted USAir97 network. As shown in Fig. 9, the weighted USAir97 network is divided into four communities. Since $Q_w$ ($Q_w = 0.20$) has a value far away from the lower bound of the recommended range, it indicates that the weighted USAir97 network does not present an explicit community structure. Due to the impact of link weights, some communities discovered in the unweighted network (see Fig. 9) are merged when link weight is considered. For example, as shown in Fig. 8, there are only 5 nodes in community 7, and they are nodes 327, 328, 329, 330 and 332, and only one node 329 is connected with other communities, see the link 329-313. However, these nodes are grouped into community 2 in the weighted USAir97 network, see Fig. 9 for details.

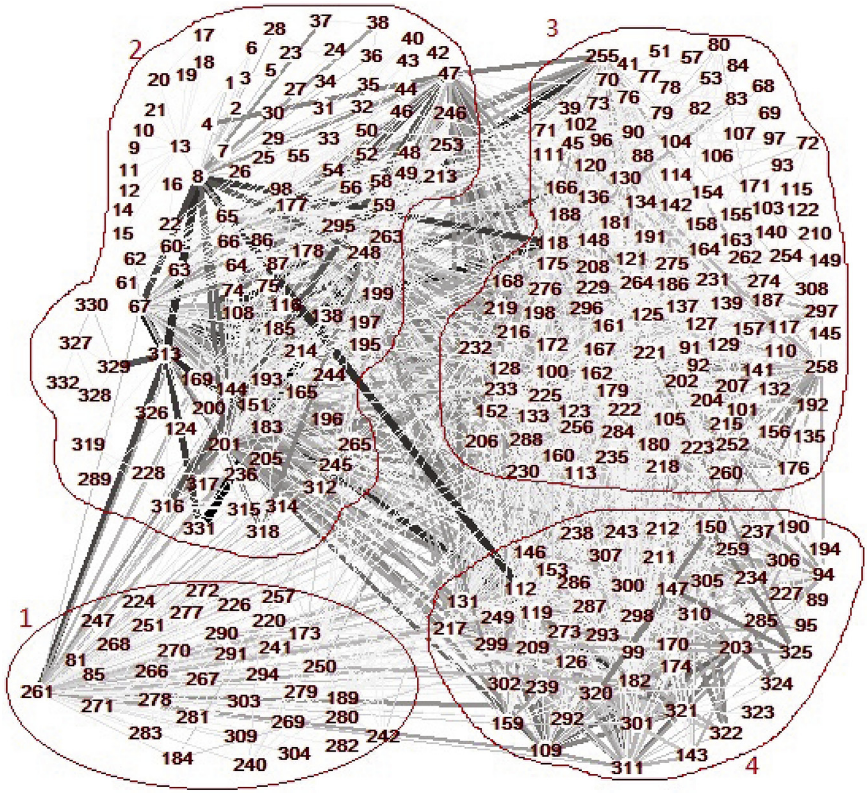


**Fig. 9.** Communities discovered in the weighted USAir97 network.

**Table 9**
The community structure vulnerability analysis in the weighted USAir97 network.

| Communities | $\|V_x\|$ | $\bar{s}_x$ | $\rho_{x'}$ | $\bar{s}_{x'}$ | $\bar{d}_{x''}$ | $v'_x$ | $R'_x$ | $R_x$ (method in Ref [16].) |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.3636 | 0.6061 | 0.3152 | 0.5227 | 0.1883 | 146.2658 | 127.2659 | 2.7503 |
| 2 | 0.737 | 1 | 0.1935 | 0.7445 | 0.4728 | 19.9210 | 17.3332 | 1.3568 |
| 3 | 0.9058 | 0.7412 | 0.2776 | 0.6381 | 0.2875 | 29.2473 | 25.4481 | 1.1040 |
| 4 | 1.00 | 0.8701 | 1.00 | 1.00 | 1.00 | 1.1493 | 1.00 | 1.00 |

The above analysis reveals the weight along each edge is an important consideration when we discover communities in networks. Given the newly discovered community structure, the vulnerability measure is updated accordingly. According to Eqs. (7) and (8), the vulnerability and relative vulnerability of four communities are calculated and the results are shown in Table 9. The relative vulnerability calculated by method in Ref. [16] is also shown at the last column of Table 9 for the sake of comparison.

As can be seen from Table 9, the vulnerability of all the communities can be ranked in the order: $R'_4 < R'_2 < R'_3 < R'_1$. Among the four communities, the first community is most vulnerable, and community 4 is the least vulnerable. Out of the five factors, community 4 has the maximum values for four of them, and they are the number of links between communities, the connection density within community 4, the degree of gateway nodes, and the strength of links. In contrast, the values related to the same four factors for community 1 are the lowest among all the communities. Thus, community 4 has the least vulnerability and community 1 is the most vulnerable community. The community vulnerability yielded by the method in Ref. [16] has an order: $R_4 < R_3 < R_2 < R_1$. It can be observed that the proposed method has a totally different ranking for the second and third community compared with the approach in Ref. [16] because the method in Ref. [16] does not consider the weight associated with each link and several other good vulnerability indicators as well. Last but not the least, the proposed method offers a better resolution in differentiating the vulnerability level related to each community in comparison with the measure developed in Ref. [16], as indicated in the last two columns of Table 9.

## 5. Conclusion

A generalized vulnerability measure of community structure is proposed in this paper based on the number and average strength of external links to other communities, the connection density and the average strength of internal links within the community, and the average degree of gateway nodes within the community. The proposed measure helps to distinguish the vulnerability values of the separate individual communities.

The proposed method is used to analyze the community vulnerability in an Italian 380KV power transmission grid network example, and the results show a clear advantage for the proposed measure in assigning a distinct rank to each community according to its vulnerability. This advantage is also reflected in the USAir97 network and the karate club network.

In our proposed method to measure the vulnerability of community structure, the relationships between the five factors are described by five parameters. The previously available metric [16] is shown to be a special case. The values of the five parameters strongly influence the vulnerability estimates. Further work is needed to verify the feasibility of using the proposed community vulnerability in practice, especially how to estimate the values of the weights for each of the five factors. In the numerical examples, all five factors were assumed to have equal weights, only for the sake of illustration. In different applications, the weights may be unequal, and expert judgment may be needed to derive the appropriate weights.

## Acknowledgment

The authors are grateful to the reviewers' thorough comments, which have helped to improve the quality of this paper to a new level. The authors are also grateful to acknowledge the financial support from China Scholarship Council, National Natural Science Foundation of China (Grant Nos. 61364030 and 61763009), and the Scientific Research Foundation of Educational Commission of Hubei Province of China (Grant No. D20151902).

## References

[1] Radicchi F, Castellano C, Cecconi F, Loreto V, Parisi D. Defining and identifying communities in networks. Proc Natl Acad Sci 2004;101(9):2658–63.
[2] Clauset A, Newman ME, Moore C. Finding community structure in very large networks. Phys Rev E 2004;70(6):066111.
[3] Labatut V, Orman GK. Community struture characterization. Encyclopedía of social network analysis and mining. 2nd Edition. Springer, 2017; 2017.
[4] Orman GK, Karadeli O, ÇalıŞır E. Overlapping communities via k-connected ego centered groups. In: Proceedings of the 2015 IEEE/ACM international conference on advances in social networks analysis and mining 2015. ACM; 2015. p. 1598–9.
[5] Rocco CM, Moronta J, Ramirez-Marquez JE, K B. Effects of multi-state links in network community detection. Reliab Eng Syst Saf 2017;163:46–56.
[6] Orman GK, Labatut V, Plantevit JF, Boulicaut M. Interpreting communities based on the evolution of a dynamic attributed network. Soc Netw Anal Min 2015;5(1):20.
[7] Orman GK, Labatut V, Naskali AT. Exploring the evolution of node neighborhoods in dynamic networks. Physica A 2017;482:375–91.
[8] Raghavan UN, Albert R, Kumara S. Near linear time algorithm to detect community structures in large-scale networks. Phys Rev E 2007;76(3):036106.
[9] Ming-Sheng S, Duan-Bing C, Tao Z. Detecting overlapping communities based on community cores in complex networks. Chin Phys Lett 2010;27(5):058901.
[10] Newman ME. Fast algorithm for detecting community structure in networks. Phys Rev E 2004;69(6):066133.
[11] Newman M. Analysis of weighted networks. Phys Rev E 2004;70(5):056131.
[12] Newman M., Moore C. "fast modularity" community structure inference algorithm. http://cs.unm.edu/∼aaron/research/fastmodularity.htm/; 2017. Online; accessed 3 October 2017.
[13] Zio E, Piccinelli R, Sansavini G. An all-hazard approach for the vulnerability analysis of critical infrastructures. In: ESREL 2011; 2011. p. 2451–8.
[14] Holme P, Kim BJ, Yoon CN, Han SK. Attack vulnerability of complex networks. Phys Rev E 2002;65(5):056109.
[15] Torres-Vera MA, Canas JA. A lifeline vulnerability study in barcelona, spain. Reliab Eng Syst Saf 2003;80(2):205–10.
[16] Rocco CM, Ramirez-Marquez JE. Vulnerability metrics and analysis for communities in complex networks. Reliab Eng Syst Saf 2011;96(10):1360–6.
[17] International Federation of Red Cross and Red Crescent Societies. What is vulnerability? http://www.ifrc.org/en/what-we-do/disaster-management/about-disasters/what-is-a-disaster/what-is-vulnerability/; 2017. Online; accessed 20 March 2017.
[18] Carrington PJ, John S, Stanley W. Models and methods in social network analysis, 28. Cambridge University Press; 2005.
[19] Crucitti P, Latora V, Marchiori M. Locating critical lines in high-voltage electrical power grids. Fluctuation Noise Lett 2005;5(02):L201–8.
[20] Zachary WW. An information flow model for conflict and fission in small groups. J Anthropol Res 1977;33(4):452–73.
[21] Wei D, Liu Q, Zhang H, Hu Y, Deng Y, Mahadevan S. Box-covering algorithm for fractal dimension of weighted networks. Sci Rep 2013;3:3049.
[22] Freeman L. Centrality in social networks conceptual clarification. Soc Netw 1979;1(3):215–39.
[23] Gideon, Rosenblatt. What is vulnerability? http://www.the-vital-edge.com/what-is-network-density; 2017. Online; accessed 20 March 2017.
[24] Borruso G. Network density and the delimitation of urban areas. Trans GIS 2003;7(2):177–91.
[25] Lancichinetti A, Kivela M, Saramaki J, Fortunato S. Characterizing the community structure of complex networks. PLoS ONE 2010;5(8):e11976.
[26] Guimera R, Amaral LAN. Cartography of complex networks: modules and universal roles. J Stat Mech 2005;2005(2):P02001.
[27] Newman ME, Girvan M. Finding and evaluating community structure in networks. Phys Rev E 2004;69(2):026113.
[28] Kernighan BW, Lin S. An efficient heuristic procedure for partitioning graphs. Bell Syst Tech J 1970;49(2):291–307.
[29] Pothen A, Simon HD, Liou K-P. Partitioning sparse matrices with eigenvectors of graphs. SIAM J Matrix Anal Appl 1990;11(3):430–52.
[30] Orman GK, Labatut V, Cherifi H. Qualitative comparison of community detection algorithms. In: International conference on digital information and communication technology and its applications. Berlin, Heidelberg: Springer; 2011. p. 265–79.
[31] Wu F, Huberman BA. Finding communities in linear time: a physics approach. Eur Phys J B 2004;38(2):331–8.
[32] Barrat A, Barthelemy RV-A, Pastor-Satorras M. The architecture of complex weighted networks. Proc Natl Acad Sci USA 2004;101(11):3747–52.
[33] Zhang X, Mahadevan S. Aircraft re-routing optimization and performance assessment under uncertainty. Decis Support Syst 2017;96:67–82.
[34] Li C, Mahadevan S. An efficient modularized sample-based method to estimate the first-order sobol' index. Reliab Eng Syst Saf 2016;153:110–21.
[35] Marrel A, Iooss B, Laurent B, Roustant O. Calculations of sobol indices for the gaussian process metamodel. Reliab Eng Syst Saf 2009;94(3):742–51.
[36] Rosato V, Bologna S, Tiriticco F. Topological properties of high-voltage electrical transmission networks. Electr Power Syst Res 2007;77(2):99–105.
[37] Zio E, Petrescu C-A, Sansavini G. Vulnerability analysis of a power transmission system. In: Proc. international probabilistic safety assessment and management conference (PSAM); 2008. p. 18–23.
[38] Manzi E, Labbé M, Latouche G, Maffioli F. Fishman's sampling plan for computing network reliability. IEEE Trans Reliab 2001;50(1):41–6.
[39] Girvan M, Newman ME. Community structure in social and biological networks. Proc Natl Acad Sci 2002;99(12):7821–6.
[40] Wilkinson SM, Dunn S, Ma S. The vulnerability of the european air traffic network to spatial hazards. Nat Hazards 2012;60(3):1027–36.
[41] Voltes-Dorta A, Rodríguez-Déniz H, Suau-Sanchez P. Vulnerability of the European air transport network to major airport closures from the perspective of passenger delays: ranking the most critical airports. Transp Res Part A 2017;96:119–45.