Critical node identification model based on cascade failure and cybernetics

Yuchao Wang[a]

[a]*School of information and Communication Engineering,University of Electronic Science and technology,Chengdu,Sichuan 610056,China*

Abstract

The study of complex network controllability is of guiding significance for network robustness analysis. Based on the controllability theory, this paper proposes an algorithm for network vulnerability analysis and vulnerable node recognition. In this algorithm, the vulnerability degree of network structure is described by the key nodes in the drive node set, and the vulnerability degree of network attributes is described by the approaching degree of nodes and their neighbors. The structural vulnerability and attribute vulnerability are integrated to identify the network vulnerable nodes under the condition of cascading failure. The experiment is carried out in real network. Compared with the existing vulnerable node identification methods, this method can excavate the nodes with strong comprehensive vulnerability. These nodes play an important role in maintaining the stability of the network. The failure of such nodes can cause greater damage to the network.

Keywords: complex networks
cascade failure
cybernetics
important node identification

1. Introduction

In recent years, research based on interactions within networks, i.e., cascading failures, has been focused on by scholars in many fields at home and abroad, and the relationship between infrastructures in various fields has become closer. This

Email address: 2019011213025@std.uestc.edu.cn (Yuchao Wang)

inseparable coupling facilitates coordinated scheduling, but can also lead to some disastrous consequences. For example, in the power failure in Italy [1] and North America [2] in 2003, the failure of one power station strained the power load in many parts of the grid. Finally, it caused a chain reaction leading to large-scale power failure, resulting in huge economic losses.

Through the analysis of these accidents, it can be found that these crash effects are often caused by the overflow of load to the surrounding nodes when one or part of the nodes in the network cannot function because of overload. This redistribution mechanism causes new failed nodes to be created until no new nodes fail, a process called cascading failure [3], or sometimes "avalanche". The same problem exists in the Internet [4], communication networks [5], transportation and logistics networks, and other networks of socioeconomic systems [6]. We simulate this phenomenon, through a variety of indicators and data of the network, find the potential impact on the network of huge nodes to protect, you can enhance the robustness and controllability of the attack, a good way to avoid these catastrophic consequences. And when the crash occurs, measures are drawn to reduce the scale of cascading failure and reduce the loss.

The importance of nodes in the network is divided by cybernetics. According to the importance of nodes, the influence of vulnerable nodes can be expanded. The existing methods focus on one aspect of the influence of structure in network cybernetics and the influence of node attributes in cascading failures. There is a lack of fusion research on cascading failure that combines the controllability of network structure with network specific attributes. In a comprehensive view, the existing studies generally study the vulnerability and controllability of the network separately, or use controllability for evaluation and follow-up study after model simulation by different attack methods. There is not much discussion on weighted networks, which lacks the application of controllability theory in real power networks. This leads to the fact that when the influence of structural controllability is studied independently, the network does not have special characteristics and the cascade failure is poor. when the node properties are studied independently, the macroscopic network is lacked and the characteristics of the network structure are ignored to achieve the optimal cascade failure.Since the actual cascading failure is a problem that comprehensively considers structure collapse and attribute degradation, it is necessary to propose a comprehensive consideration method to find the most vulnerable node set of structure and attribute synthesis. This helps to improve the robustness of the network and better analyze the direction of defending against attacks.

To address the above issues, this paper explores the relationship between node cascade failure capability and its controllability from the perspective of network controllability, and proposes a node vulnerability ranking index for network controllability derived from structural controllability. The number and set of driver nodes in a network can explain the degree of controllability of this network. According to this principle, the impact on the network driver nodes when each link is changed can be approximated as the structural influence of this link: removing links, more driver nodes, less controllable network, and more vulnerable network; and vice versa more robust. Since this paper analyzes the cascading failure of node attacks, the value of the link property within n hops of each node can be used as the structural importance of this node. Then combined with the node collapse value to describe the node vulnerability, a number of points with higher attribute ranking are selected for simulation, and it is concluded that the influence of the network node decreases with the collapse value and link attribute value.

Based on the above research content, this paper proposes an integrated vulnerability algorithm, a method for assessing the importance of network nodes based on driving nodes and collapse values. We summarize three contributions of this paper.

- We use critical edge theory to evaluate the importance of nodes by considering critical edges, redundant edges and common edges as well as the tendency of nodes to near collapse. It makes the local neighborhood information of nodes effective for large-scale networks
- The algorithm effectively transformed the controllability of network edge structure in cybernetics and the cascading impact of links in cascading failure at the theoretical level. And link criticality is used as a feature at the structural level. Combined with the attribute feature of node collapse value, the invulnerability performance of cascaded network is described.
- The empirical analysis of real power grids shows that the performance of the method in this paper is better than that of the degree centrality algorithm, current betweenness centrality algorithm, WL algorithm and CI algorithm in the cascading failure process of power grids of all sizes.

The IEEE standard test system with 30, 118, 300 nodes is simulated, and compared with some traditional and classical methods, the combined application of the two aspects of vulnerability and structural controllability in power network cascading failure is analyzed. The specific guiding significance of the driver node

to the network cascading failure is studied, and the correctness and efficiency of the conclusion are proved.

2. Related Work

The main direction of current research for cascade failure is to use the identification of vulnerable nodes and then implement pre-protection as a way to stop the explosive behavior of cascade failure in the whole network. Schneider C M et al [7] proposed a systematic strategy to select the minimum number of autonomous nodes based on the measurement of the network mesonumber. This strategy requires five times less autonomous nodes compared to random selection. Faramondi et al [8] proposed a multi-objective problem by simultaneously minimizing the number of attacking nodes and the connectivity in the remaining network after the attack (Pairwisc Connectivity, PWC). Two existing multi-objective evolutionary algorithms, NSGA-Il [9] and MIDACO [10], are then used to find the non-dominated solution. It can be seen that either using cascade failure or solving the collapse problem requires an analysis of the nodes and structure, which is called network performance assessment. This assessment is mainly divided into two levels: attributes (node vulnerability) and structure (node controllability), and the results of the assessment can screen key nodes according to different characteristics and conditions.

The vulnerability and controllability of the network is an important part of the evaluation of the network performance, for example, based on the study of traditional structural controllability by Liu [11] et al. Liu et al. Combined the theory of structural controllability with the theory of complex networks, which can control the whole network system by controlling a minimum number of nodes,which are called drive nodes. The symmetry of the basic matching problem is linked to the linear correlation, which solves the structural controllability problem of large complex networks at the structural level. As another example: Huang et al [12] established a fully dependent network with adjustable cluster structure and found through the study that the cluster structure greatly increases the vulnerability of the network. M Shao, S et al [13] also investigated the effect of clustering properties on partially dependent networks, based on percolation theory, and analyzed in depth the property characteristics of clustered networks with partial dependencies.

For the research on the theoretical framework of network controllability, Liu et al. introduced the concept of structural controllability proposed by Lin Ching-Tai [14], and pointed out that the structural controllability can be considered by

ignoring the weights in the absence of edge weight information. At the same time, it is proved that the minimum number of driver nodes needed to control the whole network depends on the maximum matching in the network, which greatly reduces the time complexity of solving the problem. literature [15] also analyzes the controllability of two self-similar networks, and the results show that for both networks the set of driving nodes and controllability of the network do not change regardless of whether the edges have weights, which provides some help for the evolution of the unweighted study to the weighted one.

In the study of network controllability optimization, Wang [16] et al. first proposed that only one driver node can control the whole network by disturbing the network structure. By finding the maximum matching path in the network and connecting it to the first place by adding edges, the purpose of ensuring the control of a single driver node can be achieved. Since this method requires increasing the number of edges in the network, which greatly increases the real network control. the literature [17] proposes to further optimize the controllability of the network by reconnecting edges, based on the classification of key edges of the network, deleting redundant edges that do not affect the number and selection of drive nodes and adding other edges to improve the controllability of the network.

In the study of robustness related to cascading failures, literature [18; 19] applied seepage theory to the robustness study of smart grid. From the perspective of seepage theory and generating function, the cascading fault seepage process between the information network and the physical power grid and the critical condition of the blackout accident are studied. In the literature [20], a heuristic algorithm based on the local information of the network topology was designed to identify the importance of nodes in an undirected unweighted network with symmetric adjacency matrix using the properties of structural holes, and the correlation degree of nodes was linked to the number of structural holes. It is important for studying the survival and robustness of the network.

For the network modeling method of power system, Motter and Lai [21] proposed to construct a dynamic cascade model considering interference according to the topological relationship of power network. On the basis of this model, Crucitti et al. [22] proposed to use the capacity coefficient as the describing characteristics of nodes, and thus constructed a more accurate fault model to describe the power grid. However, these topological relationships are unweighted and undirected graphs, and their models cannot accurately describe the real power grid. Bompard et al [23] explored the limitations of using complex network

knowledge to analyze the topology of power grids because such analysis tools ignore the actual electrical laws of power grids, and thus the analysis results may be very different from the actual network. Since then, a large number of researchers have started to add actual electrical quantities to the grid topology for further improvement [24; 25].

3. Cybernetics Theory

The concept of complex network controllability was first introduced by R.E. Kalman [26] in 1960. For a complex network system or a linear time-invariant system, the

$$\dot{x} = Ax + Bu \tag{1}$$

where A and B are respectively $N \times N$ and $N \times N_D$ constant coefficient matrices, representing the adjacency matrix of the network and the nodes controlled by the external controller of the LTI system, respectively, x is the state vector and u is the control input, the Kalman criterion suggests that a sufficient necessary condition for this system to be controllable is the controllability matrix of the system,

$$C = [B \quad AB \quad A^2B \quad \cdots \quad A^{N-1}B] \tag{2}$$

satisfying row full rank, the canonicality is called the state canonicality of a complex network. The concept of structural controllability is a generalization of state controllability. For constant parameter coefficient matrices A and B, a system is said to be structurally controllable if there exists a set of nonzero parameter values that ensure that the system is state-controllable. For an energetically controllable system, its state vector x can be controlled by an appropriate control input u ($N \times 1$ vector), driven from any initial state to any target state.

The following are some basically relevant definitions and explanations.

Definition 1 (Degree) The degree of a node in a network is defined as the number of neighboring edges of that node. For a directed network, the degree of a node is divided into two types: in degree and out degree, where the in degree of a node is defined as the number of neighboring edges pointing to the node from other nodes, and the out degree of a node is defined as the number of neighboring edges pointing to the node from the node. The degree k of a node in the network The degree k of a node in a network can be calculated by the following equation:

$$k_i = k_i(in) + k_i(out) \tag{3}$$

$$k_i(in) = \sum_{j=1}^{N} a_{ji} \tag{4}$$

$$k_i(out) = \sum_{j=1}^{N} a_{ij} \tag{5}$$

a_{ij} The degree of a node is a measure of the importance of the node in the network, and the greater the degree, the greater the role of the node in the network dynamics.

Definition 2 (Average Degree) The average degree $< k >$ of a network is defined as the arithmetic mean of the degrees of all nodes in the network, i.e.:

$$< k >= \frac{1}{N} \sum_{i=1}^{N} k_i = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} \tag{6}$$

The average degree reflects the density of edges in the network, and a network with a small average degree is usually called a sparse network; conversely, a network with a large average degree is called a dense network.

Definition 3 (Controllability) Controllability of a network n_D is defined as the ratio of the number of its driving nodes N_D to the ratio of the network size N, i.e.

$$n_D = \frac{N_D}{N} \tag{7}$$

Based on this definition, the controllability of the network can be measured by the number of drive nodes.

Definition 4 (Minimum Input Theorem) The minimum set of inputs to a network (N_I) is equivalent to the minimum set of driving nodes (N_D). If the network is perfectly matched, the minimum input set is any node in the network; otherwise, it is equal to the set of nodes that are not matched after the maximum matching of the network. It can be expressed by the following equation.

$$|N_I| = |N_D| = max1, N - |M^*| \tag{8}$$

$N - |M^*|$ is the maximum number of matching nodes in the directed network. Based on the minimum input theorem, we used Hopcroft–Karp Maximum matching algorithm [27] to calculate the number of drive nodes as well as to count critical, redundant and common edges.

Definition 5 (critical edge) Based on the definition of control robustness by Liu and Barabasi et al. and in order to understand how well the network can be

controlled under unavailable link failures, we classify each link into one of the following three categories (Figure 1a, b, c): "critical" if in the absence of a link we need to increase the redundant" if it can be removed without affecting the current set of driver nodes; and "normal" if it is neither critical nor redundant.

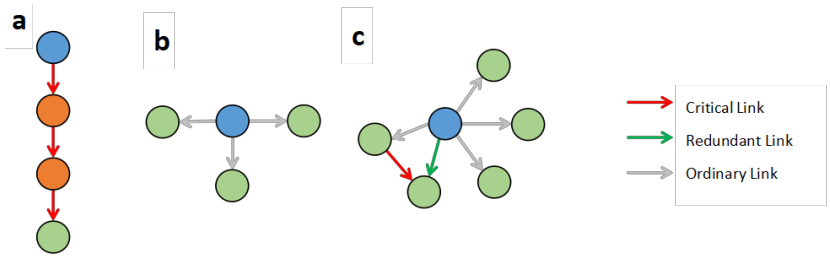

Figure 1: Network Link Classification

4. Cascade model for node failure

4.1. Node Status

The node state is represented by the ability of the available nodes to accommodate and transfer the load, which contains the initial load, load capacity, and load remaining accommodation capacity of the node.

First consider the remaining capacity of the node, and assume that the real-time load of node i is F_r and the capacity of the node is C_i then the residual capacity of the node's load L_i, can be expressed by the following equation.

$$L_i = \begin{pmatrix} C_i - F_r, & F_r \leq C_i \\ 0, & F_r > C_i \end{pmatrix} \tag{9}$$

That is, the residual capacity of node load is the difference between the node capacity and the existing load, and when the load exceeds the capacity, the residual capacity of the load is 0, and the node has failed at that time. The remaining capacity of node load reflects the ability of node to handle or accommodate the load. The larger the value is, the stronger the processing capacity is and the higher the percentage should be when the load is redistributed, and the opposite is the lower. It is constantly changing with the change of node load.

A fairly intuitive example of reflecting the state of a node is in a traffic network, where a road is blocked for unexpected reasons, and the traffic on the road will

shift to other roads adjacent to it, and the traffic management center will suggest drivers to move to roads that can carry more vehicles based on real-time monitoring data, but the carrying capacity of the road is affected by the diversion capacity of the adjacent roads.

The initial load of a node is the initial value of the node load in the load network, and the initial load of node i is assumed to be F_i The classical equation of node initial load is used, which is defined as a function of the degree of the node k_i as a function of

$$F_i = \rho k_i^{\tau} \tag{10}$$

Among them , ρ and τ are adjustable parameters, which control the magnitude of the initial load of node i. This dimensionless "structural load" is a reasonable and effective way to study the destructiveness of complex networks and to evaluate the importance of nodes in cases where it is difficult to determine the actual physical load on the network.

In a real network, the capacity of each node to handle the load is usually affected by technical and economic factors. The load capacity of a node is determined on a "demand-driven" basis, so the load capacity of a node is considered to be C_i is proportional to its initial load.

$$C_i = (1+\alpha)F_i \tag{11}$$

where is the capacity factor of the network, which indicates the capacity of the node to handle the additional load.

4.2. Load redistribution method

The load redistribution method is a rule that describes the transfer of load on a node to other nodes in the network after its failure. The load redistribution method is shown in Fig. 2. After the failure of node f, the load (e.g., information flow, current, traffic flow) originally intended to pass through this node f will choose a new path and the load on it will be transferred to the node adjacent to it to ensure the efficient operation of the whole network. The whole network will undergo a full update of the load due to the failure of the node, i.e.

$$F_i' = F_i + \Delta F_i, \qquad i \in \Gamma_f \tag{12}$$

where F_i is the load before the update, and F_i' is the load after the update, and ΔF_i is the increment at update, and Γ_f is the set of all neighboring nodes of the failed node

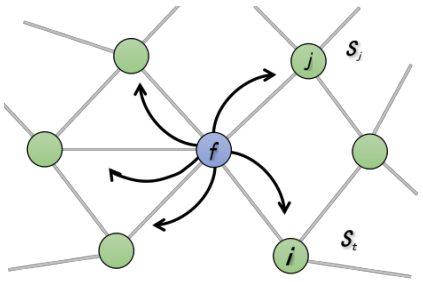

Figure 2: Load reallocation based on nodes' state evolution

f (and the failed nodes are not counted in the set of neighboring nodes). Suppose the real-time load of the failed node f is F_r , the load increment redistributed on the neighboring nodes can be expressed by the following equation.

$$\Delta F_i = \frac{F_r}{\Gamma_f} \tag{13}$$

In this paper, we use the nearest neighbor uniform shared load rule, which is a simplification of the nearest neighbor redistribution approach, and it is related to the node state but not to the node load.

5. Integrated vulnerability attack algorithm

The combined vulnerability metric used in this paper $At(i)$ is a combination of the node's nearby link criticality and the node collapse value, if only the resultant controllability is considered, the node has propagation only among the unprivileged network and cannot measure its influence in the privileged network or say that its influence in the privileged network is not the most considerable; if only the load and weight properties of the node are considered, it will strengthen the influence of a single or small range of nodes and cannot spread this influence. Therefore, this paper associates to consider the structural scopicity of cybernetics together with the important attributes of nodes in the entitled network, and the two values are normalized and defined as follows.

$$At(i) = \beta \times Cr(i) + \gamma \times Co(i) \tag{14}$$

where the nearby link criticality $Cr(i)$ for structural fragility, and the collapse value $Co(i)$ for state fragility, the β and γ are normalization operations performed on the

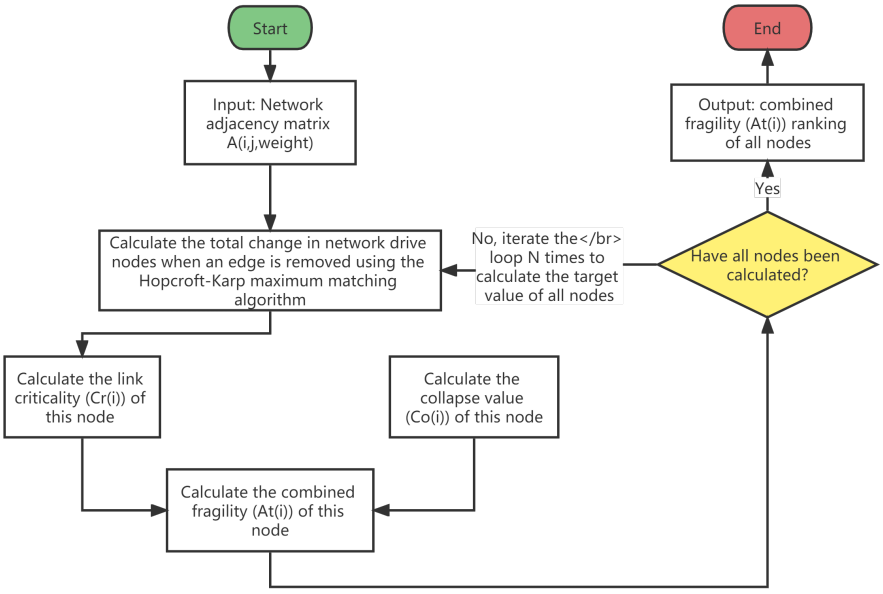

Figure 3: Integrated vulnerability algorithm flow chart

two values so that both have the same magnitude. To evaluate the degree of link impact on the structure, the nearby link criticality is specified as the sum of the link criticalities within n hops of a node. In the attempts on this experimental network, the results of the various methods obtained by choosing n = 2 are more promising, and it is easy to distinguish the differences between the various algorithms without the need for more global information.

The link criticality of node i is defined as follows.

$$Cr(i) = \frac{\sum_{i \in \phi(n)} \Delta D_i + c}{|\phi(n)|} \tag{15}$$

where $\phi(n)$ denotes the total set of all links within n hops, the ΔD_i denotes the amount of drive node variation in the total network after removing link i, i.e., the link criticality, and c is the smoothing value.

According to the definition of controllability, the ratio of drive nodes that describes the strength of network controllability, the removal of a link, the increase of drive nodes, the weakening of network controllability, proves that the changed link plays a positive role in the stability of the whole network and is a critical link,

the impact of destroying such edge on the network will be huge; on the contrary, if the number of drive nodes does not change or only the set of drive nodes changes, the network controllability hardly changes, proving that the changed links are dispensable to the network, and the network can still be controlled without them, and breaking such edges will not cause a large impact. This impact is mainly from the characteristics of the structure and applies to so directed networks, but for some networks where the nodes have load weight distinctions, to find more critical nodes more effectively, it is necessary to combine the nodes' own properties for the fragility of the node states and define the collapse value as follows Eq.

$$Co(i) = \frac{\sum_{i \in \Gamma_f} \frac{\Delta_i}{L_i}}{|\Gamma_f|} \tag{16}$$

where Γ_f denotes the set of neighboring nodes directly affected after attacking node f, and Δ_i denotes the amount of load change of node i after attacking node f, and L_i denotes the difference between the capacity and load of node i (remaining accommodation capacity).

If the overflowing load of node f is greater than the remaining holding capacity of node i, that is, we get $\frac{\Delta_i}{L_i} > 1$, at this time the node will overload and fail, if the load overflowed by node f is less than or equal to the remaining capacity of node i $\frac{\Delta_i}{L_i} < 1$, the node normally accommodates the load at this time, so that the partial cascade failure stops. The definition of collapse value is a measure of the average accommodation capacity of nodes around a node, the larger the Co(i), the worse the average accommodation capacity, the easier it is for such neighboring nodes to collapse and fail, i.e., the impact of this node is relatively large from a small structure of one hop range, and vice versa, the impact is smaller.

According to the integrated fragility of the ranking results, the nodes can be divided into critical and non-critical nodes, and the removal of critical nodes can have a great impact on the stability of the network.

6. Experiments and analysis of results

6.1. Data Description

To evaluate the performance of the proposed method, we apply it to real networks. The real network includes IEEE30-bus, IEEE118-bus and IEEE300-bus node system [28][29]. The network statistical characteristics of the three networks

Network	N	M	<k>	L	C
IEEE30-bus	30	41	2.733	3.306	0.235
IEEE118-bus	118	186	3.153	6.309	0.165
IEEE300-bus	300	373	2.487	9.651	0.095

Figure 4: Table 1 Basic statistical characteristics of three real networks, including network size (N), number of network edges (M), node average degree (<k>), average shortest path length L, and clustering coefficient C.

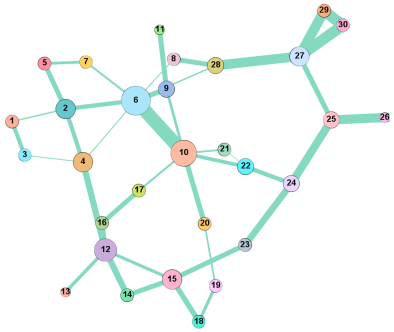

Figure 5: IEEE 30 Node Network Topology Schematic

are shown in the table. As can be seen from the table, each of the three real data sets used in this paper has its own characteristics and is representative, which can verify the effectiveness of the algorithm well. The node size is positively correlated with the node degree, i.e., the larger the node degree, the larger the node, and vice versa; the link width is positively correlated with the load flowing through the re-link, i.e., the larger the load of the link, the wider the width, and vice versa the narrower the width. The numbers marked in the nodes indicate the serial numbers of the nodes, and the arrangement is YiFan Hu.

6.2. Benchmark Methods

We use several classical and popular heuristics to compare with this paper, which are also studied entirely based on the network topology. These methods include: the degree centrality algorithm, the current betweenness centrality algorithm, the WL algorithm, and the CI algorithm.

1. Degree centrality algorithm

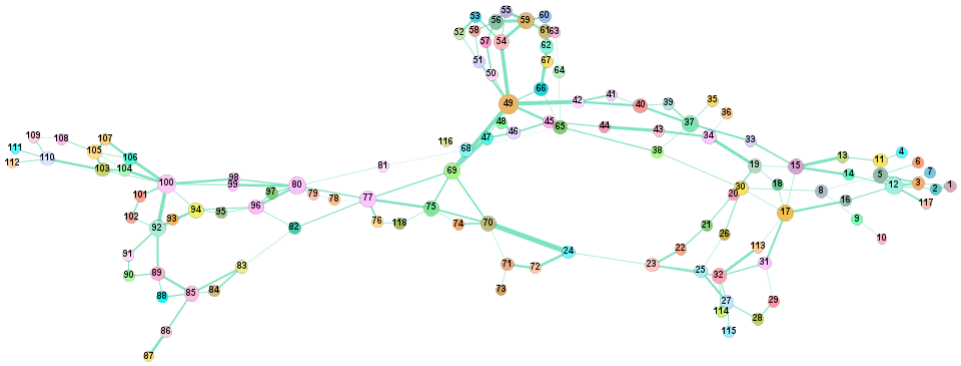

Figure 6: IEEE 118 Node Network Topology Schematic

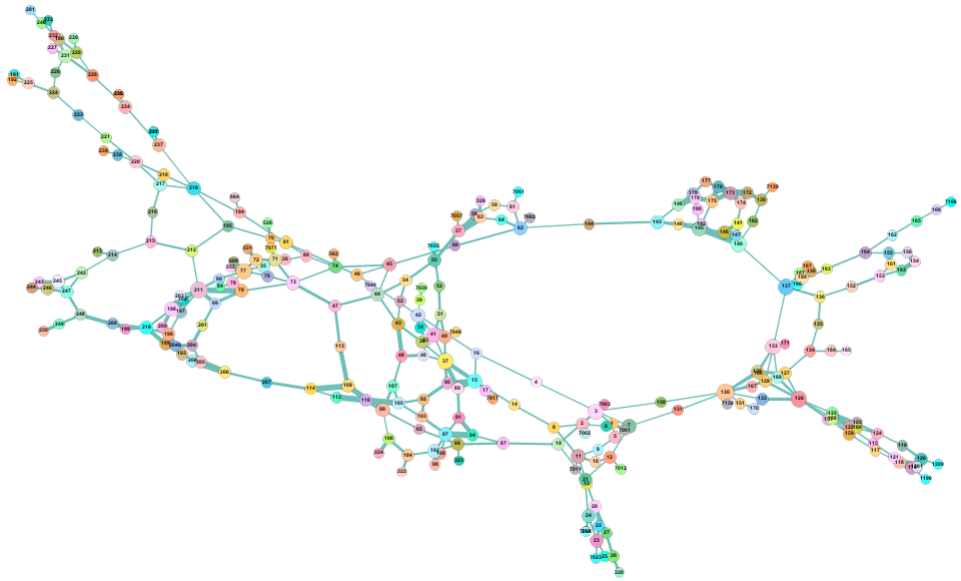

Figure 7: IEEE 300 Node Network Topology Schematic

Linden C. Freeman [30], a research professor in the Department of Sociology and the Institute of Mathematical and Behavioral Sciences at the University of California, Irvine, USA, formally introduced the concept of degree centrality in 1979.

Degree centrality is the most direct and basic metric to characterize nodes The most direct and basic metric of centrality, he can measure the importance of a node in the network according to the number of degrees of the node, a brief description is that the greater the node degree of a node means the higher the degree centrality of the node, the more important the node is in the network, the degree of node i can be expressed as

$$k_i = \sum_{j=1}^{N} a_{ij} \tag{17}$$

where a_{ij} denotes the relationship between node i and the other N-1 j nodes that are directly connected.

2. Current Median Centrality Algorithm

The current betweenness centrality algorithm [31] was proposed by Brandes and Fleischer in 2005 and can be used to estimate the flow state in a spatial network and is commonly used to measure the mediation of nodes in a network. The current mediator centrality of vertex i is defined as the amount of electrical flow flow through i in this network. The current permittivity of vertex i is the average of the currents of all node pairs (s source-target pairs).

$$C_{CB}(i) = \frac{\sum_{s \neq t \in i} I_i^{st}}{\frac{1}{2}n(n-1)} \tag{18}$$

where n(n - 1)/2 is a normalization constant and I_i^{st} is the current through vertex i between s and t. Thus, the current betweenness centrality measures the ratio of currents through vertex i between all possible pairs of nodes in the network.

The main drawback of this centrality is its high computational cost, especially for very large networks. But at the same time its computational cost drawback makes it suitable for networks with a large number of nodes, and is introduced here as a large number of network measurement methods to compare with the method in this paper.

3.WL algorithm

The Weisfeiler-Lehman algorithm [32], a method for testing isomorphic graphs.The main idea of the WL algorithm is that the importance of the nodes in the network

and the importance of the edges connected to the nodes are closely related. The weights of the edges ij can then be expressed as

$$\omega_{ij} = k_i \times k_j \tag{19}$$

k_i is the degree value of node i. The weights of the nodes can be expressed as

$$\omega_i = \sum_{j \in \Gamma_i} \omega_{ij} \tag{20}$$

Γ_i is the set of neighbors of node i. Thus the importance of nodes can be expressed as

$$\omega(i) = \frac{\omega_i}{\sum_{j \in N} \omega_j} \tag{21}$$

4.CI algorithm

The CI algorithm identifies the critical nodes that are most effective in disrupting network connectivity after deletion, as indicated by.

$$CI_i = (k_i - 1) \sum_{j \in \varphi ball(i,l)} k_j - 1 \tag{22}$$

where $\varphi ball(i,l)$ is the sphere with node i as the center of the circle and the radius the boundary of the sphere consisting of all nodes up to distance i from of all nodes. In this paper, the CI radius is set to 2.

6.3. Evaluation Criterion of Algorithms

We used lapse rate, network residual load, and network efficiency to evaluate the severity of the network under attack in order to assess the effectiveness of the node importance identification algorithm.

1. Failure of efficiency

Failure rate is the most direct parameter to measure the proportion of remaining effective nodes in the network and the degree of network failure. The change of failure rate during dynamic cascade failure directly reflects the number of failed nodes in each round of the network, and the failure rate is the ratio of the number of deleted nodes to the total number of nodes, i.e.

$$\eta_1 = \frac{N_d}{N} \tag{23}$$

N is the total number of nodes, and N_d denotes the number of remaining effective nodes, and the faster the failure rate decreases, the more nodes are crashed.

2. Network load

After load redistribution, the proportion of the total load of the remaining nodes in the network after load redistribution to the total initial network load, i.e.

$$\eta_2 = \frac{\sum_{i=1}^{M} F_r'}{\sum_{i=1}^{N} F_r} \tag{24}$$

The faster the network load drops, the less load flows through the network links and nodes, the less information the network delivers, and the more pronounced the network failure.

3. Maximum connectivity factor

The maximum connectivity coefficient G [33] is an important measure of network resistance to destruction and can be calculated by:

$$G = \frac{R}{N} \tag{25}$$

Where R denotes the maximum number of nodes connected to the network after the attack and N denotes the total number of nodes in the network. the faster the value of G decreases, the more effective the attack strategy is.

6.4. Analysis of results

In order to verify the effectiveness of the comprehensive vulnerability index to measure the importance of nodes, we compare the proposed method with the degree centrality algorithm, the current betweenness centrality algorithm, the WL algorithm, and the CI algorithm. The effect of intentional attacks on real networks is simulated by selectively removing nodes in order of their importance.

The number of attack nodes for the three real networks are 3, 10, and 15, and the node removal rates are 0.1, 0.085, and 0.05, respectively. In the case of small node removal rate or number of attack nodes, the advantage of the integrated vulnerability approach is obvious.

From the figure, it can be seen that the failure rate of the integrated vulnerability algorithm increases the fastest and always outperforms the other algorithms. In addition to this it can be noticed that the WL algorithm basically maintains the attack effect second only to the algorithm in this paper.

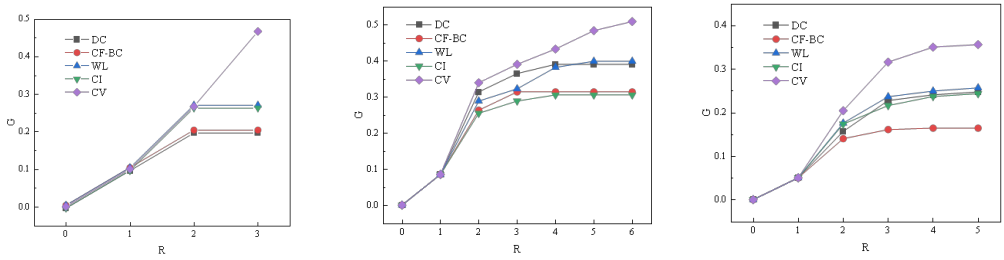

Figure 8: Failure rate varies with the number of rounds
The y-axis represents the failure rate of the network The x-axis represents the rounds of cascaded failure load redistribution R. (a) IEEE-30bus (b) IEEE-118bus (c) IEEE-300bus. all algorithms include: degree centrality algorithm (DC), WL algorithm based on nodes and their neighbor degrees (WL), current mediator centrality algorithm (CF-BC), CI algorithm based on percolation principle (CI) , integrated vulnerability algorithm (CV). Based on the variation of the failure rate, it can be seen that in the three real networks, the integrated vulnerability algorithm damages the network tremendously, achieves considerable results with few nodes, and always fails faster than the other algorithms.

To make the results easier to compare, methods that do not stop crashing before the last round delay the final state until the last round for comparison.

In the failure rate change results, in the small-scale network, the network changes in the first two rounds are almost the same. However, the CV algorithm has more rounds of reassignment than the other algorithms, which leads to a significantly higher failure rate of the results than the other algorithms. It proves that the vulnerable nodes selected by the algorithm have more destructive power and influence. In the larger network, the algorithm is in the leading position in each round of failure. The reason is that the link criticality based on cybernetics greatly improves the collapse propagation, and the collapse value based on the adjacent collapse attribute greatly improves the failure degree of the selected nodes and the spreading nodes. CF-BC algorithm is more suitable for extremely large networks, so it does not work well in general networks. Compared with the CV algorithm, the other algorithms have little difference. In general, WL algorithm is second only to CV algorithm.

In the network load change results, the impact on the network load is also basically synchronized with the failure rate. It is proved that the failure nodes hit by CV algorithm are also concentrated in the nodes with heavy load and great influence on the network. And the final node load failure degree is much larger than other algorithms, basically realizing the loss of half of the network load. And

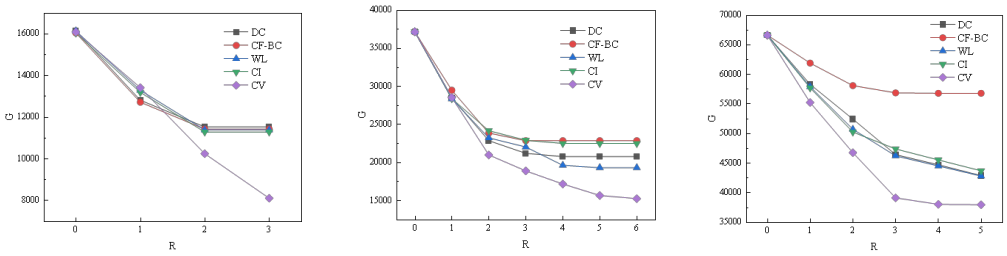

Figure 9: Network load varies with number of rounds
The y-axis represents the network load , and the x-axis indicates the rounds of cascading failed load redistribution R. (a) IEEE-30bus (b) IEEE-118bus (c) IEEE-300bus. it can be seen that the CV algorithm has the greatest impact on the network load, making a large number of loads in the network fail, which greatly affects the amount of information and functionality of the network, and the load disappears faster than other algorithms.

it has a good effect in the small and large network world.

Looking at the three networks, the CV algorithm does not directly break the connectivity significantly at the beginning. Instead, a large area strike is applied to other critical nodes, which in the final stage will destroy the cluster of nodes that directly affect the connectivity. Finally, the maximum disruption of connectivity is reached. If the key nodes affecting connectivity are disabled too much at the beginning, it will lead to the inability to further expand the impact and make the damage propagate fastest.

After some experimental attempts, it is found that the algorithm only requires neighborhood information within two hops of a node to obtain respectable results in a general-sized network, and no global information is required. And the more large-scale network attack trials, the better the performance of the integrated vulnerability algorithm can be. And the nodes selected for attack are all between 0.05-0.1 of the total number of nodes, and the method in this paper has more significant advantages in practical applications where the cost of network attacks is limited.

7. Summary and Outlook

This paper discusses the impact of network controllability on network structure vulnerability and its specific application in power networks. Aiming at the analysis of comprehensive vulnerability, the critical links that affect the network controllability are selected as the evaluation objects of link structure vulnerability. A more

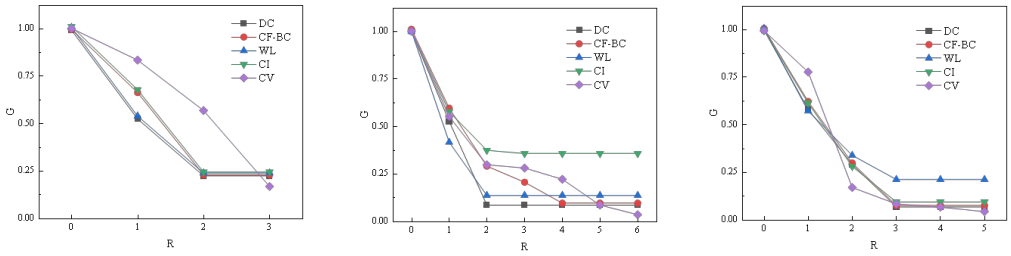

Figure 10: Variation of network connectivity coefficient with number of rounds
The y-axis represents the maximum connectivity factor G of the network and the x-axis represents the number of rounds of cascading failure load redistribution R. (a) IEEE-30bus (b) IEEE-118bus (c) IEEE-300bus. The impact of the CV algorithm on network connectivity is greater than the other algorithms, and the algorithm for connectivity predicts that although the process and results do not differ significantly in value, the overall connectivity the impact on the overall connectivity is significant.

comprehensive vulnerability index is established by combining the attribute vulnerability of node collapse tendency degree, and the vulnerable node set is constructed. Then, simulations were carried out with IEEE30, IEEE118 and IEEE300 power systems as examples. And various comparisons were made with other classical importance ranking methods to verify that the integrated vulnerability is more effective in identifying vulnerable nodes.

Aiming at the application of controllability analysis, this paper makes some attempts and experiments to apply the node control ability derived from network controllability to vulnerability analysis. A new idea of vulnerability analysis is proposed, which is to carry out attacks according to the order of control ability of nodes. It is finally concluded that the control capability possessed by a critical link is large and linearly related to its own vulnerability, and the more critical links in the corresponding nodes and neighboring connected edges, the greater the role played in the topology and the greater the impact after failure. This method of node importance identification has important theoretical and practical significance for improving the robustness of the network or finding out the vulnerable nodes of the network.

References

[1] M. Sforna and M. Delfanti, “Overview of the events and causes of the 2003 italian blackout,” in *Power Systems Conference and Exposition, 2006. PSCE*

'06. 2006 IEEE PES, 2006.

[2] U. T. Force, "Interim report: Causes of the august 14th blackout in the us and canada," 2003.

[3] L. Zhao, K. Park, and Y. C. Lai, "Attack vulnerability of scale-free networks due to cascading breakdown," *Physical Review E*, 2004.

[4] J. W. W Peng, "Pastor-sat orras r, vazquez a, ves pi gnani a. dynami cal and correl ation p roperti es of the in tern," *Advances in Applied Mathematics*, 2001.

[5] R. Guimera, A. Arenas, A. Diaz-Guilera, and F. Giralt, "Dynamical properties of model communication networks - art. no. 026704," *Physical review.E.Statistical physics, plasmas, fluids, and related interdisciplinary topics*, no. 2 Pt.2, p. 66, 2002.

[6] D. J. Watts, "A simple model of global cascades on random networks," 2011.

[7] C. M. Schneider, T. A. Kesselring, J. S. Andrade, and H. J. Herrmann, "box-covering algorithm for fractal dimension of complex networks a box-covering algorithm for fractal dimension of complex networks," 2018.

[8] M. Schlüter, M. Gerdts, and J.-J. Rückmann, "A numerical study of midaco on 100 minlp benchmarks," *Optimization*, vol. 61, no. 7, pp. 873–900, 2012.

[9] Y. Y. Liu, Slotine, Jean-Jacques, Barabási, and Albert-László, "Controllability of complex networks." *Nature*, 2011.

[10] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii, ieee trans. on evol," *IEEE Transactions on Evolutionary Computation*, vol. 6, 2002.

[11] L. Faramondi, G. Oliva, S. Panzieri, F. Pascucci, M. Schlueter, M. Munetomo, and R. Setola, "Network structural vulnerability: A multiobjective attacker perspective," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 49, no. 10, pp. 2036–2049, 2019. [Online]. Available: https://doi.org/10.1109/TSMC.2018.2790438

[12] X. Huang, S. Shao, H. Wang, S. V. Buldyrev, H. E. Stanley, and S. Havlin, "The robustness of interdependent clustered networks," *Epl*, vol. 101, no. 1, p. 18002, 2013.

[13] S. Shao, X. Huang, H. E. Stanley, and S. Havlin, "Robustness of partially interdependent network formed of clustered networks," 2013.

[14] C.-T. Lin, "Structural controllability," *IEEE Transactions on Automatic Control*, vol. 19, no. 3, pp. 201–208, 1974.

[15] Deng, Cong-Zheng, Xu, Chuan-Yun, Ming, Wang, Huan, Cao, and Ke-Fei, "Analytical controllability of deterministic scale-free networks and cayley trees," *The European physical journal, B. Condensed matter physics*, 2015.

[16] W. X. Wang, X. Ni, Y. C. Lai, and C. Grebogi, "Optimizing controllability of complex networks by minimum structural perturbations," *Physical Review E Statistical Nonlinear & Soft Matter Physics*, vol. 85, no. 2 Pt 2, p. 026115, 2012.

[17] L. Hou, S. Lao, B. Jiang, and B. Liang, "Enhancing complex network controllability by rewiring links," in *Third International Conference on Intelligent System Design & Engineering Applications*, 2013.

[18] H. Zhen, W. Cheng, S. Ruj, M. Stojmenovic, and A. Nayak, "Modeling cascading failures in smart power grid using interdependent complex networks and percolation theory," in *2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA)*, 2013.

[19] L. I. Wenguo, S. Deng, L. I. Jiasheng, and W. Xiao, "Defense strategy of cascading failures between information network and physical power grid," *High Voltage Engineering*, vol. 39, no. 11, pp. 2714–2720, 2013.

[20] H. Yang and S. An, "Critical nodes identification in complex networks," *Symmetry*, vol. 12, no. 1, pp. 123–, 2020.

[21] "Proceedings. first international conference on the quantitative evaluation of systems," in *International Conference on Quantitative Evaluation of Systems*, 2004.

[22] A. E. Motter and Y. C. Lai, "Cascade-based attacks on complex networks," 2003.

[23] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Physical Review E Statistical Nonlinear & Soft Matter Physics*, vol. 69, no. 4 Pt 2, p. 045104, 2003.

[24] E. Bompard, W. Di, and X. Fei, "Structural vulnerability of power systems: A topological approach," *Electric Power Systems Research*, vol. 81, no. 7, pp. 1334–1340, 2011.

[25] H. Feng, L. I. Huaqiang, Y. Wang, and Y. Luo, "Comprehensive vulnerability assessment method for nodes considering anti-interference ability and influence," *Proceedings of the CSU-EPSA*, 2017.

[26] J. I. Xingpei, W. Bo, D. Zhaoyang, C. Guo, L. Dichen, W. Daqian, and W. Xunting, "Vulnerability assessment and edge protection strategies for power information physical interdependence networks(inchinese)," 2016.

[27] J. E. Hopcroft and R. M. Karp, "Algorithm for maximum matching in bipartite graphs," *soc.ind.appl.math.j.comptation*, 1973.

[28] A. Abaza, A. Fawzy, R. A. El-Sehiemy, A. S. Alghamdi, and S. Kamel, "Sensitive reactive power dispatch solution accomplished with renewable energy allocation using an enhanced coyote optimization algorithm," *Ain Shams Engineering Journal*, vol. 12, no. 1, 2020.

[29] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1978.

[30] J. E. Hopcroft and R. M. Karp, "An $n^ 5/2 $ algorithm for maximum matchings in bipartite graphs," *SIAM Journal on Computing*, 1973.

[31] U. Brandes and D. Fleischer, "Centrality measures based on current flow," in *Proceedings of the 22nd annual conference on Theoretical Aspects of Computer Science*, 2005.

[32] B. Y. Weisfeiler and A. A. Leman, "A reduction of a graph to a canonical form and an algebra arising during this reduction (in russian)," 1968.

[33] H. Pavlović, "Graphs with maximum connectivity index," *Computational Biology and Chemistry*, 2003.