IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 18, 2018, accepted April 22, 2018, date of publication April 25, 2018, date of current version May 24, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2829979

Analysis of Cascade Fault Optimization Based on Regional Fault and Traffic Reallocation in Complex Networks

LANLAN RUI, (Member, IEEE), XIAOMEI WANG, AND XUESONG QIU, (Member, IEEE)
State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China
Corresponding author: Lanlan Rui (llrui@bupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61302078 and Grant61702048 and in part by the 863 Program under Grant 2011AA01A102.

ABSTRACT With the development of complex networks and demand for information services, more and more businesses are turning to complex networks for their hosting needs. Due to the growing service, attention must be paid to the problem of complex network cascade failures. A cascade failure is a large-scale failure caused by several small-scale failures. This phenomenon initially attracted attention in the power grid field. Currently, complex network cascade failures also deserve attention. In this paper, we focus on mitigating the impact of cascade failures. First, we propose a model to simulate the cascade fault propagation process based on a virus propagation model. We introduce an SDN control plane to the communication network to monitor the overall network. We consider regional failures and optimize traffic redistribution strategies to reduce the impact of cascade failure. Finally, we compare several simulation results based on proposed fault propagation models. The mechanism, we proposed, shows better performance in the remaining available nodes and the remaining network traffic.

INDEX TERMS Cascade failure, regional failure, traffic redistribution, virus propagation model, complex network.

I. INTRODUCTION

With the rapid development of the information society, the service and traffic carried by communication networks have also seen exponential growth. Modern communication networks have gradually evolved into a typical complex network structure. Fig.1 shows the schematic complex network structure. In the modern complex network, different devices and service are connected by network. Through routers/switches and links, we can communicate with other people and many different service that can make people's lives easier. With the increase in various forms of network, more and more researchers focus on the study of network structures in real world systems particularly from two perspectives: dynamical processes taking place on networks and how network structure impacts such dynamics [1]–[6]. Cascade failure is one of the issues that aroused interest in complex networks. The occurrence of cascade failure means that the network may not be able to provide services or may even be paralyzed. This is the result of a reallocation of important services after network failure. In communication networks, some important links and nodes have backup equipment. When important nodes and links are interrupted, their service can be transferred to backup equipment. Therefore, important service can be quickly recovered. Other nodes and links do not have backup equipment, so they need to redistribute services, which are sent to other normal working nodes and links. After redistribution, the nodes and links that have increased service needs may be out of capacity, which triggers additional interruptions and a new round of traffic redistribution. There are many different solutions to the problem from a variety of perspectives.

Research into cascading failures mainly focuses on two aspects: the cascading failure model and the cascading failure optimization mechanism. From a modelling point of view, methods based on branching process parameters [7] or statistical decision theory frameworks [8] can be used to analyse cascade failure. From the optimization perspective, Gupta *et al.* [8] identify potential critical links that can cause cascade failure and find ways to avoid the failure.

Although many related studies on cascade failure have been pursued, there are still many problems that need to be solved. For example, mitigating the impact of cascade failure

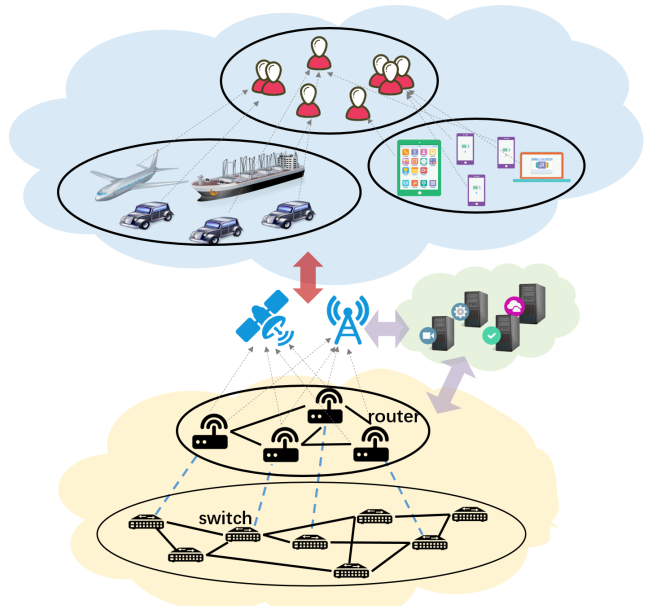

FIGURE 1. **The composition of complex communication networks.**

after an occurrence is an important task. In this paper, our contribution is as follows:

a) We propose a model that describes the cascade failure propagation process based on a virus propagation model. The virus propagation model has three types of nodes: susceptible nodes, infected nodes and recovery nodes. These nodes correspond to working nodes, failed nodes and repaired nodes, respectively, in a cascade failure. Moreover, the dissemination of cascade failure resembles the spread of a virus. The virus propagation model is similar to fault propagation in complex networks. Thus, we choose the virus propagation model to simulate cascade failure. Compared with other models, the virus propagation model has better performance in terms of simulating fault propagation. We verify this result in the simulation section.

b) Based on the cascade failure propagation model, we consider regional failures and optimize traffic reallocation mechanisms to ease the impact. We introduce the concept of survivability to measure the probability of link-avoidance regional failures. While considering the process of reallocation, we determine the appropriate nodes according to survivability parameters. Moreover, reallocation can ensure that the backup nodes and links we choose avoid being influenced by regional failures. The nodes which have higher survivability have higher reliability. In the process of traffic reallocation, we consider distributing the traffic according to clustering coefficients. If the nodes have a higher clustering coefficient and better connectivity, then the better connectivity guarantees that the node will have a better traffic distribution if the node fails. These two measures can effectively reduce the impact of cascade failure. Compared with other schemes, the scheme proposed here can also ensure service to some extent.

The remaining structure of this paper is as follows. In Section II, we discuss related work on cascade failure. In Section III, we model the network and propose a cascade fault propagation model. Then, we propose improved approaches for regional failure and traffic redistribution. In Section IV, we show the simulation result and analysis optimization. Finally, we conclude the paper and discuss future work in Section V.

II. RELATED WORK

In this section, we discuss related work on cascade failure models and optimization measures. Cascade failure originates in the power grid. Many studies about cascade failure are based on the power grid. Nevertheless, modern communication networks have gradually become complex networks, and cascade failure problems have emerged. Therefore, the study of cascade failure in communication networks is also important. First, we discuss the modelling of cascade failure.

Dey *et al.* [7] studied the basic topology of power networks and calculated the average propagation of faults under changing topology conditions as the branch processed parameters. They also studied the relationships among the network topological characteristics. Dey *et al.* [9] used the branch computing process as a model to propose the use of RT-Lab real-time cascade data generation and new packet technology to assess the branch process parameters. Gupta *et al.* [8] proposed a smart power grid probabilistic framework with statistical decision theory for assessing system performance in steady state and dynamic conditions and determining the potential for cascading. Sanghavi *et al.* [10] designed an effective algorithm to construct a Markov reliability system with cascading failures. In addition, they also derived two new dependability measures related to the distribution of the size of a cascade. Zhang *et al.* [11] use a circuit-based power flow model to study the cascade fault propagation process; they then combine it with a stochastic model to describe an uncertain moment of failure to produce a model that can be used to describe cascading failures in a circuit network. Yu [12] established model interactions between the virus propagation and the cascade failure. However, they only verified the simulated and theoretical values and did not provide any optimization measures. By combining the stochastic process with the state transition description, Liu *et al.* [13] proposed a method for describing the network coupling of the smart grid. La [14] studied the cascade failures in a system comprising interdependent networks/systems, with particular attention to the variability and dependence of node degrees. In [15], a novel interdependent Markov-chain framework was proposed, with the goal of predicting the resilience in response to cascade failure. Reference [16] proposed an approach to examine the vulnerability of a specific type of complex network against cascading failure threats that is based on an extended topological metric.

Much of the existing research concerns modelling cascade failures. However, there are still several methods to mitigate the impact of cascade failures. In [17], a cascade fault model is applied to the system. The system was subsequently analysed in terms of flow redistribution, the vulnerability and robustness of a circuit system with random faults, and the maximum load fault. Gupta *et al.* [8] determined ways to predict critical links, which may lead to system blackouts based on the cascade failure prediction model. La [14] studied the relationship between node degrees and cascade failure; they proposed a method of improving node degrees to mitigate the failure. Iwaori and Hayashi [18] proposed a new measure to analyse the effects of countermeasures against cascading failures of a telecommunications network. The method enables the analysis of cascading failures without ignoring important features such as node pairs. Liu *et al.* [19] proposed a submodular optimization approach for mitigating cascading failures by partitioning the system into internally stable islands (a process known as controlled islanding). In [20], the paper describes load-dependent cascading failures in random networks consisting of a large but finite number of components. Under a random single-node attack, a framework is developed to quantify the damage at each stage of the cascade. Estimations and analyses for the fraction of failed nodes are shown to evaluate the time-dependent system damage resulting from the attack. The results provide guidelines for choosing the load margin to avoid a cascade of failures. Xu *et al.* [21] proposed a method for power system cascade failures. The approach uses simulated trajectories of the system to determine uncertainties. In [22], a node importance indicator considered the load sharing rule is proposed to identify the critical nodes and mitigate cascade failure. Based on these aspects, we can propose our optimization algorithm, there are also some researches about optimization algorithm in [23] and [24]. Referring these references, we propose the optimization algorithm in cascade failure propagation.

In this paper, the improvement of the traffic reallocation process and the consideration of the regional fault are proposed based on the cascade fault propagation process. Finally, the optimization is verified based on the virus propagation model.

III. NETWORK MODELLING

A. PROBLEM DESCRIPTION

In modern communication, complex networks carry a variety of services; each node and link transmits the information used to provide services for the users. To guarantee the stability of the service provided by the network, service providers secure backup equipment for those important nodes. If these nodes fail, the traffic they carry can be led to the backup equipment so that network stability and user experience will not be affected. Because resource constraints can be costly, each node and each link in the network cannot be backed up. Although these links and nodes have a lower importance degree or higher security degree, there are still several important faults. To ensure that the network can provide services normally, the traffic that is originally destined to be transmitted by the faulty node or link needs to be reallocated to other nodes and links. Those links and nodes have carried these services before, as they are reallocated new services. In this case, those links and nodes may overload and breakdown. Thus, a new reallocation process will be initiated. The scope of the fault is expanded. This is called cascade failure.

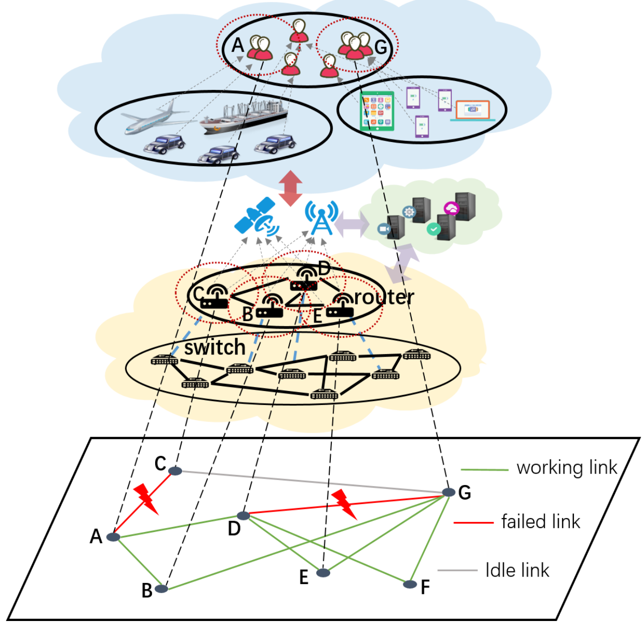

FIGURE 2. Cascade failure diagram.

Fig. 2 shows a schematic of a cascade. There are messages that A wants to send to G. At first, the messages are carried on the link A-C-G. However, Router C has failed for some reason. The link A-C is blocked and the link C-G is idle. Under this circumstance, to successfully deliver the message, we need to choose a different backup path. (The red link in the figure denotes a failed link, the grey link denotes an idle link, and the green link denotes a working link.) Therefore, the traffic is assigned to the Routers D and B. First, certain traffic is carried on link D-G. The link overloads due to the newly distributed traffic. The traffic needs to redistribute to other nodes and links. Thus, the traffic assigned to D is reassigned to Routers E and F. Finally, the message is delivered to destination G via path A-B, A-D-E-G and A-D-F-G. The process of spreading faults from link A-C to link D-G is called cascade failure.

To effectively cope with cascading failures that may occur in the network, we need to predict the network conditions after the fault occurs. When simulating the network cascade fault, a virus propagation model is used. Based on this model, we introduce some improvement measures, such as regional faults. In the reallocation process, after the cascade fault occurs, we choose the nodes that are furthest away from the fault nodes. This process ensures that the affected nodes will

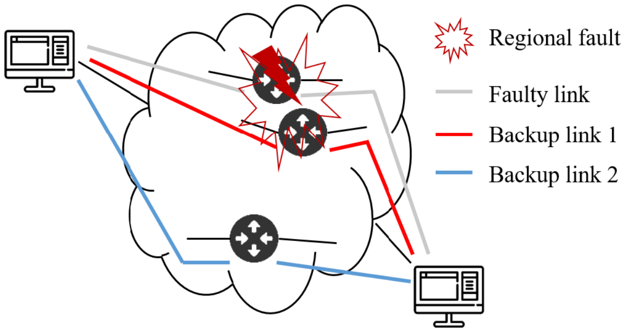

FIGURE 3. Regional failure.

TABLE 1. Parameter Explanation.

Symbol	Meaning
L_i	initial load of node i initial load of link i
D_i	degree of node i
Cp_i	maximum load capacity of node i maximum load capacity of link i
ΔL_i	increase load of node i
C_i	clustering coefficient of node i
E_i	the actual number of edges between node I and neighbour nodes
k_i	the number of neighbour nodes of node i
$C_{k_i}^2$	the theoretical number of edges between mode i and neighbour nodes
$d_{i,j}$	the distance between i and j
fp_i	the failure probability of path i
wp_i	the working probability of path i
S(t)	the number of susceptible nodes at t time
I(t)	the number of infected nodes at t time
R(t)	the number of recover nodes at t time
δ	the fault propagation speed
f	the proportion of faulty nodes
δ_{node}	the parameter that controlled initial load of nodes
ς_{node}	the parameter that controlled initial load of nodes
δ_{link}	the parameter that controlled initial load of links
ς_{link}	the parameter that controlled initial load of links
ε_{node}	the parameter that controlled capacity of nodes
ε_{link}	the parameter that controlled capacity of links

only be affected regionally. Thus, it can alleviate network congestion after failure.

Fig. 3 shows the regional failure situation. In the figure, we can see that the router that is closer to the failed router is affected by a regional failure. Backup link 1 will likely fail. Thus, we prefer to choose backup link 2 instead of backup link 1.

In this paper, the related parameters are shown in the Table 1.

B. LIMITATION FACTOR

The following definitions are required in the analysis of cascade fault network propagation models.

Definition 1: We define the initial load of each node i in the complex network, L_i. For L_i, there is

$$L_i = \varsigma_{node} D_i^{\delta_{node}}, \tag{1}$$

where D_i represents the degree of node i. δ_{node} and ς_{node} are adjustable parameters. They are used to control initial load of the node, and both their values are greater than 0 [25].

Definition 2: We define the initial load of each link i in the complex network, L_i. For L_i, there is

$$L_i = \varsigma_{link} D_i D_j^{\delta_{link}}, \tag{2}$$

where D_i and D_j represent the degree of link endpoints i and j. δ_{node} and ς_{link} are adjustable parameters. They are used to control the initial load of the link, and both value are greater than 0.

Definition 3: We define that the max load of each node is restricted by the cost of the node. Thus, we define Cp_i as the capacity of node i.

$$Cp_i = (1 + \varepsilon_{node}) L_i, \tag{3}$$

where L_i is initial load, and constant ε_{node} is the redundancy factor representing the ability of the node to handle redundant loads [26].

Definition 4: We define that the max load of each link is restricted by the cost of the link. Therefore, we define Cp_i as the capacity of link i.

$$Cp_i = (1 + \varepsilon_{link}) L_i, \tag{4}$$

where L_i is the initial load and constant ε_{link} is the redundancy factor representing the ability of the link to handle redundant loads.

C. MODELLING BASED ON THE VIRUS PROPAGATION MODEL

1) THE NETWORK ARCHITECTURE IS BASED ON SDN

In the study of the cascade fault propagation model, we introduce the SDN framework and virtualization technology based on the communication complex network. The virtualization technology can virtualize the network resources and provide the required functions to the upper layer. The SDN framework provides a unified control platform. We can acquire the overall network information from the SDN control platform. If we can know the overall topology of the network, we can easily calculate the backup path when failures occur. As shown in Fig. 4, there is a combination of SDN and virtualization technology communication network architecture.

In the figure 4, the bottom layer is the physical device layer, such as the IP router and optical multiplexer, and the upper layer is the virtual network resource layer. The resources provided by the physical equipment, such as the storage and calculation of the virtual escape as a resource pool for the upper SDN controller platform, include Path Computation Elements (PCE) and Traffic Engineering Databases (TED). Those two functional modules combine to provide the function of calculating the alternate path.

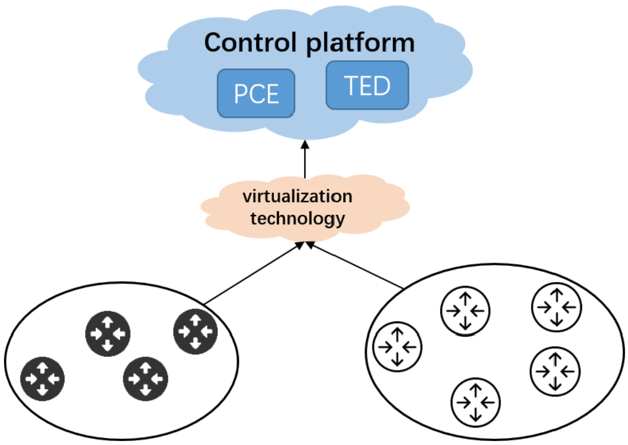

FIGURE 4. The communication network architecture.

2) LOAD REALLOCATION MECHANISM

When cascade failures occur, the traffic on the failed nodes needs to be allocated to neighbour nodes. We assume that the redistribution of this process is not evenly distributed but is based on the relationship between the nodes. We define the increase in the load on the neighbour node j as ΔL_j:

$$\Delta L_j = \frac{C_j L}{\sum_{s \in (neighborhood\ of\ i)} C_s}, \tag{5}$$

$$C_j = \frac{E_i}{C_{k_i}^2} = \frac{2E_i}{k_i(k_i - 1)}, \tag{6}$$

where Cj represents the clustering coefficient of node j, i represents failed node. And L means the traffic which needs to be re-allocated. Besides, C_s represents the clustering coefficient of the nodes which are neighbourhood of the failed node. The clustering coefficient is introduced into the load reassignment so that the fault propagation process can be better simulated. The larger the clustering coefficient is among nodes, the easier it is to propagate the traffic. During the process of load redistribution, the load should be preferentially assigned to the nodes with large clustering coefficients, which can effectively reduce the scope and the possibility of further cascading failures.

3) REGIONAL ANALYSIS

While the reallocation process is started because of network failure, we can add a regional concept to effectively avoid the deterioration of the situation. In the analysis of regional concepts, we introduce the parameter Path Survivability to evaluate the probability that a normal path may fail due to a regional failure when another path fails. We use the distance between paths to define the path survivability.

We define the distance between the two links as equivalent to the distance between the midpoints of two links, such as in (7).

$$d_{l_1,l_2} = d_{m_1,m_2}, \tag{7}$$

where l1 and l2 represent two links and m1 and m2 represent the midpoint of each link, respectively. Similarly, we can compute the distance between path Pl consisting of the link $\{l_1, l_2 \ldots l_L\}$ and the link l.

$$d_{P_l,l} = \frac{\sum_{i=1}^{L} d_{l,l_i}}{L}. \tag{8}$$

In short, the distance between path Pl and link l is equal to the average of the distance between all the links constituting the path and link l. If link l fails, the survivability of path Pl can be calculated according to the above.

$$fp_{P_l} = e^{-\beta (d_{P_l,l})^2}, \tag{9}$$

$$wp_{P_l} = 1 - fp_{P_l} = 1 - e^{-\beta (d_{P_l,l})^2}, \tag{10}$$

where β is a constant and is defined as a failure factor. The term fp_{P_l} represents the failure probability that is affected by the area fault of path Pl. Therefore, we can easily estimate the probability of normal work (10). We define wp_{P_l} as the survivability of path Pl.

4) ANALYSIS OF CASCADE FAULT PROPAGATION

After the network failure occurs, the faulty working path preferentially selects its alternate equipment to carry the service route. However, due to the limited resources and cost, there may only be some important links and nodes that support backup equipment in a network topology. Most of the them do not have backup devices. Therefore, we focus on building cascading fault propagation models for those that do not have alternate equipment. Consider the topology below as an example.

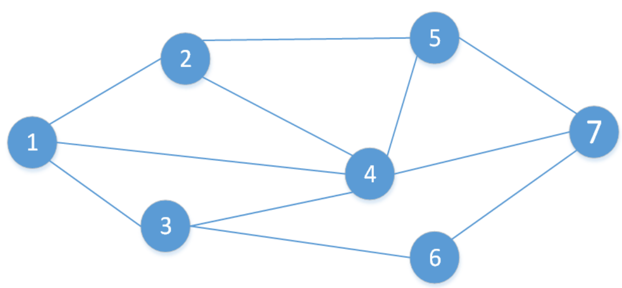

FIGURE 5. Network topology.

Figure 5 is a simple network topology with 7 nodes and 11 links. Assuming that there is a service path 1-4-7 in the topology of Figure 5, where node 4 is faulty due to a certain reason, the service carried on route 1-4-7 needs to be transferred to other working paths to ensure that the service can be provided normally. Now, the optional nodes are node 2 and node 3. First, we calculate the path survivability of nodes 2 and 3, then sort by the value. After that, we calculate the clustering coefficients of nodes 2 and 3 and allocate the traffic carried on the faulty node 4 according to the clustering coefficients. Now, if node 2 is allocated to the traffic of $\Delta\alpha$ and it has carried a certain traffic flow, it may be beyond the maximum capacity that node 2 can carry. Thus, a new round of the reallocation process begins. The figure below describes this process.

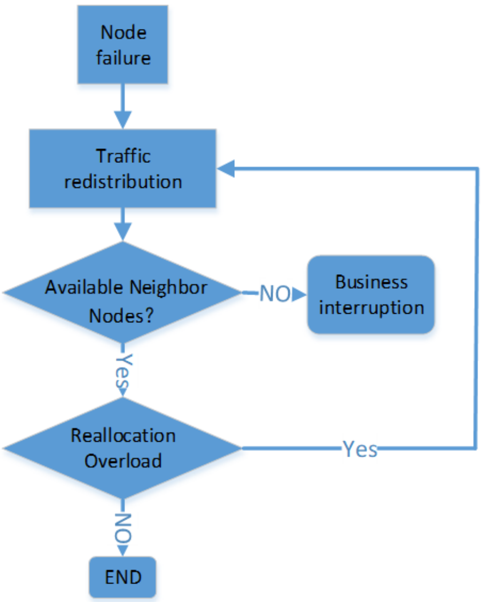

FIGURE 6. Cascade failure process.

Fig.6 briefly describes the process of fault propagation. In general, when failure occurs in the complex network, due to the network repair mechanism the traffic on the failure node will be allocated to other working node. This is the basis process of cascade failure.

In order to simulate the cascade fault propagation process, we introduce the virus propagation model. We set S(t) as the number of susceptible nodes, I(t) as number of infectious nodes and R(t) as the number of immune nodes at time t. In addition, we have $S(t) + I(t) + R(t) = 1$. According to the virus propagation model, we can obtain the following differential equations.

$$\begin{cases} \frac{dS}{dt} = -\alpha IS \\ \frac{dI}{dt} = \alpha SI - \beta I \\ \frac{dR}{dt} = \beta I, \end{cases} \tag{11}$$

We assume that there is only one or several infected nodes in the initial state, and there is no immune node. Therefore, in the initial state, the number of susceptible nodes is S0, the number of infected nodes is I0, and the number of immune nodes is 0. We set the virus transmission rate to $\delta = \alpha/\beta$. According to (11), we arrive at

$$\frac{dI}{dS} = \frac{\beta - \alpha S}{\alpha S} \tag{12}$$

$$\mathrm{dI} = \frac{\beta - \alpha S}{\alpha S}\mathrm{dS} = \frac{\beta}{\alpha S}dS - dS \tag{13}$$

$$\mathrm{I} = 1 - \mathrm{R} - S_0 e^{(-R\frac{\alpha}{\beta})} \tag{14}$$

$$\frac{dR}{dt} = \beta(1 - R - S_0 e^{(-R\delta)}) \tag{15}$$

$$\mathrm{dR} = \beta(1 - \mathrm{R} - S_0 e^{(-R\delta)})\mathrm{dt} \tag{16}$$

$$\mathrm{R} = 1 - S_0 e^{(-R\delta)} \tag{17}$$

The availability of the initial conditions suggests the following:

$$\mathrm{R} = 1 - e^{-R\delta} \tag{18}$$

There are many propagation theories, like the power flow model proposed in [11] to simulate cascade failure in power network. However, we use virus propagation model to simulate the cascade failure propagation in complex network. Compared with the model proposed in [11], the virus propagation model is more suitable for complex network. And the power flow model is more about power grid characteristics. The simulation of virus propagation is more in line with the actual situation. We will verify it in the simulation part.

So that if we know the speed of cascade failure propagation, we can get the number of failed nodes at some point according to Eq.18. We map different state network nodes as different nodes in virus model, and we can consider the speed of cascade failure as virus propagation. It is obviously that if we control the speed of cascade failure propagation we can relief the impact of cascade failure. Therefore, we take into account the optimization measures that include avoiding regional fault and reasonable traffic redistribution. We will discuss specific optimization algorithm in the next section.

5) THE CASCADE FAULT OPTIMIZATION REASSIGNMENT ALGORITHM

The above analysis is about cascading failure propagation. On this basis, we can propose our optimization reassignment algorithm as Algorithm 1.

The detailed algorithm is as follows.

1) When failure occurs in the network, the traffic on the failure node need to be reallocated to other working nodes. So the reassignment algorithm is triggered.

2) We first store the availability neighbour nodes in a set S and information of links between the nodes and failure node in set L.

3) We use the information about links stored in set to calculate the link survivability. Moreover, we select the links, which satisfy the threshold and extract the endpoint of the links.

4) Calculate the clustering coefficient of appropriate nodes and reallocated the traffic according to the clustering coefficient.

5) The control plane send the information to related nodes.

The brief algorithm is as follows.

In Fig. 7, we can see the simple process of traffic reallocation. We set the scenarios which has two reallocation process. In the beginning, the left red node and the red link fail. The L traffic on the red node needs to be reallocated. There are three neighbour nodes connected to it. We first calculate the link survivability of each available link and select the appropriate node, allocating traffic according to the clustering coefficient of each node. Nevertheless, the upper node is overloaded after reallocation. Therefore, after one step, the second-round reallocation distributes traffic again.

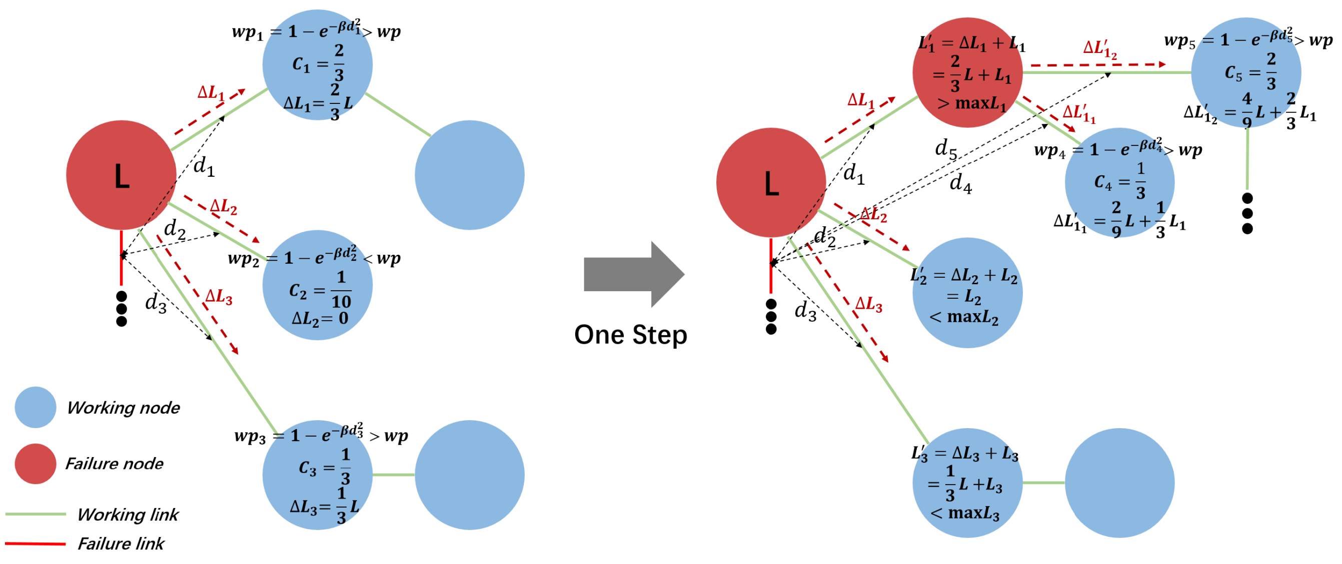

FIGURE 7. The traffic reallocation process.

As same as the previous step, we find the two neighbour nodes connected to the failed node. After calculation, the two neighbour nodes both meet the requirement of link survivability. So we reallocate the traffic to the two neighbour nodes according to the clustering coefficient.

6) ANALYSIS OF ALGORITHM COMPLEXITY

In the proposed mechanism, there are two stages. First, we calculate the survivability of each candidate node and contrast it with a threshold. This process only traverses the candidate node once. The complexity of this process is O(n). After selecting nodes, we need to calculate the clustering coefficient of each selected node. The complexity of the whole process is $O(n^2)$. Finally, we calculate the traffic assigned to each node based on the clustering coefficient. Therefore, we can obtain the complexity of the whole mechanism as follows.

$$\Omega = O\left(n^2\right) + \mathrm{O}\,(\mathrm{n}) \tag{19}$$

IV. SIMULATION

A. CASCADE FAULT DYNAMIC ANALYSIS

We use random network topology to simulate network cascade fault propagation. As shown in Figure 8, the network topology has 20 nodes and 41 edges. In terms of the simulation, we do not consider the existence of standby equipment. The green node represents normal work node and red node represents fault node. The blue edge indicates the normal working link, and the grey edge indicates the faulty link. In Fig. 8 a), the network is in the initial state, and each node is in the normal working state. We randomly selected a node as a fault node and set the virus transmission rate δ as 1. When a node fails, the business it carried is spread out according to the clustering coefficient and the survivability. Fig. 8 b) indicates the network state at a certain moment, and the final network state is shown in Fig. 8 c). All traffic through the network is completely interrupted.

B. SIMULATION ENVIRONMENT

In this paper, using Matlab, we randomly generate a network topology that is composed of 1000 nodes. To facilitate the simulation experiment, the topology is divided into 10 parts. As shown in Figure 9, 10 regions are connected by trunks (blue links), and each area is composed of 100 nodes. The 10 regional sub-networks constitute the 1000-node network topology used by the simulation experiment.

To facilitate the analysis of cascading faults in the simulation process, the parameter f is defined as the ratio of the number of failed nodes to the total number of nodes in the network.

$$\mathrm{f} = \frac{number\ of\ fault\ points}{number\ of\ totall\ points} \tag{20}$$

C. SIMULATION RESULT

1) FAULT PROPAGATION PROCESS

Based on the network topology proposed in the previous section, we compare the method of cascade fault modelling used in this paper with the method of [11].

In Fig. 10, the horizontal axis represents the step. We define each step as one change of the network. The ordinate is the percentage of the number of failed nodes and the total number of points. We can see from the figure that the model used in this paper has a relatively flat node growth rate compared with the model used as a reference at the beginning of the paper. The model we proposed has an earlier start time and later smooth time. The total process of cascading failures in the network is longer. The model we proposed is more in line with the actual situation of complex network cascade fault propagation.

Algorithm 1 Cascade Fault Optimization Reassignment Algorithm

Input D={d1,d2,d3...}, ki, Ei, Loadi;
Output $\Delta\alpha = \{\Delta\alpha 1, \Delta\alpha 2, \Delta\alpha 3 \ldots\}$
When cascade failure occurs, trigger the reassignment algorithm.
Stored availability point **in** S={s1,s2,s3...} and L={l1,l2,l3...};
For li in L which endpoint in S
Calculate survivability value of each link in L which endpoint in S**;**

$$wp_{P_l} = 1 - e^{-\beta(d_{P_l,l})^2}$$

Reorder L according to the value of survivability
Extract endpoint of each link stored in P;
For sj in P
Calculate clustering coefficient

$$C_j = \frac{2E_i}{k_i(k_i - 1)}$$

Calculate increased loads according to the loads on failure point Loadi

$$\Delta L_j = \frac{C_j L}{\sum_{s \in (neighborhood\ of\ i)} C_s}$$

Stored in SET $\Delta\alpha$;
End

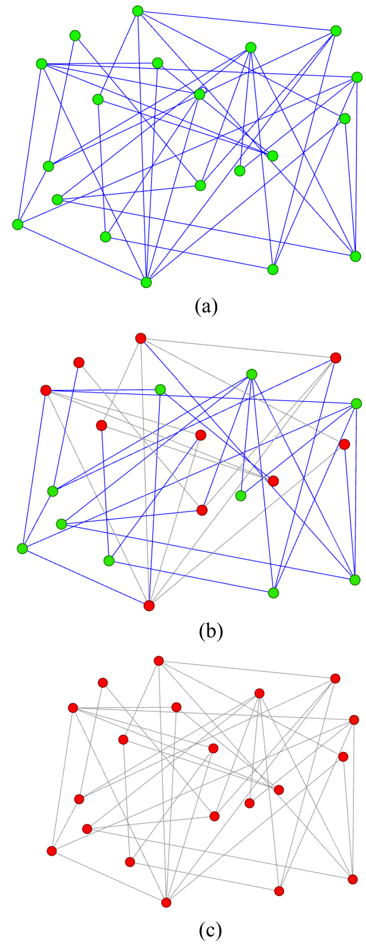

FIGURE 8. The simulation of the cascade fault propagation process. a) The initial network topology. b) The network topology at some point. c) The final network topology.

2) SIMULATION ANALYSIS OF FAULT PROBABILITY BASED ON DIFFERENT REGIONS

Fig. 11 a) shows the trends of failure nodes at different probabilities of regional failure. In the figure, we can see that the number of fault nodes increased significantly during steps 0 to 300, and it tends to be smooth after 300 steps. The final proportion of the failure nodes is basically stable at 79%. The shallower the colour in the figure, the greater the number of fault nodes is, and the deeper the colour is, the fewer fault nodes there are. The failure node growth rate is the same in the case of different regional fault probability. Because the optimized reallocation algorithm considers the regional fault, the differences in the fault nodes between the different regional fault probabilities are small. Overall, if the probability of a regional fault is large, the number of nodes affected by the cascade failure will increase.

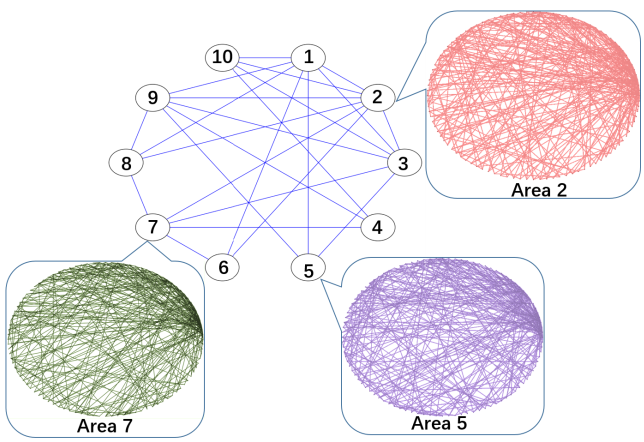

FIGURE 9. The simulation topology.

3) SIMULATION ANALYSIS BASED ON DIFFERENT INITIAL LOADS

The cascade fault mitigation mechanism is analysed according to the different initial loads. We select four different loads: normal load, 5% load, 7% load and 10% load. The horizontal axis represents the time step, and the vertical axis represents number of failed nodes. It can be seen from Figure 12 that the failure node growth rate is basically the same in the case of different loads. In addition, the number of fault nodes increases with the increase in the initial load in a certain step. In the event of a higher initial load, there are also more affected nodes.

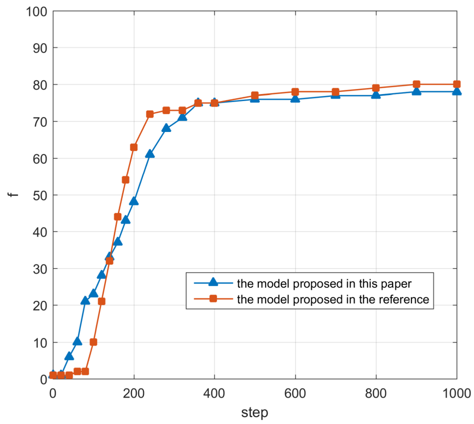

FIGURE 10. The comparison of two cascade fault model.

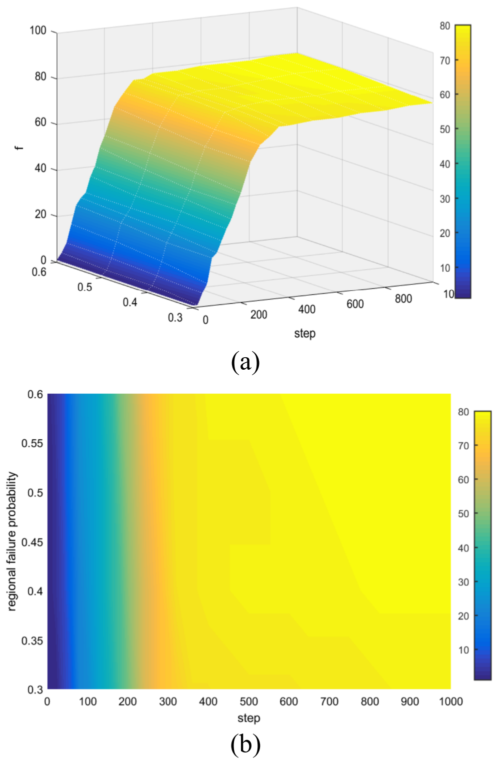

FIGURE 11. The analysis of cascade failure under different probabilities of regional fault occurring.

From the perspective of the remaining business, the situation of higher initial business loads shows a larger reduction in Figure 13. However, in the case of a final stable situation, one of the higher initial loads has higher remaining service.

4) SIMULATION ANALYSIS BASED ON AN OPTIMIZED REALLOCATION PROCESS

Figure 14 shows a comparison of the improved reallocation algorithm and the unimproved algorithm. It can be seen from the figure that the growth rate of the fault nodes is the same, whether it is the traditional cascade fault process or the cascade fault process using the improved

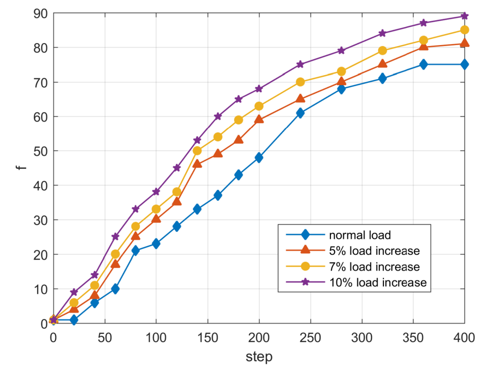

FIGURE 12. Analysis of the failure nodes of cascade failure under different initial loads.

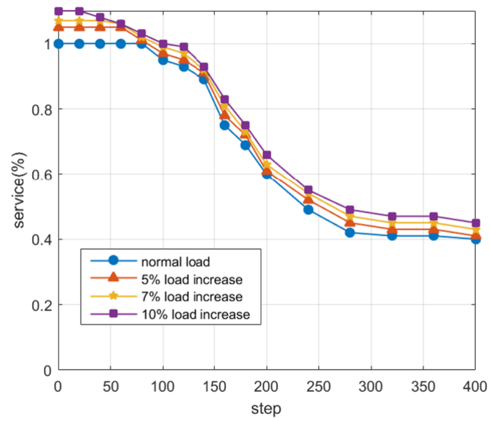

FIGURE 13. The service analysis of cascade failure under different initial loads.

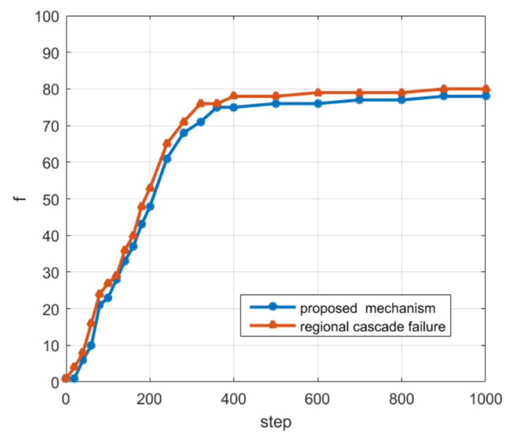

FIGURE 14. The failure nodes analysis of cascade failure under different reallocation mechanisms.

reallocation algorithm. The network topology reaches a steady state at approximately 300 steps. However, in the same step, the traditional cascade fault process has a higher number of fault nodes than the optimization algorithm proposed in this paper. The proposed improved redistribution mechanism can play a certain mitigation role in the cascade fault propagation process.

In Figure 15, from the perspective of the remaining business, the performances of those two mechanisms are similar.

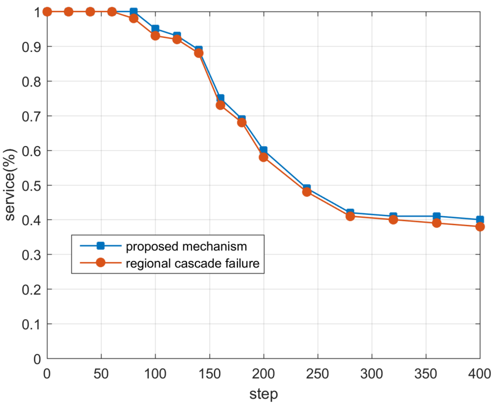

FIGURE 15. The service analysis of cascade failure under different reallocation mechanisms.

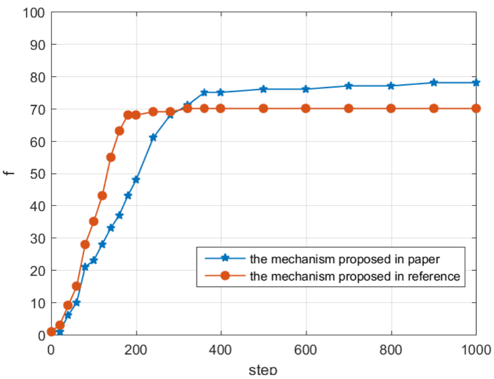

FIGURE 16. The analysis of failure nodes of cascade failure under two optimization mechanisms.

However, in the same step case, the remaining service of the proposed mechanism is higher than the traditional one. No matter what kind of mechanism, they tend to be smooth after 400 steps.

5) COMPREHENSIVE ANALYSIS OF CASCADE FAILURE OPTIMIZATION MEASURES

Parandehgheibi *et al.* [27] proposed another measure to mitigate cascade failure. The main operation is to reduce the nodes affected by the cascading failures and reduce the services from the source. This can alleviate the effect of cascading failures to a certain extent. We compare those two mechanisms from two perspectives, services and the number of affected nodes. The results are shown in Figs. 16 and 17.

In the Fig. 16, the horizontal axis represents the time step, and the vertical axis is the number of affected nodes. It can be seen from the figure that in the initial stage of the cascade fault propagation, the method proposed in [27] manually cancelled the effected nodes. Therefore, the growth rate is higher than that from the algorithm proposed in this paper. However, because of the control of traffic, the network topology is stabilized earlier, and the number of fault nodes is lower than in the proposed algorithm. Anyway, from the perspective

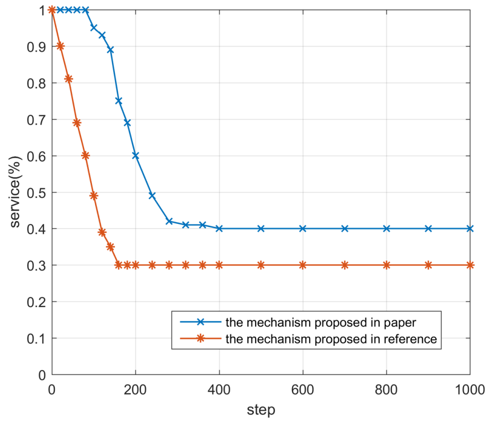

FIGURE 17. The service analysis of cascade failure under two optimization mechanisms.

of the remaining business, the proposed mechanism in this paper is better than the mechanism in [27]. In the Fig. 17, the vertical axis shows the percentage of the services. In the beginning of the cascading failures process, the mechanism in [27] has a rapid downward trend and is lower than the mechanism proposed in paper. In the final network topology, there are more working nodes in the mechanism in [27], but fewer services than with the mechanism proposed in paper because of the service control. In today's complex networks, people are more concerned about whether the business can perform better. Thus, with this factor in mind, selecting the mechanism proposed in this article can be more conducive to controlling cascade failure.

V. CONCLUSION

In this paper, we use the virus propagation model to analyse the propagation process of cascade failure. We analyse the optimization of cascade failure from two perspectives, regional fault and traffic reallocation. We introduce path survivability to evaluate the probability that the path will work properly in the case of regional failure. In the process of optimizing traffic reallocation, we consider reallocating the traffic according to node clustering coefficient. In the final simulation section, the advantages of the mechanism are verified. In future work, we will further study methods of improving cascade failure.

REFERENCES

[1] A.-E. Baert, V. Boudet, and A. Jean-Marie, "Data replication optimisation in grid delivery network," *Int. J. Grid Utility Comput.*, vol. 1, no. 4, pp. 287–295, 2009.
[2] A. Damian, A. Costan, V. Cristea, and I. Legrand, "A visualisation technique for network topology transformation within MonALISA monitoring framework," *Int. J. Grid Utility Comput.*, vol. 2, no. 2, pp. 119–129, 2011.
[3] S. Sawamura, A. Barolli, A. Aikebaier, M. Takizawa, and T. Enokido, "Design and evaluation of algorithms for obtaining objective trustworthiness on acquaintances in P2P overlay networks," *Int. J. Grid Utility Comput.*, vol. 2, no. 3, pp. 196–203, 2011.
[4] E. Niewiadomska-Szynkiewicz and A. Sikora, "Simulation-based design of self-organising and cooperative networks," *Int. J. Space-Based Situated Comput.*, vol. 1, no. 1, pp. 68–75, 2011.
[5] O. Boyinbode, H. Le, and M. Takizawa, "A survey on clustering algorithms for wireless sensor networks," *Int. J. Space-Based Situated Comput.*, vol. 1, nos. 2–3, pp. 130–136, 2011.

[6] H. Peng, Y. Tian, J. Kurths, L. Li, Y. Yang, and D. Wang, "Secure and energy-efficient data transmission system based on chaotic compressive sensing in body-to-body networks," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 3, pp. 558–573, Jun. 2017.
[7] P. Dey, R. Mehra, F. Kazi, S. Wagh, and N. M. Singh, "Impact of topology on the propagation of cascading failure in power grid," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1970–1978, Jul. 2016.
[8] S. R. Gupta, F. S. Kazi, S. R. Wagh, and N. M. Singh, "Probabilistic framework for evaluation of smart grid resilience of cascade failure," in *Proc. Innov. Smart Grid Technol.-Asia*, May 2014, pp. 255–260.
[9] P. Dey, M. Parimi, A. Yerudkar, and S. R. Wagh, "Real-time estimation of propagation of cascade failure using branching process," in *Proc. IEEE, Int. Conf. Power Eng., Energy Elect. Drives*, May 2015, pp. 629–634.
[10] M. Sanghavi, S. Tadepalli, T. J. Boyle, M. Downey, and M. K. Nakayama, "Efficient algorithms for analyzing cascading failures in a Markovian dependability model," *IEEE Trans. Rel.*, vol. 66, no. 2, pp. 258–280, Jun. 2017.
[11] X. Zhang, C. Zhan, and K. T. Chi, "Modeling the dynamics of cascading failures in power systems," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 2, pp. 192–204, Jun. 2017.
[12] Z. Yu, "Research on cascading failure of complex system and its virus propation," in *Proc. 9th Int. Conf. Adv. Comput. Intell.*, Feb. 2017, pp. 115–120.
[13] S. Liu, B. Chen, T. Zourntos, D. Kundur, and K. Butler-Purry, "A coordinated multi-switch attack for cascading failures in smart grid," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1183–1195, May 2014.
[14] R. J. La, "Cascading failures in interdependent systems: Impact of degree variability and dependence," *IEEE Trans. Netw. Sci. Eng.*, to be published.
[15] M. Rahnamay-Naeini and M. M. Hayat, "Cascading failures in interdependent infrastructures: An interdependent Markov-chain approach," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1997–2006, Jul. 2016.
[16] J. Yan, H. He, and Y. Sun, "Integrated security analysis on cascading failure in complex networks," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 3, pp. 451–463, Mar. 2014.
[17] Y. Fan, Y. Cheng, Y. Chen, and Y. Yang, "Analysis of cascading failure of circuit systems based on load-capacity model of complex network," in *Proc. 2nd Int. Conf. Rel. Syst. Eng.*, Jul. 2017, pp. 1–4.
[18] S. Iwaori and M. Hayashi, "Analysis of area-to-area cascading failure rate for telecommunications networks," in *Proc. IEEE Int. Symp. Local Metropolitan Area Netw.*, Jun. 2017, pp. 1–2.
[19] Z. Liu, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran, "A submodular optimization approach to controlled islanding under cascading failure," in *Proc. Int. Conf. Cyber-Phys. Syst.*, Apr. 2017, pp. 187–196.
[20] D. Lv, A. Eslami, and S. Cui, "Load-dependent cascading failures in finite-size Erdös-Rényi random networks?" *IEEE Trans. Netw. Sci. Eng.*, vol. 4, no. 2, pp. 129–139, Apr. 2017.
[21] Z. Xu, A. A. Julius, and J. H. Chow, "Robust testing of cascading failure mitigations based on power dispatch and quick-start storage," *IEEE Syst. J.*, to be published.
[22] Z.-H. Li and D.-L. Duan, "Node importance of complex networks based on cascading failure dynamic," in *Proc. Int. Conf. Intell. Netw. Collaborative Syst.*, Sep. 2016, pp. 31–35.
[23] L. Li, H. Peng, J. Kurths, Y. Yang, and H. J. Schellnhuber, "Chaos–order transition in foraging behavior of ants," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 23, pp. 8392–8397, 2014.
[24] Y. Chen, L. Li, H. Peng, J. Xiao, and Q. Wu, "Dynamic multi-swarm differential learning particle swarm optimizer," *Swarm Evol. Comput.*, vol. 39, pp. 209–221, Apr. 2018.
[25] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 4, p. 045104, 2004.
[26] W. Jian-Wei and R. Li-Li, "Cascading failures on complex networks based on the local preferential redistribution rule of the load," *Acta Phys. Sinica*, vol. 6, pp. 3714–3721, Jun. 2009.
[27] M. Parandehgheibi, E. Modiano, and D. Hay, "Mitigating cascading failures in interdependent power grids and communication networks," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Nov. 2014, pp. 242–247.

LANLAN RUI received the Ph.D. degree in computer science technology from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2010. She is currently an Associate Professor with the State Key Laboratory of Networking and Switching Technology, BUPT, China. Her research interests include IOT, MEC, content-based measurement and analysis, quality of service (QoS) and intelligent theory, and technology of network services. As a result of the standardization in network management, she received the 3GPP SA5 Outstanding Contribution Award.

XIAOMEI WANG is currently pursuing the master's degree with the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China. Her research interests include network reliability assessment and network failure recovery.

XUESONG QIU received the Ph.D. degree in computer science from the Beijing University of Posts and Telecommunications (BUPT) in 1999. He is currently a Professor with BUPT and the Deputy Director of the State Key Laboratory of Networking and Switching Technology. He has authored about 150 SCI/EI index papers and received 13 national and provincial scientific and technical awards. His research interests include communication network and communication software, Internet routing and applications, smart service provisioning in mobile social network, network management, and information-centric networking. He received the Beijing Higher Education Young Elite Teacher Plan in 2005. He is a Web Chair in 2016 IEEE/ACM IWQoS.