IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 5 June 2022, accepted 30 June 2022, date of publication 5 July 2022, date of current version 11 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3188659

RESEARCH ARTICLE

An Exceeding Recovery Model for Enhancing Network Resilience Against Cascading Failures

JIE LI[1], YING WANG[1], AND JILONG ZHONG[2]
[1]Air Traffic Control and Navigation College, Air Force Engineering University, Xi'an 710038, China
[2]National Institute of Defense Technology Innovation, PLA Academy of Military Science, Beijing 100071, China

Corresponding author: Jilong Zhong (z_jilong@sina.cn)

This work was supported in part by the Project of the National Natural Science Foundation of China under Grant 72001213, and in part by the National Social Science Foundation of China under Grant 19BGL297.

ABSTRACT Failure and recovery underlie many complex systems ranging from critical infrastructures to organisms. In many real complex systems, the reliability of repaired components is improved due to the exceeding recovery mechanism, and such systems typically have enhanced failure resistance. The main motivation of this study lies in developing an exceeding recovery model to capture the exceeding recovery mechanics of complex network systems. In the proposed model, cascading failure and exceeding recovery perform concomitantly. The network resilience analysis is performed in the Barabási–Albert and Erdős–Rényi networks by focusing on the exceeding recovery process from random and targeted attacks. The results show that for a given initial failure size, there is a critical value of the exceeding recovery coefficient above which the network can restore to the normal state, but below this value, the network abruptly collapses. The proposed model is compared with the conventional recovery model. The comparison indicates that the proposed model can recover to a significantly high level in a short recovery time and at a low recovery cost. The exceeding recovery mechanism strongly affects the failure–recovery property, which is expressed as reduced risk of a secondary failure at the micro level and enhanced heterogeneity of the load distribution at the macro level. These findings provide a guideline to address the exceeding recovery problem of a network and can help to design networks with better resilience against cascading failures.

INDEX TERMS Complex network, cascading failure, network recovery, network resilience.

I. INTRODUCTION

Cascading failure is a universal phenomenon of large–scale failures that can lead to catastrophic events in many complex network systems [1]–[3]. The dynamic mechanisms of a cascading failure have received increasing attention in recent years [4]–[6]. It has been demonstrated that spontaneous recovery or deliberate repair is often accompanied by a cascading failure [7]–[9]. Many models have been developed to depict this recovery property of a network [10]–[12]. Buzna *et al.* [13] studied the effectiveness of recovery strategies for a dynamic model and proposed a few recovery strategies based on the network's state and topology. Shang [14] proposed a general model of localized recovery, where some of the neighboring nodes are repaired in an invasive way from a seed node. Recently, to model more realistic recovery adequately, scientists have paid great attention to the dynamic recovery model, where failure and recovery of system components usually occur concomitantly and have a competing relationship [7]–[9], [15], [16]. Di Muro *et al.* [7] proposed a model in which the recovery process was applied from the first step of the cascade of failures, capturing the competition between the cascading failure and recovery of interdependent networks. Wu *et al.* [17] have improved on this method and proposed a new preferential recovery strategy based on connectivity links to determine the recovery impact of mutual boundary nodes in interdependent networks. Majdandzic [8] and Böttcher1 *et al.* [9] developed a model where spontaneous failure and recovery perform concomitantly; this model

The associate editor coordinating the review of this manuscript and approving it for publication was Yichuan Jiang.

can mitigate a spontaneous failure and damage spread in dynamical networks. La Rocca *et al.* [18] proposed a dynamic recovery strategy where the isolated network component is reconnected to the functional giant component with a probability γ at each time-step of the cascading failure process. Recent research has revealed that dynamic recovery models are effective to improve network resilience against abrupt spontaneous and cascading failures [15], [16], [19].

The existing studies on network recovery have been mostly based on the assumption that failed system components can recover to their pre-failure levels of functioning [20], [21]. However, in practice, a repaired system component can be even more reliable than it was before the failure occurred [22], [23]. There have been many examples of this phenomenon. As described in the susceptible–infected–recovered (SIR) epidemiological model [24], when infectious individuals infect susceptible individuals, some of the susceptible individuals recover and develop an enhanced protective immunity to secondary infections. When a virus infects its host, the arms race between the immune system and the virus may enhance immunocompetence [25]. Due to the presence of exceeding recovery mechanisms, such systems typically have enhanced failure resistance and adaptability feature, but this phenomenon has not been systematically explored yet. Inspired by the mentioned phenomenon, this study develops an exceeding recovery model (ERM) of a complex network to capture the exceeding recovery mechanics of complex network systems.

In this work, the ERM is developed following three fundamental assumptions: (i) a node's failure and recovery are alternated, (ii) the recovery process follows the exceeding recovery rule that repaired nodes have an improved node capacity, (iii) the load redistribution follows the preference rule that repaired nodes are assigned more additional load than failure nodes. The interplay between these three assumptions results in a phase diagram of network resilience, where the network state switching is observed. In addition, compared with the conventional recovery model, the proposed ERM can achieve a high level of functional recovery in a short recovery time and at a low recovery cost. The correlation analysis of the proposed ERM is performed, and the results indicate that lower secondary failure probability and higher heterogeneity of the load distribution are effective in improving the network resilience against cascading failures. The proposed ERM model provide a useful insight into designing a highly reliable network. The main contributions of this study are as follows: 1) an exceeding recovery model of networks to capture the exceeding recovery mechanics of complex network systems is proposed; 2) a network resilience indicator considering both the structure and the nodal load is developed; 3) the effectiveness of the exceeding recovery for improving the network resilience against cascading failures is verified.

The rest of this paper is organized as follows: In Section II, we describe ERM model in detail and give the theoretical analysis of the recovery strategy. In Section III, we develop four performance assessment metrics, including network capability ($Q(t)$), network resilience (R), secondary failure percentage (p_{sf}), and recovery cost (Rc). The numerical examples and characteristics analysis of ERM model are discussed in Section IV. Section V concludes the work.

II. METHODOLOGY

A. ERM MODEL

In many real–world networks, the load and capacity of a node are positively correlated with the node degree. For instance, on the Internet [26] (transportation [27] or wireless sensor networks [28]), a router with a larger number of connection edges (corresponding to the large node degree) allows more information (corresponding to the larger load) to be routed through it, and it is often more powerful (corresponding to the larger capacity) to process information flow to avoid crowding phenomena. Many load–capability models based on the local information have been developed to capture the association between the node load and capacity and node degree [29], [30]. Motivated by these, in this study, the initial load $L_i(0)$ of the ith node is correlated with its degree k_i, which is expressed is as $L_i(0) = k_i^{\alpha}$, and node capacity C_i^N has a linear relationship with the initial load, which is expressed as $C_i^N = (1+\beta) L_i(0)$; α reflects the heterogeneity of the node load, and β represents the redundancy of the load capacity. The ratio $L_i(0)/C_i^N$ is defined as a nodal load rate l_r, and it is calculated by $l_r = 1/1+\beta$. The initial random removal of nodes will trigger a cascading failure, and load of failed nodes is redistributed to the neighbor nodes according to the following rule:

$$\Delta_{ji} = \frac{L_j(0)}{\sum_{n\in\Gamma_i} L_n(0)} L_i(0), \tag{1}$$

where Γ_i denotes a set of neighboring nodes of node i, and node j belongs to Γ_i.

After the first round of load redistribution, we assume that the load of node y is greater than its capacity. Here, we define these overloaded nodes that have not been removed from the network as NTF. Before node y redistributes its load to its normal neighboring nodes, the proposed exceeding recovery model performs a recovery process to avoid or delay the collapse of the network. In this process, failed nodes are repaired as follows:

1) A fraction ρ of failed nodes, which are neighbors of the NTF, are repaired in descending order of the node degree. Here, we define the recovery strategy as KBS.

2) The repaired nodes are reconnected to the remaining normal nodes according to the initial network topology.

3) The load and node capacity of the repaired nodes at time t are respectively calculated as follows:

$$L_i(t) = 0, \tag{2}$$

$$C_i^R = (1+\gamma)\cdot C_i^N, \tag{3}$$

where γ is the exceeding recovery coefficient reflecting the node recovery ability.

After the recovery process, NTF redistributes its load to neighbors and switches from the to–be–failed state to the failed state. To maximize the potential of the repaired nodes and thus reduce the impact of the cascading failure, an NTF redistributes load to its neighbors following the preference rule, which is defined as follow:

$$\Delta_{jy} = \begin{cases} \dfrac{(1+\gamma)\, L_j(0)}{(1+\gamma)\sum\limits_{i\in V_y} L_i(0) + \sum\limits_{n\in\Gamma_y} L_n(t)} L_y(t) & j \in V_y \\ \dfrac{L_j(t)}{(1+\gamma)\sum\limits_{i\in V_y} L_i(0) + \sum\limits_{n\in\Gamma_y} L_n(t)} L_y(t) & j \notin V_y \end{cases}, \tag{4}$$

where Γ_y is the set of neighboring nodes of node y, and node j belongs to it, and V_y represents the set of repaired nodes in Γ_y. When $j \in V_y$, an NTF redistributes load to node j according to the load weight, so the node load rate l_r is the same as its initial value of $l_r = 1/1+\beta$. When $j \notin V_y$, an NTF redistributes load to node j, referring to the actual node load weight. The larger the value of γ is, the more additional load from an NTF is redistributed to the repaired nodes. For the overall execution, in addition to setting the initial failure as a separate time step, it is assumed that a time step corresponds to one failure–recovery–redistribution cycle, in which the overloaded nodes are first removed from the network (failure stage); then, some of the failed nodes are repaired (recovery stage), and finally, the overloaded nodes redistribute load to their normal neighboring nodes (redistribution stage). The current–moment load of repaired nodes is $L_j(t) = 0$, so the node load at the next time moment $(t+1)$ is:

$$L_j(t+1) = \begin{cases} \Delta_{jy} & j \in V_y \\ L_j(t) + \Delta_{jy} & j \notin V_y \end{cases}. \tag{5}$$

The above procedure is repeated until an equilibrium state is reached, where no more new failure nodes occur. The schematic illustration of the failure propagation and exceeding recovery processes is presented in FIGURE 1. The specific steps are as follows:

Stage 1: An initial attack is added to the second and sixth nodes, and their load is redistributed to their neighboring nodes according to the rule given by Eq. (1).

Stage 2: The second and sixth nodes fail, and all edges connected to them are removed. The load of the first and fifth nodes exceeds their node capacity due to the load redistribution, and they are converted into NTFs.

Stage 3: Before the NTFs (i.e., the first and fifth nodes) affect other normal nodes' states by redistributing their load, the second node is repaired according to the KBS.

Stage 4: The load of NTFs (i.e., the first and fifth nodes) is redistributed following the preference rule given by Eq. (4), and the node load is updated by Eq. (5). If at least one of the zeroth and second nodes fail, the failure–recovery process continues; otherwise, this process terminates.

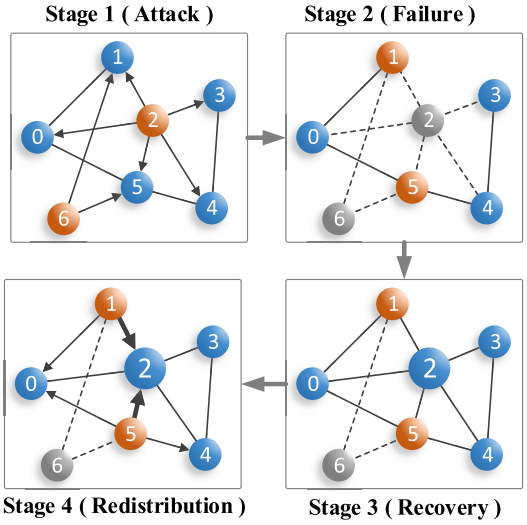

FIGURE 1. The schematic illustration of the failure propagation and exceeding recovery processes; blue, red, and gray nodes represent normal, NTF, and failed nodes, respectively. A node with a large area represents a repaired node having a large node capacity. Solid lines indicate the edges between nodes, and arrows show the flow directions of the load. The width of the solid lines indicates the relative amount of load redistributed through the edges. Dotted lines represent the failed edges.

B. THEORETICAL BASIS OF RECOVERY PROCESS

When failed nodes are repaired, the KBS gives priority to failed nodes with the greater degree. Here, assume node y fails at time t, and its load is redistributed to its neighboring nodes according to Eq. (4). Then, the failure force of node y is given by:

$$FS(t) = \frac{L_y(t)}{\sum\limits_{n\in\Gamma_y} C_n - L_n(t)}, \tag{6}$$

where $L_y(t)$ is the actual load of node y, and $\sum\limits_{n\in\Gamma_y} C_n - L_n(t)$ represents the sum of the capacity redundancies of all neighboring nodes of node y. If $FS(t) > 1$, some of the neighboring nodes can fail, and the larger the value of $FS(t)$ is, the greater the influence of the failure of node y is. Thus, the value of $FS(t)$ should be minimized as much as possible by repairing important failed nodes. Combining Eqs. (3), and (6), we have:

$$FS(t) = \frac{L_y(t)}{\sum\limits_{i\in V_y} \gamma(1+\beta) L_i(0) + \sum\limits_{n\in\Gamma_y} (1+\beta) k_n^{\alpha} - L_n(t)}. \tag{7}$$

The values of $L_y(t)$ and $L_n(t)$ are certain at time t, $FS(t)$ can be calculated as

$$FS(t) \propto \frac{1}{\sum\limits_{i\in V_y} \gamma(1+\beta) L_i(0)}. \tag{8}$$

By substituting $L_i(0)$ into Eq. (8), we obtain:

$$FS(t) \propto \frac{1}{\sum_{i \in V_y} k_i^{\alpha}}. \tag{9}$$

According to Eq. (9), there is an inverse relationship between $FS(t)$ and k_i. In addition, the local maximum degree node plays a crucial role in maintaining the network connectivity [31]. Thus, failed nodes with the greater node degree should be repaired preferentially.

III. ERM PERFORMANCE MEASUREMENT

To measure the network resistance to failure, the following performance indexes are introduced: network capability, network resilience, secondary failure probability, and recovery cost.

A. NETWORK CAPABILITY

Network capability, which reflects the real network performance affected by some internal and external factors, is related to both the network structure and node load. This is because the structural integrity of a network is fundamental for maintaining normal network operation, and a node's capability is usually expressed in the form that includes the node load. As mentioned in the "Model" section, the node function and failure propagation are performed in the form of a node load, while edges are used only to characterize the interrelationships between nodes. Therefore, the network capability is estimated only from the perspective of nodes. Considering both the network structure and node load, the normalized network capability function at time t is defined as follow:

$$Q(t) = \begin{cases} G(t) \cdot S(t) \\ S(t) = \dfrac{\sum_{i \in V_n} L_i(t)}{\sum_{i}^{N} L_i(0)} \end{cases}, \tag{10}$$

where $G(t)$ and $S(t)$ are related to the network structure and node load, respectively; $G(t)$ indicates the relative size of the giant component at time t, which is usually related to network connections. $G(t)$ is typically used to measure the structural integrity of networks. $S(t)$ represents the ratio between the total load of the normal nodes and the total load of all nodes at the initial time, which is used to measure the functional integrity of nodes. From the definition of $L_i(t)$ and Eq. (10), $S(t)$ is related to the node degree. V_n is a set of normal nodes; N is the network size; $Q(t) = 1$ at the initial time of $t = 0$.

B. NETWORK RESILIENCE

Network resilience reflects the recovery ability of the network, which is usually quantified according to the performance response process after a perturbation. A time-dependent metric proposed by Reed *et al.* quantifies the network resilience based on the network capability and recovery time of the network [32]. This definition reflects not only the level of network recovery but also the speed of network recovery. On this basis, we propose a resilience indicator according to the integral of the $Q(t)$ with respect to t as follow:

$$R = \frac{\int_{t_0}^{T^*} Q(t)\, dt}{T^* - t_0}, \tag{11}$$

where R represents the network resilience; T^* is the end time of the failure–recovery process; t_0 denotes the initial moment, and $t_0 = 0$. Network resilience usually varies with some underlying factors such as disturbance magnitude and place of impact. During the framing of the network resilience, no direct connection is modeled between the network resilience and these factors, but an indirect influence is assumed via network capability Q(t).

C. SECONDARY FAILURE PROBABILITY

During the failure–recovery process, repaired nodes face the risk of a secondary failure. A severe secondary failure can cause wastage of recovery resources and reduce recovery efficiency. To measure the risk of a secondary failure of repaired nodes, indicator P_{sf} is defined as follows:

$$P_{sf} = \frac{\sum_{i \in V_f} F_f^i - \|V_f\|}{\|V_f\|}, \tag{12}$$

where F_f^i indicates the failure frequency of a node i, which represents the total number of failures of node i during the entire cascading failure process, is actually a measurable value. In both simulation experiments and real systems, the value of F_f^i can be obtained directly by recording the number of failures of node i, and $F_f^i \geq 2$ when the node i is a secondary failure node; V_f is a set of failed nodes, and $\|V_f\|$ represents the number of failed nodes.

D. RECOVERY COST

During the exceeding recovery stage, repaired nodes are given a larger node capacity, which results in a higher recovery cost. The recovery cost Rc is calculated as follow:

$$Rc = \frac{\sum_{i \in V_r} (1+\gamma)(1+\beta) L_i(0) \cdot F_r^i}{\sum_{i}^{N} (1+\beta) L_i(0)}, \tag{13}$$

where V_r is a set of repaired nodes, and F_r^i is the recovery frequency of node i. γ represents the exceeding recovery coefficient, and β represents the redundancy of the load capacity. $L_i(0)$ is the initial load of node i. N represents the network size. The recovery cost relates to all failed nodes being repaired, including secondary failed nodes.

IV. NUMERICAL SIMULATIONS

The numerical simulations are conducted on Barabási–Albert (BA) and Erdős–Rényi (ER) networks, which have been commonly used to depict real–world network systems.

The BA network is one of the most familiar growing network models for generating scale-free networks, where the degree distribution follows a power law. The ER network is a typical random graph model, where N labeled nodes are connected with L randomly placed links [33]. In this paper, the BA and ER networks are generated using *NetworkX* (https://networkx.org/), which is a Python package for the construction, manipulation, and analysis of complex networks. The networks contain 10,000 nodes, and the average degree is $\langle k \rangle = 10$. The load heterogeneity parameter is set to $\alpha = 1.3$, and the node capacity parameter is set to $\beta = 0.5$. The recovery fraction is fixed to $\rho = 0.02$ and $\rho = 0.05$ in the BA and ER networks, respectively. The results are averaged over 3,00 iterations.

A. NETWORK RESILIENCE ANALYSIS

Network resilience usually varies with the size and location of the disturbance. In this section, we perform the network resilience analysis under the different initial failure sizes and failure modes (i.e., random attack and targeted attack).

We first test the network resilience under the random initial failure. The correlation between the network resilience (R) and the initial failure size (q) for different values of the exceeding recovery coefficient (γ) is presented in FIGURE 2(a) and 2(b). The results show that when $\gamma = 0$ (i.e., failed nodes are repaired with the initial node capacity), R decreases most as q increases. However, as γ increases, there is a gradual increase in R. This indicates that the network resilience can be enhanced by increasing the exceeding recovery coefficient. To examine the effect of the exceeding recovery coefficient on the network resilience further, the change in R with γ is analyzed, as shown in FIGURE 2 (c) and 2(d). The results showed that for a fixed value of q, R gradually increases with γ, and when the value of γ reaches a certain value, the increase in R is not obvious. This is because that the effect of the recovery is close to saturation.

Next, the resilience analysis is performed under targeted attacks ordered by the node degree to verify the effects of the attack location on the network resilience. As shown in FIGURE 3(a) and 3(c), for the BA network, R is almost the same for different γ when q is fixed. This is because when hub nodes are attacked, the BA network will completely collapse within a few steps [34], which restricts the full usage of the recovery process. For the ER network, FIGURE 3(b) and 3(d), the results are similar to those under the random initial failure, i.e., R decreases with q, and R can be enhanced by increasing γ. In addition, by comparing the results in FIGURE 2 and FIGURE 3, we find that the network possesses better resilience under random attacks compared to targeted attacks.

To further validate the effect of γ and q on R, the phase diagrams under the random initial failure are shown in FIGURE 4. In this study, the critical value of the network resilience is set to $R_c = 0.25$, but it can differ among application scenarios and network properties. Here, the network will be considered to fail completely as the real network resilience is less than R_c; otherwise, it is able to maintain essential functions. The results showed that there are three regions delimited in FIGURE 4 by the solid yellow curve, which represents the critical value of γ as a function of q, and the dashed orange line, which represents the critical value of q for $\gamma = 0$. The leftmost region (**Normal**) indicates the network does not crash for any value of γ. In the middle

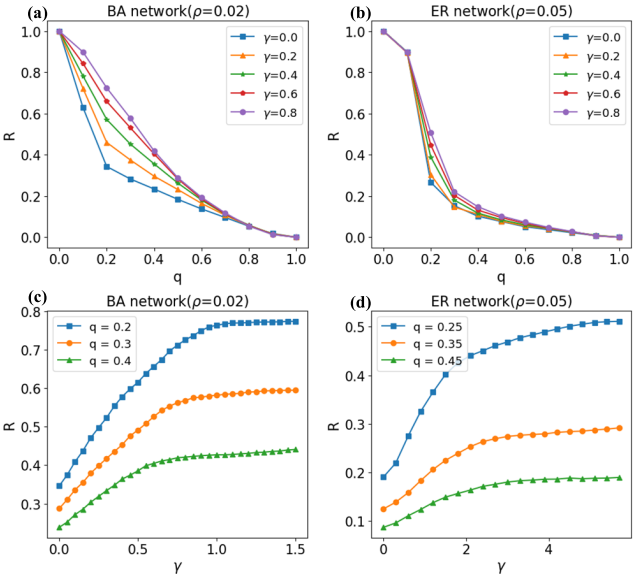

FIGURE 2. Network resilience under the random attack; R, q, and γ represent the network resilience, the initial failure size, and the exceeding recovery coefficient, respectively. (a), (b) The correlation between R and q for different values of γ; (c), (d) The correlation between R and γ under different values of q.

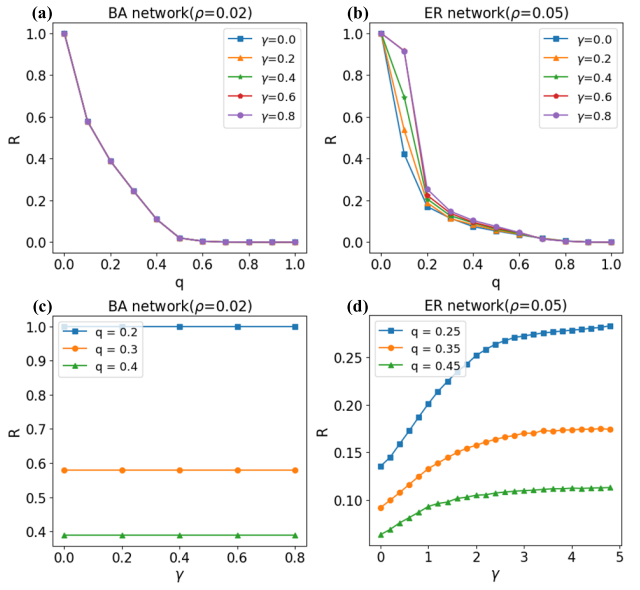

FIGURE 3. Network resilience under the targeted attack. (a), (b) The correlation between R and q for different γ; (c), (d) The correlation between R and γ under different q.

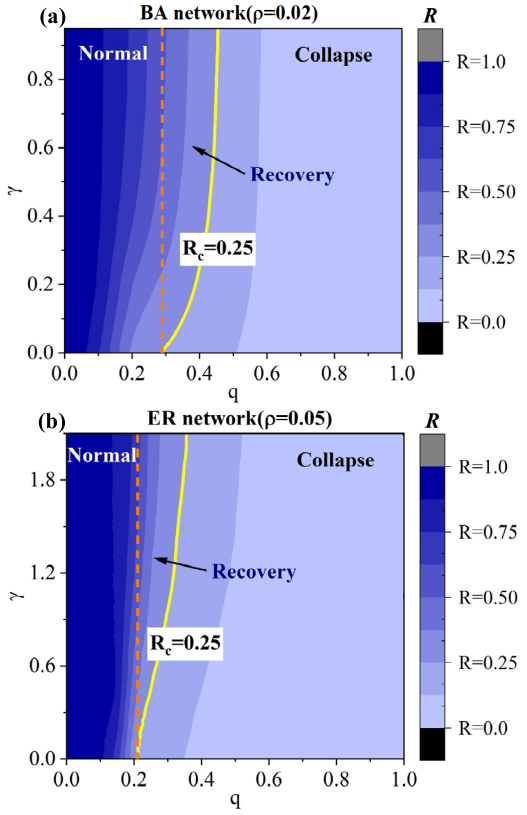

FIGURE 4. Phase diagrams in the $\gamma - q$ plane for BA (a) and ER (b) networks. The solid yellow line represents the critical value of γ as a function of q; the dashed orange line represents the critical value of q for $\gamma = 0$.

region (**Recovery**), the minimum γ is needed to prevent the network from falling out completely. Finally, in the rightmost region (**Collapse**), the exceeding recovery model cannot avoid network collapse. In FIGURE 4, the solid yellow curve corresponds to the minimum value of the exceeding recovery coefficient γ, which ensures the network will not collapse. The results indicate that the exceeding recovery model can effectively improve the network resilience performance, thus making the network more resilient against cascading overload failures. The presented results provide a suitable foundation for the calculation of the optimal exceeding recovery coefficient to keep a network functional.

B. ERM RECOVERY SUPERIORITY ANALYSIS

The contrast experiments under random initial failure are conducted to examine the recovery properties of the ERM. The ERM will degenerate into the conventional recovery model (CRM), characterized by repaired nodes with the initial node capacity, and this model is used as a reference model in the comparison. The comparison results of the CRM and ERM under different γ values are presented in FIGURE 5, where the dashed blue line indicates the CRM, and solid colored lines correspond to the ERM; $Q(T^*)$ represents the network capacity at the end time of the failure–recovery process T^*, and Rc denotes the recovery cost. According to the results in FIGURE 5(a), $Q(T^*)$ decreases monotonously with the increase of the initial failure size (q). Furthermore, compared with the CRM, $Q(T^*)$ of the ERM recoveries to high levels, and a larger $Q(T^*)$ meant a better network capability to recover. The total duration and recovery cost of the failure–recovery process are presented in FIGURE 5(c) and 5(e), respectively, where it can be seen that under the same model parameters, the ERM requires shorter recovery time and lower recovery cost to achieve a stable state than the CRM; these advantages are even more obvious as the value of γ increased. However, as shown in the subgraph in FIGURE 5(e), when q increases, the results were the opposite. We find that the larger γ lead to higher consumption of recovery resources when q is larger than the critical value, which could serve as a trigger to switch the model to achieve the cost–minimization. That is, when q is large than the critical value, the CRM is less costly than the ERM; otherwise, the ERM performs better than the CRM. Consequently, compared to the CRM, the ERM could recover to a state with a significantly high network capability in a short recovery time at a low recovery cost. The same results are observed for ER networks, as shown in FIGURE 5(b), 5(d), and 5(f).

C. ERM CORRELATION ANALYSIS

In the ERM, failed nodes with a greater node degree are repaired preferentially, and they are given larger node capacity and node load. The proposed recovery mechanism profoundly affects the failure–recovery process at both the micro level (node level) and the macro level (network level). From the micro–level perspective, more recovery resources are deployed at the early stages to alleviate the influence of failed nodes before a network collapse occurs. In this way, the risk of node failures, including both the initial failure of normal nodes and the secondary failure of repaired nodes, is significantly reduced. To validate this assumption, the secondary failure probability (P_{sf}) of the ERM is evaluated for the BA and ER networks. As shown in FIGURE 6(a) and 5(b), P_{sf} decreases monotonously with the increase in γ under different q values, and P_{sf} is almost zero when γ is large enough. This indicates that the ERM has an excellent performance to avoid the secondary failure, which cannot only decrease the consumption of recovery resources but also improve the network recovery efficiency. In addition, as shown in FIGURE 6(b), the distribution line at $q = 0.2$ significantly different from that at other q values. The reasons for this are explained below. As mentioned in the "Model" section, a time step corresponds to one failure–recovery–redistribution cycle. As shown in FIGURE 5(d), the highest peak of the end time $T*$ occurs near $q = 0.2$. This means that when $q = 0.2$, the network will experience a larger amount of recovery than under other q values. This allows the recovery model to exert its node repair function fully to avoid or postpone a network crash and thus avoid secondary failures efficiently. However, as shown in

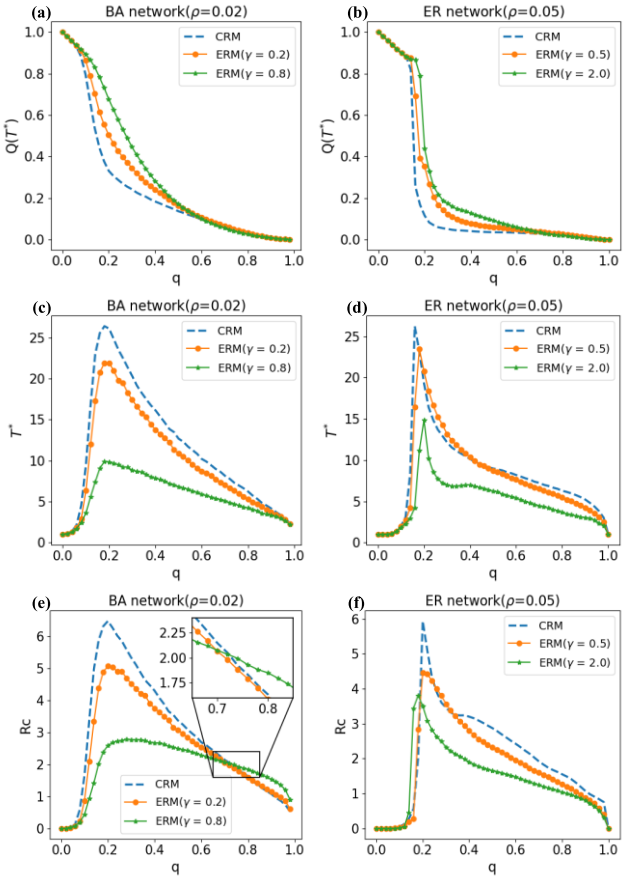

FIGURE 5. Diagrams of the BA and ER networks under different γ values: (a), (b) Q(T∗)–q plane; (c), (d) T∗–q plane; (e), (f) Rc–q plane; T∗ is the end time of the failure–recovery process; Q(T∗) denotes the network capability at time T∗; Rc is the recovery cost. The dashed blue line indicates the CRM; solid colored lines indicate the ERM.

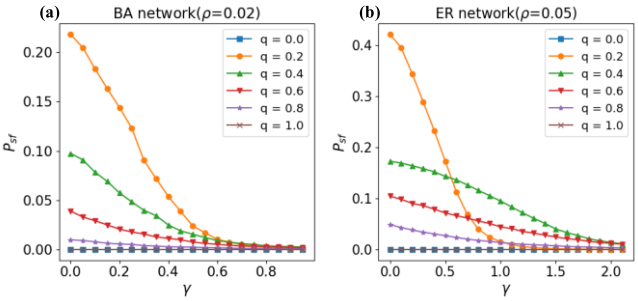

FIGURE 6. Secondary failure probability of the ERM: (a) the BA network; (b) the ER network.

FIGURE 5(b) and 5(d), when q is larger than 0.2, $Q(T*)$ and $T*$ decrease rapidly, which indicates that the network will completely collapse within a few steps, and the effect of recovery is no longer evident. Thus, as γ increases, the secondary failure probability P_{sf} at $q = 0.2$ decreases more rapidly than that at other q values.

From the macro–level perspective, increased load will flow to the repaired nodes with larger node capacity, contributing greatly to the heterogeneity of the load distribution. The load distributions of the BA and ER networks before and after the failure–recovery process are presented in FIGURE 7, where blue dots represent the load distribution before failure, and orange, green, and red dots represent the load distribution at different γ values after the failure–recovery process. The results show that the load is mostly concentrated in small fractions of nodes after the failure–recovery process, and the heterogeneity of the load distribution is more obvious as the value of γ increases. In Section 4.2, a strong positive correlation between the ERM performance and γ is demonstrated. Hence, it could be concluded that higher heterogeneity of the load distribution is effective in improving the network performance against failures. In addition, the results show that the BA network can recover to a high level due to the high heterogeneity of the load distribution.

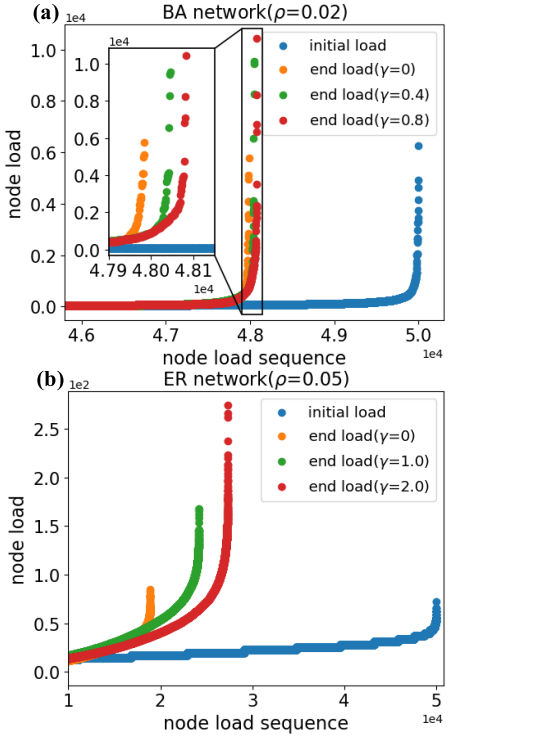

FIGURE 7. The load distribution before and after the failure–recovery process: (a) the BA network; (b) the ER network. The node sequence is arranged in ascending order of node load. Blue dots represent the load distribution before the failure occurs, and the orange, green, and red dots represent the load distributions at different γ values after the failure–recovery process.

D. NETWORK RESILIENCE ANALYSIS CASE STUDY: GNUTELLA NETWORK

The Internet offers a great case study for the exceeding recovery of networks. For the Internet, the mismatch between the network demand and the routing capacity caused by local load perturbations will result in function losses of a router [3]. To avoid further crowding phenomena, failed routers can resort to capacity expansion or backup facilities to ensure they function properly [22], [23]. In fact, these contingency mechanisms essentially aim at enhancing the capacity of a failed router and making it more reliable to allow more

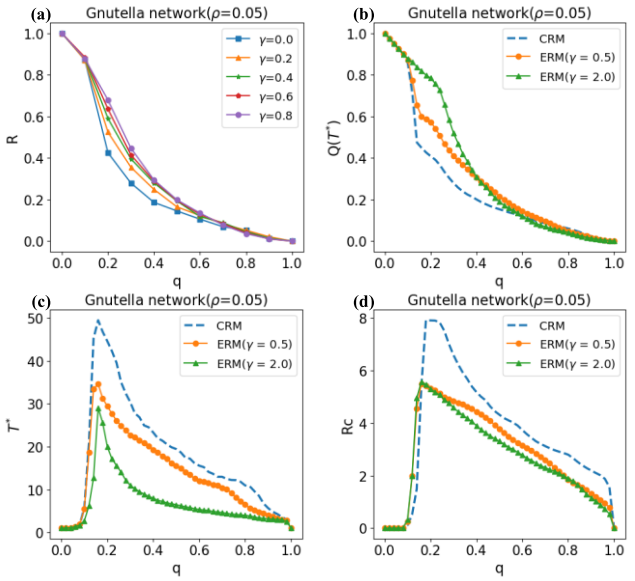

FIGURE 8. Performance of ERM in the Gnutella network. Network topology of the Gnutella network consisting of 6,301 nodes and 20,777 edges. (a) R–q plane; (b) Q(T∗)–q plane; (c) T∗–q plane; (d) Rc–q plane. The dashed blue line indicates the CRM; the solid colored lines indicate the ERM.

information to be routed through it. The proposed ERM model can capture the exceeding recovery mechanics existing on the Internet to a certain extent.

We test the performance of the ERM on the Gnutella network (http://snap.stanford.edu/data/#p2p), which is an actual Internet network based on the Gnutella routing protocol. In this work, nodes represent hosts in the Gnutella network topology, and edges denote connections between the Gnutella hosts. The Gnutella network contains 6,301 nodes and 20,777 edges. The correlation between R and q for different γ under a random attack is shown in FIGURE 8(a), where it can be seen that the network resilience decreases with the initial failure size, and it can be enhanced by increasing the exceeding recovery coefficient. In addition, the results in FIGURE 8(b)–8(d) are similar to those of the BA and ER networks; that is, the ERM can recover to a state with a significantly high network capability in a short recovery time at a low recovery cost.

V. CONCLUSION

This study proposes the ERM to capture the exceeding recovery mechanics of complex network systems. The proposed ERM model is verified by simulations for two network types. The simulation results show that for a given initial failure size q, there is a critical value of the exceeding recovery coefficient γ above which the network can restore to the normal operating state, but below this value, the network abruptly collapses. The results demonstrate that there are three regions in the network resilience phase diagram: a region where the network does not crash for any γ value, a region where the network restores to the normal state, and a region where even under the exceeding recovery, network collapse cannot be avoided. In addition, with the increase in γ, network resilience R gradually increases until reaching a saturation value. This indicates that the exceeding recovery mechanism can effectively improve the network resilience, thus making the network more resilient against cascading overload failures. The CRM, which represents the degenerated ERM model at the exceeding recovery coefficient of $\gamma = 0$, was used in the comparison test. The results indicate that compared to the CRM, the ERM can achieve high levels of functional recovery in a short recovery time and at a low recovery cost, and these advantages become even more obvious as the value of γ increases. The exceeding recovery mechanism profoundly affects the failure–recovery process, which is expressed through the reduced risk of a secondary failure at the micro level and improved heterogeneity of load distribution at the macro level. The lower secondary failure probability and higher heterogeneity of the load distribution are effective in improving the network resilience against cascading failures. In conclusion, the proposed ERM can be effective for repairing the damaged complex systems and provide a useful insight into designing a highly reliable network. Considering that the available recovery resources are finite in the real world, the optimal recovery strategies will be studied in future work.

REFERENCES

[1] X. Fu and Y. Yang, "Modeling and analysis of cascading node-link failures in multi-sink wireless sensor networks," *Rel. Eng. Syst. Saf.*, vol. 197, May 2020, Art. no. 106815, doi: 10.1016/j.ress.2020.106815.
[2] M. Packer, "What causes sudden death in patients with chronic heart failure and a reduced ejection fraction?" *Eur. Heart J.*, vol. 41, no. 18, pp. 1757–1763, May 2020, doi: 10.1093/eurheartj/ehz553.
[3] L. Xing, "Cascading failures in Internet of Things: Review and perspectives on reliability and resilience," *IEEE Internet Things J.*, vol. 8, no. 1, pp. 44–64, Jan. 2021, doi: 10.1109/JIOT.2020.3018687.
[4] 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, 2004, Art. no. 045104, doi: 10.1103/PhysRevE.69.045104.
[5] Y. Jin, Y. Chen, Z. Lu, Q. Zhang, and R. Kang, "Cascading failure modeling for circuit systems using impedance networks: A current-flow redistribution approach," *IEEE Trans. Ind. Electron.*, vol. 68, no. 1, pp. 632–641, Jan. 2021, doi: 10.1109/TIE.2020.2967672.
[6] Y. Wu, Z. Chen, X. Zhao, H. Gong, X. Su, and Y. Chen, "Propagation model of cascading failure based on discrete dynamical system," *Rel. Eng. Syst. Saf.*, vol. 209, May 2021, Art. no. 107424, doi: 10.1016/j.ress.2020.107424.
[7] M. A. Di Muro, C. E. La Rocca, H. E. Stanley, S. Havlin, and L. A. Braunstein, "Recovery of interdependent networks," *Sci. Rep.*, vol. 6, no. 1, p. 22834, Mar. 2016, doi: 10.1038/srep22834.
[8] A. Majdandzic, B. Podobnik, S. V. Buldyrev, D. Y. Kenett, S. Havlin, and H. E. Stanley, "Spontaneous recovery in dynamical networks," *Nature Phys.*, vol. 10, no. 1, pp. 34–38, Jan. 2014, doi: 10.1038/nphys2819.
[9] L. Böttcher, M. Luković, J. Nagler, S. Havlin, and H. J. Herrmann, "Failure and recovery in dynamical networks," *Sci. Rep.*, vol. 7, no. 1, pp. 1–9, Feb. 2017, doi: 10.1038/srep41729.
[10] T. Afrin and N. Yodo, "A concise survey of advancements in recovery strategies for resilient complex networks," *J. Complex Netw.*, vol. 7, no. 3, pp. 393–420, Jun. 2019, doi: 10.1093/comnet/cny025.
[11] W.-J. Jiang, R.-R. Liu, T.-L. Fan, S.-S. Liu, and L.-Y. Lü, "Overview of precaution and recovery strategies for cascading failures in multilayer networks," *Acta Phys. Sinica*, vol. 69, no. 8, 2020, Art. no. 088904, doi: 10.7498/aps.69.20192000.

[12] L. M. Shekhtman, M. M. Danziger, and S. Havlin, "Recent advances on failure and recovery in networks of networks," *Chaos, Solitons Fractals*, vol. 90, pp. 28–36, Sep. 2016, doi: 10.1016/j.chaos.2016.02.002.
[13] L. Buzna, K. Peters, H. Ammoser, C. Kühnert, and D. Helbing, "Efficient response to cascading disaster spreading," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 75, no. 5, May 2007, Art. no. 056107, doi: 10.1103/PhysRevE.75.056107.
[14] Y. Shang, "Localized recovery of complex networks against failure," *Sci. Rep.*, vol. 6, no. 1, p. 30521, Sep. 2016, doi: 10.1038/srep30521.
[15] F. Chaoqi, W. Ying, Z. Kun, and G. Yangjun, "Complex networks under dynamic repair model," *Phys. A, Stat. Mech. Appl.*, vol. 490, pp. 323–330, Jan. 2018, doi: 10.1016/j.physa.2017.08.071.
[16] C. Fu, Y. Wang, Y. Gao, and X. Wang, "Complex networks repair strategies: Dynamic models," *Phys. A, Stat. Mech. Appl.*, vol. 482, pp. 401–406, Sep. 2017, doi: 10.1016/j.physa.2017.04.118.
[17] W. Jia-Jian, G. Kai, W. Cong, and W. Lei, "Enhancing resilience of interdependent networks against cascading failures under preferential recovery strategies," *Acta Phys. Sinica*, vol. 67, no. 8, 2018, Art. no. 088901, doi: 10.7498/aps.67.20172526.
[18] C. E. La Rocca, H. E. Stanley, and L. A. Braunstein, "Strategy for stopping failure cascades in interdependent networks," *Phys. A, Stat. Mech. Appl.*, vol. 508, pp. 577–583, Oct. 2018, doi: 10.1016/j.physa.2018.05.154.
[19] G. Moutsinas and W. Guo, "Node-level resilience loss in dynamic complex networks," *Sci. Rep.*, vol. 10, no. 1, p. 3599, Feb. 2020, doi: 10.1038/s41598-020-60501-9.
[20] H. Sanhedrai, J. Gao, A. Bashan, M. Schwartz, S. Havlin, and B. Barzel, "Reviving a failed network through microscopic interventions," *Nature Phys.*, vol. 18, no. 3, pp. 338–349, Mar. 2022, doi: 10.1038/s41567-021-01474-y.
[21] S. Xu, Y. Xia, and M. Ouyang, "Effect of resource allocation to the recovery of scale-free networks during cascading failures," *Phys. A, Stat. Mech. Appl.*, vol. 540, Feb. 2020, Art. no. 123157, doi: 10.1016/j.physa.2019.123157.
[22] J. Xie, Y. Yuan, Z. Fan, J. Wang, J. Wu, and Y. Hu, "Eradicating abrupt collapse on single network with dependency groups," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 29, no. 8, Aug. 2019, Art. no. 083111, doi: 10.1063/1.5093077.
[23] X. Yuan, Y. Hu, H. E. Stanley, and S. Havlin, "Eradicating catastrophic collapse in interdependent networks via reinforced nodes," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 13, pp. 3311–3315, Mar. 2017, doi: 10.1073/pnas.1621369114.
[24] F. Avram, R. Adenane, and D. I. Ketcheson, "A review of matrix SIR arino epidemic models," *Mathematics*, vol. 9, no. 13, p. 1513, Jun. 2021, doi: 10.3390/math9131513.
[25] F. Broecker and K. Moelling, "Evolution of immune systems from viruses and transposable elements," *Frontiers Microbiol.*, vol. 10, p. 51, Jan. 2019, doi: 10.3389/fmicb.2019.00051.
[26] J.-W. Wang and L.-L. Rong, "Cascading failures in Barabási–Albert scale-free networks with a breakdown probability," *Int. J. Modern Phys. C*, vol. 20, no. 4, pp. 585–595, Apr. 2009, doi: 10.1142/S0129183109013819.
[27] Y. Wu, G. Hou, and S. Chen, "Post-earthquake resilience assessment and long-term restoration prioritization of transportation network," *Rel. Eng. Syst. Saf.*, vol. 211, Jul. 2021, Art. no. 107612, doi: 10.1016/j.ress.2021.107612.
[28] X. Fu, Y. Yang, and O. Postolache, "Invulnerability of clustering wireless sensor networks against cascading failures," *IEEE Syst. J.*, vol. 13, no. 2, pp. 1431–1442, Jun. 2019, doi: 10.1109/JSYST.2018.2849779.
[29] Z. Guo, Y. Wang, J. Zhong, C. Fu, Y. Sun, J. Li, Z. Chen, and G. Wen, "Effect of load-capacity heterogeneity on cascading overloads in networks," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 31, no. 12, Dec. 2021, Art. no. 123104, doi: 10.1063/5.0056152.
[30] J. Wang, L. Rong, L. Zhang, and Z. Zhang, "Attack vulnerability of scale-free networks due to cascading failures," *Phys. A, Statist. Mech. Appl.*, vol. 387, no. 26, pp. 6671–6678, Nov. 2008, doi: 10.1016/j.physa.2008.08.037.
[31] R. Albert, H. Jeong, and A.-L. Barabási, "The internet's achilles' heel: Error and attack tolerance of comple," *Nature*, vol. 406, pp. 378–382, Jan. 2000, doi: 10.1038/35019019.
[32] D. A. Reed, K. C. Kapur, and R. D. Christie, "Methodology for assessing the resilience of networked infrastructure," *IEEE Syst. J.*, vol. 3, no. 2, pp. 174–180, Jun. 2009, doi: 10.1109/JSYST.2009.2017396.
[33] J. Loscalzo and A. L. Barabási, *Network Science*. Cambridge, U.K.: Cambridge Univ. Press, 2016.
[34] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, pp. 378–382, Jul. 2000, doi: 10.1038/35019019.

JIE LI received the B.Sc. degree in electrical engineering from Air Force Engineering University, Xi'an, China, in 2017, where he is currently pursuing the Ph.D. degree with the Air Traffic Control and Navigation College. His current research interests include complex networks reliability modeling and complex systems.

YING WANG received the B.Sc. degree in system science from the China University of Science and Technology, in 1990, and the M.Sc. and Ph.D. degrees from Xi'an Jiaotong University, in 1995 and 2002, respectively. She is currently a Professor with the Air Traffic Control and Navigation College, Air Force Engineering University. Her research interests include complex systems, uncertainty theory, and airspace management.

JILONG ZHONG received the Ph.D. degree from Air Force Engineering University, Xi'an, China, in 2019. He is currently a Researcher with the National Institute of Defense Technology Innovation, PLA Academy of Military Science. His current research interests include complex networks, cyber–physical system reliability modeling, and cascading failure.