

Complexity

# Modeling and Quantification of Resilience in Complex Engineering Systems

Lead Guest Editor: Seyedmohsen Hosseini

Guest Editors: Riccardo Patriarca and Md Sarder





---

# **Modeling and Quantification of Resilience in Complex Engineering Systems**

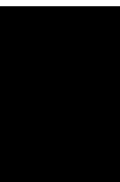
Complexity

---

## **Modeling and Quantification of Resilience in Complex Engineering Systems**

Lead Guest Editor: Seyedmohsen Hosseini

Guest Editors: Riccardo Patriarca and Md Sarder



---

Copyright © 2019 Hindawi Limited. All rights reserved.

This is a special issue published in "Complexity." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Chief Editor

Hiroki Sayama, USA

## Editorial Board

Oveis Abedinia, Kazakhstan  
José Ángel Acosta, Spain  
Carlos Aguilar-Ibanez, Mexico  
Mojtaba Ahmadiéh Khanesar, United Kingdom  
Tarek Ahmed-Ali, France  
Alex Alexandridis, Greece  
Basil M. Al-Hadithi, Spain  
Juan A. Almendral, Spain  
Diego R. Amancio, Brazil  
David Arroyo, Spain  
Mohamed Boutayeb, France  
Átila Bueno, Brazil  
Arturo Buscarino, Italy  
Ning Cai, China  
Guido Caldarelli, Italy  
Eric Campos, Mexico  
M. Chadli, France  
Émile J. L. Chappin, The Netherlands  
Yu-Wang Chen, United Kingdom  
Diyi Chen, China  
Giulio Cimini, Italy  
Danilo Comminiello, Italy  
Sergey Dashkovskiy, Germany  
Manlio De Domenico, Italy  
Pietro De Lellis, Italy  
Albert Diaz-Guilera, Spain  
Thach Ngoc Dinh, France  
Jordi Duch, Spain  
Marcio Eisencraft, Brazil  
Joshua Epstein, USA  
Mondher Farza, France  
Thierry Floquet, France  
José Manuel Galán, Spain  
Lucia Valentina Gambuzza, Italy  
Harish Garg, India  
Bernhard C. Geiger, Austria  
Carlos Gershenson, Mexico  
Peter Giesl, United Kingdom  
Sergio Gómez, Spain  
Lingzhong Guo, United Kingdom  
Xianggui Guo, China  
Sigurdur F. Hafstein, Iceland  
Chittaranjan Hens, India

Giacomo Innocenti, Italy  
Sarangapani Jagannathan, USA  
Mahdi Jalili, Australia  
Peng Ji, China  
Jeffrey H. Johnson, United Kingdom  
Mohammad Hassan Khooban, Denmark  
Abbas Khosravi, Australia  
Toshikazu Kuniya, Japan  
Vincent Labatut, France  
Lucas Lacasa, United Kingdom  
Guang Li, United Kingdom  
Qingdu Li, China  
Chongyang Liu, China  
Xinzhi Liu, Canada  
Xiaoping Liu, Canada  
Rosa M. Lopez Gutierrez, Mexico  
Vittorio Loreto, Italy  
Noureddine Manamanni, France  
Didier Maquin, France  
Eulalia Martínez, Spain  
Marcelo Messias, Brazil  
Ana Meštrović, Croatia  
Ludovico Minati, Japan  
Saleh Mobayen, Iran  
Christopher P. Monterola, Philippines  
Marcin Mrugalski, Poland  
Roberto Natella, Italy  
Sing Kiong Nguang, New Zealand  
Nam-Phong Nguyen, USA  
Irene Otero-Muras, Spain  
Yongping Pan, Singapore  
Daniela Paolotti, Italy  
Cornelio Posadas-Castillo, Mexico  
Mahardhika Pratama, Singapore  
Luis M. Rocha, USA  
Miguel Romance, Spain  
Avimanyu Sahoo, USA  
Matilde Santos, Spain  
Ramaswamy Savitha, Singapore  
Michele Scarpiniti, Italy  
Enzo Pasquale Scilingo, Italy  
Dan Selișteanu, Romania  
Dehua Shen, China  
Dimitrios Stamovlasis, Greece




---

Samuel Stanton, USA  
Roberto Tonelli, Italy  
Shahadat Uddin, Australia  
Gaetano Valenza, Italy  
Jose C. Valverde, Spain  
Alejandro F. Villaverde, Spain  
Dimitri Volchenkov, USA  
Christos Volos, Greece  
Zidong Wang, United Kingdom  
Qingling Wang, China  
Wenqin Wang, China  
Yan-Ling Wei, Singapore  
Honglei Xu, Australia  
Yong Xu, China  
Xinggang Yan, United Kingdom  
Zhile Yang, China  
Baris Yuce, United Kingdom  
Massimiliano Zanin, Spain  
Hassan Zargarzadeh, USA  
Rongqing Zhang, China  
Xianming Zhang, Australia  
Xiaopeng Zhao, USA  
Quanmin Zhu, United Kingdom

# Contents

## **Modeling and Quantification of Resilience in Complex Engineering Systems**

Riccardo Patriarca and Seyedmohsen Hosseini 



Editorial (2 pages), Article ID 1038908, Volume 2019 (2019)

## **Sink-Convergence Cascading Model for Wireless Sensor Networks with Different Load-Redistribution Schemes**

Xiuwen Fu , Haiqing Yao , and Yongsheng Yang 


Research Article (9 pages), Article ID 7630168, Volume 2019 (2019)

## **A Resilience Toolbox and Research Design for Black Sky Hazards to Power Grids**

Dmitry Borisoglebsky  and Liz Varga 

Research Article (15 pages), Article ID 1065419, Volume 2019 (2019)

## **Simulating Environmental Innovation Behavior of Private Enterprise with Innovation Subsidies**

Hongjun Guan, Zhen Zhang, Aiwu Zhao , and Shuang Guan

Research Article (12 pages), Article ID 4629457, Volume 2019 (2019)

## **Measuring Component Importance for Network System Using Cellular Automata**

Li He , Qiyao Cao , and Fengjun Shang 

Research Article (11 pages), Article ID 3971597, Volume 2019 (2019)

## **Metrics for Assessing Overall Performance of Inland Waterway Ports: A Bayesian Network Based Approach**

Niamat Ullah Ibne Hossain, Farjana Nur, Raed Jaradat, Seyedmohsen Hosseini , Mohammad Marufuzzaman, Stephen M. Puryear, and Randy K. Buchanan


Research Article (17 pages), Article ID 3518705, Volume 2019 (2019)

## **Stability and Complexity of a Novel Three-Dimensional Environmental Quality Dynamic Evolution System**

LiuWei Zhao  and Charles Oduro Acheampong Otoo




Research Article (11 pages), Article ID 3941920, Volume 2019 (2019)

## **Cascading Failures Analysis Considering Extreme Virus Propagation of Cyber-Physical Systems in Smart Grids**

Tao Wang , Xiaoguang Wei , Tao Huang , Jun Wang, Luis Valencia-Cabrera, Zhennan Fan , and Mario J. Pérez-Jiménez

Research Article (15 pages), Article ID 7428458, Volume 2019 (2019)

## **Identification of Two Vulnerability Features: A New Framework for Electrical Networks Based on the Load Redistribution Mechanism of Complex Networks**

Xiaoguang Wei, Shibin Gao , Tao Huang , Tao Wang , and Wenli Fan

Research Article (14 pages), Article ID 3531209, Volume 2019 (2019)

## Editorial

# Modeling and Quantification of Resilience in Complex Engineering Systems

**Riccardo Patriarca<sup>1</sup> and Seyedmohsen Hosseini<sup>2</sup>** 

<sup>1</sup>*Sapienza University of Rome, Dept. of Mechanical and Aerospace Engineering, Via Eudossiana, 18, 00184 Rome, Italy*

<sup>2</sup>*University of Southern Mississippi, Industrial Engineering Technology, Long Beach, MS 39560, USA*

Correspondence should be addressed to Seyedmohsen Hosseini; [mohsen.hosseini@usm.edu](mailto:mohsen.hosseini@usm.edu)

Received 26 August 2019; Accepted 29 August 2019; Published 10 December 2019

Copyright © 2019 Riccardo Patriarca and Seyedmohsen Hosseini. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The identification of criticalities is an important aspect for any engineering systems. Such identification plays a central role for safety and security analyses of modern socio-technical systems. Nowadays, electrical network as well as information and transportation networks can be enhanced in reliability, efficiency, safety and security through the adoption of automation solutions, and cyber-physical systems. In this regard, it becomes even more important to model and quantify the resilience of such systems with respect to a variety of disruptions. In response to this target, this special issue aims at providing a forum to present recent developments in terms of models and metrics for understanding, assessing, and enhancing system's resilience. Main challenges in the field include identification of vulnerability features, component importance measures, cascading failures modelling, environmental modelling, and overall performance assessment. Such aspects are managed through innovative approaches combining reliability engineering, network theory, social sciences, Bayesian network, genetic algorithms, and artificial intelligence.

This special issue has attracted high-quality submissions from scholars worldwide in the areas of resilience engineering, industrial engineering, business management, systems engineering, civil and environmental engineering, electrical engineering, energy, logistics, social sciences, and computer sciences. The researchers utilized their expertise and competences and match up to the challenges of developing solutions for managing the ever-increasing complexity of modern systems.

The total number of submissions is 16. After single-blind peer-review by at least two reviewers, 8 papers were finally accepted to be published. The acceptance rate is 50%. The average number of authors for each accepted paper is 4.1. The affiliated institutes of authors are from China (20 contributors), USA (8 contributors), and UK, Spain, and Italy (2 contributors each). These accepted papers can be organized into two major groups, which in reality are also inter-related.

The first group of papers is about resilience modeling. The paper titled "A Resilience Toolbox and Research Design for Black Sky Hazards to Power Grids" by D. Borisoglebsky and L. Varga presents a simulation model utilizing a resilience assessment equation for iteratively selecting the most appropriate tool for power grids. The paper titled "Sink-Convergence Cascading Model for Wireless Sensor Networks with Different Load-Redistribution Schemes" by X. Fu et al. proposes a realistic sink-convergence cascading model for wireless sensor networks with two load-redistribution schemes (i.e., idle redistribution and even redistribution). About cascading models, T. Wang et al. focus on cascading failures in interdependent systems from the perspective of cyber-physical security for smart grids in the paper titled "Cascading Failures Analysis Considering Extreme Virus Propagation of Cyber-Physical Systems in Smart Grids." The paper titled "Stability and Complexity of a Novel Three-Dimensional Environmental Quality Dynamic Evolution System" by L. W. Zhao and C. O. A. Otoo describes an environmental quality dynamic system based on Bayesian estimation and neural network to effectively



identify the system parameters for calibration of various variables and official data. The paper titled “Simulating Environmental Innovation Behavior of Private Enterprise with Innovation Subsidies” by H. Guan et al. adopts a social science perspective to model behaviors of private enterprises and simulate their evolution process in different market mechanisms, product competitions, and innovation subsidies.

The second group focuses more specifically on metric definition and quantification. In the transportation domain, the paper titled “Metrics for Assessing Overall Performance of Inland Waterway Ports: A Bayesian Network Based Approach” by N. U. I. Hossain et al. discusses a novel multidimensional metric to assess maritime port resilience. The paper titled “Identification of Two Vulnerability Features: A New Framework for Electrical Networks Based on the Load Redistribution Mechanism of Complex Networks” by X. Wei et al. describes a new framework to analyze two vulnerability features, impactability and susceptibility, in electrical networks by the perspective of load redistribution mechanisms. Lastly, the paper titled “Measuring Component Importance for Network System Using Cellular Automata” by L. He et al. concentrates on component importance measures of a network whose arc failure rates are not deterministic and imprecise.

In summary, the research papers cover a wide range of applications for resilience modelling in complex networks, as well as for metric definition able to support decision-making processes at different organizational levels. Based on the outcomes of this special issue, more research is still required to further progress the scientific field, especially considering more in detail the cyber-physical aspects, as well as combining the proposed technocentric approaches with a dominant sociotechnical perspective. In practical terms, such targets call for the development of efficient algorithms and frameworks to solve larger and more dynamic models.

### **Conflicts of Interest**

The editors declare that they have no conflicts of interest.

### **Acknowledgments**

The editors acknowledge the precious contribution of Md Sarder, who guest edited the issue together with them.

*Riccardo Patriarca  
Seyedmohsen Hosseini*

## Research Article

# Sink-Convergence Cascading Model for Wireless Sensor Networks with Different Load-Redistribution Schemes

Xiuwen Fu , Haiqing Yao , and Yongsheng Yang 

*Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai 201306, China*

Correspondence should be addressed to Xiuwen Fu; [fuxiuwen1987@163.com](mailto:fuxiuwen1987@163.com)

Received 27 December 2018; Revised 22 April 2019; Accepted 28 May 2019; Published 19 June 2019

Guest Editor: Riccardo Patriarca

Copyright © 2019 Xiuwen Fu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Existing cascading models are unable to depict the sink-convergence characteristic of WSNs (wireless sensor networks). In this work, we build a more realistic cascading model for WSNs, in which two load-redistribution schemes (i.e., idle redistribution and even redistribution) are introduced. In addition, failed nodes are allowed to recover after a certain time delay rather than being deleted from the network permanently. Simulation results show that the network invulnerability is positively correlated to the tolerance coefficient and negatively correlated to the exponential coefficient. Under the idle-redistribution scheme, the network can have stronger invulnerability against cascading failures. The extension of the recovery time will exacerbate the fluctuation of the cascading process.

## 1. Introduction

Wireless sensor network (WSN) is one of the most important components of the Internet of Things (IOT) system, because it has the characteristics of simple deployment, low cost, self-organization, and so on [1, 2]. In actual WSNs, sensor nodes are characterized by limited capacity. If the traffic load of a sensor node is greater than its capacity, its performance will be severely affected and all or part of its load will be rerouted to other sensor nodes, further leading to a redistribution of traffic load across the network. During this process, there may be new sensor nodes being failed due to overload. We call this dynamic process the cascading failures. In WSNs, due to the existence of cascading failures, even though most failures emerge very locally, the entire network can be largely affected or even collapsed globally [3–5].

Existing cascading models for WSNs usually used the degree or betweenness value of a sensor node to represent their traffic load. These assumptions are reasonable enough in the peer-to-peer networks, but they cannot apply to WSNs as they ignored the impacts of the sink node on network traffic distribution. Sink convergence is the most evident characteristic that can distinguish WSNs from other networks. Therefore, this paper proposes a more realistic cascading

model for WSNs. The main contributions of this paper are as follows:

- (1) A cascading model that can depict the sink-convergence characteristic of WSNs is proposed.
- (2) Two load-redistribution schemes (i.e., even-redistribution scheme and idle-redistribution scheme) are introduced.
- (3) We evaluate the impacts of key parameters in this model and compare two load-redistribution schemes.

The remainder of the paper is organized as follows. Section 2 describes recent related work. In Section 3, the cascading model is proposed. In Section 4, simulation results are given. Finally, conclusion and the future work are presented.

## 2. Related Work

In the real world, cascading failures are very common in actual network systems, such as power grid network, supply chain network, and communication network. Many researchers attempted to model the cascading process of actual networks [6]. Wang et al. [7] developed an under-load cascading model of supply chain networks, where each node is characterized by a capacity with upper and lower bounds. Rohden et al. [8] studied the cascading invulnerability of

TABLE 1: Summary of existing cascading models.

Models	Network types	Load Metrics	Models	Network types	Load Metrics
[7]	supply chain networks	degree	[14]	WSNs	betweenness
[8]	electricity grids	simulated current	[15]	WSNs	degree
[9]	communication networks	betweenness	[16]	WSNs	betweenness
[10]	interdependent networks	betweenness	[17]	WSNs	exponential degree
[11]	transmission networks	betweenness	[18]	WSNs	betweenness
[12]	cyber-physical systems	betweenness	[19]	WSNs	cluster degree
[13]	transportation networks	degree	[20]	WSNs	number of messages

electricity grids based on the alternating current model. Ren et al. [9] proposed a stochastic model to study the cascading dynamics in communication networks and identified the vital nodes from the perspective of network invulnerability. Chen et al. [10] investigated the cascading process in interdependent power grids and communication networks. Wu et al. [11] analyzed the impacts of link capacity on the cascading process in general transmission networks and found that a bifurcation point may exist in some cases which divides regions of opposite robustness behavior. Tu et al. [12] investigated the cascading invulnerability of cyberphysical systems and observed that two coupling networks have different sensitivity to the failure propagated from the other network. Candelieri et al. [13] investigated the cascading invulnerability of public transportation networks against directed attacks.

WSNs have also received a lot of attention in terms of cascading failures. Liu et al. [14] proposed a betweenness-oriented cascading model. In this model, the traffic of a sensor node is defined as its betweenness value. Yin et al. [15] studied the cascading process of scale-free WSNs and assumed that the traffic load of sensor nodes is correlated to their degrees. Li et al. [16] used the probability generation function to analyze the critical load of scale-free WSNs. In this work, the load is set to be closely correlated to the betweenness value. Ye et al. [17] proposed a fault-tolerant scheme to resist cascading failures in WSNs. They assumed that the load of sensor nodes is correlated to their degrees in an exponential way. Hu et al. [18] analyzed the cascading process of WSNs under random attacks based on the betweenness-load model. In [19], we presented a cascading model for hierarchical WSNs. In this model, the nodes' load is determined by its intercluster degree and its inner-cluster degree. In [20], we proposed a routing-based cascading model of WSNs in which the load of sensor nodes is defined as the real-time number of messages they carry.

Table 1 summarizes the mentioned cascading models. Although many cascading models have been proposed, they do not apply to realistic WSNs because they cannot reflect the sink-convergence characteristic of WSNs. A sample of sink convergence in WSNs is shown in Figure 1. In realistic WSNs, all the data packets collected by general sensor nodes will eventually be collected at the sink node and then be uploaded to the cloud; thus WSNs follow a typical many-to-one transmission paradigm, which makes them exhibit completely different traffic characteristics from other networks.

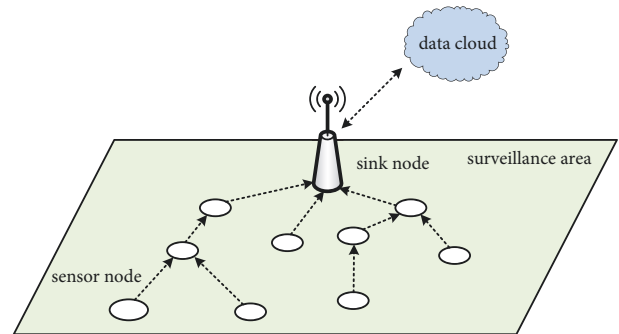


FIGURE 1: A sample of sink convergence in WSNs.

Therefore, it is necessary to develop a more realistic cascading model of WSNs.

### 3. Cascading Model

**3.1. Traffic Metric.** In [21], we have proposed a traffic metric “sink-oriented betweenness” to characterize the load distribution of WSNs. Its effectiveness and soundness have been verified through extensive simulations in [21]. Therefore, in this work we still use this traffic metric, as described below:

$$C_i(t) = \frac{\sum_{j \in V} g_{i,j}(t) / g_j(t)}{N}, \quad (1)$$

where  $g_{i,j}(t)$  is the number of the shortest paths from node  $j$  to the sink node passing through node  $i$  at time  $t$ .  $g_j(t)$  is the number of the shortest paths from node  $j$  to the sink node at time  $t$ .  $V$  and  $N$  are the set of sensor nodes and the total number of sensor nodes in the network, respectively.

**3.2. Load and Capacity.** As discussed in the last section, in actual WSNs, sensor nodes' initial load is correlated to the number of shortest paths from all the other sensor nodes to the sink node passing through it in the network, so it is reasonable to define the nodes' initial load as a function of the sink-oriented betweenness. For this consideration, we define the initial load of node  $i$  as

$$L_i(0) = C_i(0)^\alpha, \quad (2)$$

where  $\alpha \geq 0$  is the load-exponential coefficient that determines the distribution of the initial load. We can easily

observe that the initial load of each sensor node bears a linear relationship with its sink-oriented betweenness value when  $\alpha = 1$ . The configuration of  $\alpha$  is closely related to the data type of WSNs. If the data type is the multimedia data, it means that the initial load will have a rapid growth in an exponential way with the increase of  $C_i(0)$ ; thus  $\alpha$  should be set to a relatively large value. If the data type is the general text data,  $\alpha$  can be a small value. It is obvious that the introduction of  $\alpha$  can provide a high flexibility for our model to apply to different types of WSNs.

In most literature [14, 16], the nodes' capacity is set to be positively correlated to their initial load, as shown in

$$W_i = (1 + \beta) L_i(0), \quad (3)$$

where  $\beta$  is the overload-tolerance coefficient. However, in WSNs, this setting is far from the realistic situations. Unlike power grids in which the nodes' capacity can be customized according to the practical demands, the nodes' capacity in WSNs is always the same. This is partly because in most cases the hardware configurations of sensor nodes within the same WSN are always the same, and partly because it is impossible to customize the nodes' capacity when hundreds and even thousands of them are deployed. Therefore, in this work, the sensor nodes' capacity is defined as

$$W_N = (1 + \beta) L_N(0) = (1 + \beta) \frac{\sum_{i=1}^N L_i(0)}{N}. \quad (4)$$

According to (4), each sensor node has the same capacity, which is positively correlated to the average load of the initial network.

**3.3. Load-Redistribution Schemes.** In case that node  $i$  fails, its load will be distributed to other nodes in the network. There are two load-redistribution schemes: (1) even-redistribution scheme; (2) idle-redistribution scheme. The even-redistribution scheme is widely used in many cascading models. Under this scheme, the load originally taken by the failed node will be redistributed to its neighboring nodes. If node  $i$  fails at time  $t$ , its neighbor  $j$  can receive extra load  $\Delta_{ji}$  at time  $t + 1$  as follows:

$$\Delta_{ji}(t) = \frac{1}{N_i(t)} L_i(t), \quad (5)$$

where  $N_i(t)$  is the number of neighbors that node  $i$  has at time  $t$ . In some routings protocols of WSNs, sensor nodes do not have the real-time state information about their neighbors and they have the same capacity. It is reasonable to assign the load of the failed node to its neighbors evenly.

With the development of routing technologies in WSNs, in some routing protocols, sensor nodes can be congestion-aware, which means that they can own the real-time state information regarding their neighbors. On this basis, the idle-redistribution scheme is proposed. If node  $i$  fails at time  $t$ , its neighbor  $j$  can receive extra load  $\Delta_{ji}$  at time  $t + 1$  as follows:

$$\Delta_{ji}(t) = \frac{W_N - L_j(t)}{\sum_{k \in \Omega_i(t)} [W_N - L_k(t)]} L_i(t), \quad (6)$$

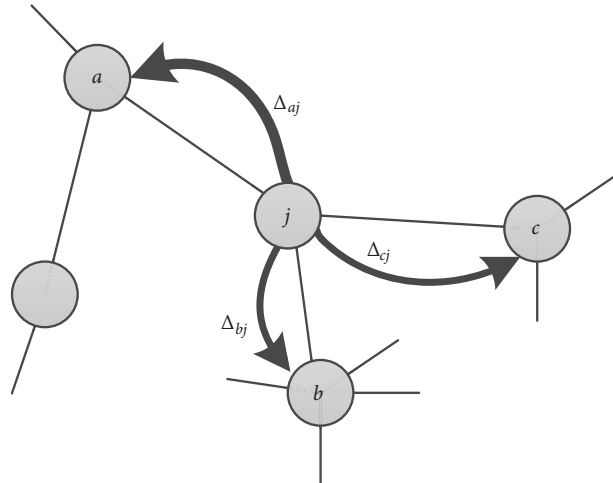


FIGURE 2: An example of load-redistribution process.

where  $\Omega_i(t)$  is the set consisting of the neighbors of node  $i$  at time  $t$ .  $W_N - L_k(t)$  is the idle capacity of node  $k$ , which can also be understood as the maximum load it can still receive. According to (6), we can easily find that, under the idle-redistribution scheme, the node with more idle capacity can be assigned more load from the failed node.

To illustrate the load-redistribution process more clearly, we present an example on a simplified network topology (shown in Figure 2). Assuming that node  $j$  fails at time  $t$ , the original load it takes will transfer to its neighbors  $a$ ,  $b$ , and  $c$  according to the load-redistribution scheme. At time  $t + 1$ , the real-time load of nodes  $a$ ,  $b$ , and  $c$  will be updated according to (7).

$$\begin{aligned} L_a(t+1) &= L_a(t) + \Delta_{aj} \\ L_b(t+1) &= L_b(t) + \Delta_{bj} \\ L_c(t+1) &= L_c(t) + \Delta_{cj}. \end{aligned} \quad (7)$$

If  $L_i(t+1) > W_N$ ,  $i \in \{a, b, c\}$ , another round of node failures will be triggered and the load of the newly failed node will transfer to its neighbors. This cascading process will not stop until the load of remaining nodes is within their capacity.

**3.4. Cascading Mechanism.** In most of the existing cascading models, sensor nodes have two states: normal and overloaded. According to their assumptions, if the node's load is beyond its capacity, then it will be removed from the network permanently. This assumption is reasonable in the network like power grids. However, in WSNs, this assumption is far from the fact. Different from the electricity overload in power grids, the overload of data packets in WSNs will not cause physical damage of sensor nodes. Overloaded nodes will reboot rather than fail permanently. When the reboot is completed, it will join the network again and function normally. Thus, in our model, the node at overloaded state will be given a recovery time  $\Delta t$ . Within  $\Delta t$ , this node cannot

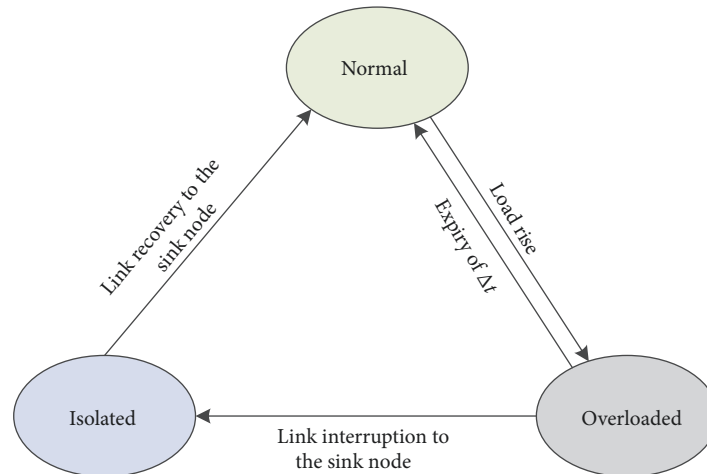


FIGURE 3: Cascading-state transition of sensor nodes.

receive, process, and transmit data packets. When  $\Delta t$  is expired, the node will become “normal” again. It is easy to understand that overloaded node can be recovered instantly when  $\Delta t$  approaches 0. Apparently, in this case, the damage caused by overload can be minimized. If  $\Delta t$  approaches  $\infty$ , our cascading scheme is equivalent to the conventional “permanent removal” scheme.

Sink convergence is the most evident feature that distinguishes WSNs from other wireless networks. In WSNs, if the link between a sensor node and the sink node is interrupted, the sensor node will be considered as an isolated node as its messaging service is not available. When cascading failures occur, some sensor nodes will become overloaded and the network connectivity will be severely impaired. During this process, some nodes will be isolated as their paths to the sink nodes are cut off. When some overloaded nodes are recovered via reboot, the network connectivity can be restored and isolated nodes will return to normal. The state transitions in the cascading mechanism is shown in Figure 3.

## 4. Analysis of the Invulnerability of WSNs

**4.1. Simulation Settings.** In the simulations, we set the network size to 300 and sensor nodes are randomly deployed in the simulation area. The wireless transmission radius of sensor nodes is set to 20m and the sink node is placed at the center of the simulation area. Figure 4 shows the network topology. In order to trigger cascading failures, we initially attack the first 10% of sensor nodes in the descending order of sink-oriented betweenness. Each node in the initial network before attack is at the normal state. We use survival ratio  $H_n(t)$  to measure the network invulnerability against cascading failures. As discussed in Section 3.4, normal nodes are the nodes that are not overloaded and can still maintain at least one effective path to the sink node.  $H_n(t)$  can be calculated by

$$H_n(t) = \frac{N(t)}{N(1 - q\%)}, \quad (8)$$

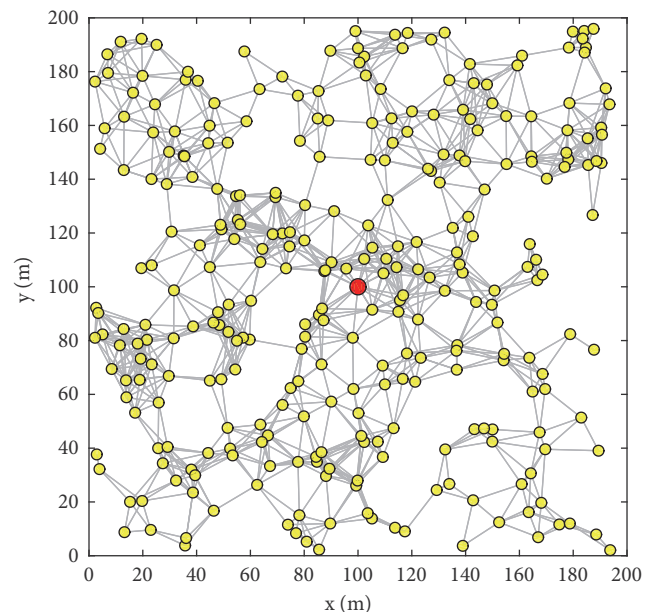


FIGURE 4: Network topology (300 sensor nodes are deployed and the sink node is marked in red).

where  $N(t)$  is the number of normal nodes at time  $t$ . Here we use  $H_n(\infty)$  to represent the survival ratio when the network reaches the steady state.

### 4.2. Simulation Results

**4.2.1. Verification of Sink-Convergence Characteristic.** The purpose of this experiment is to verify the sink-convergence characteristic of the proposed cascading model. In actual WSNs, since the sensor nodes around the sink node need to undertake more message-forwarding tasks, their load will be significantly higher than that of the nodes far from the sink

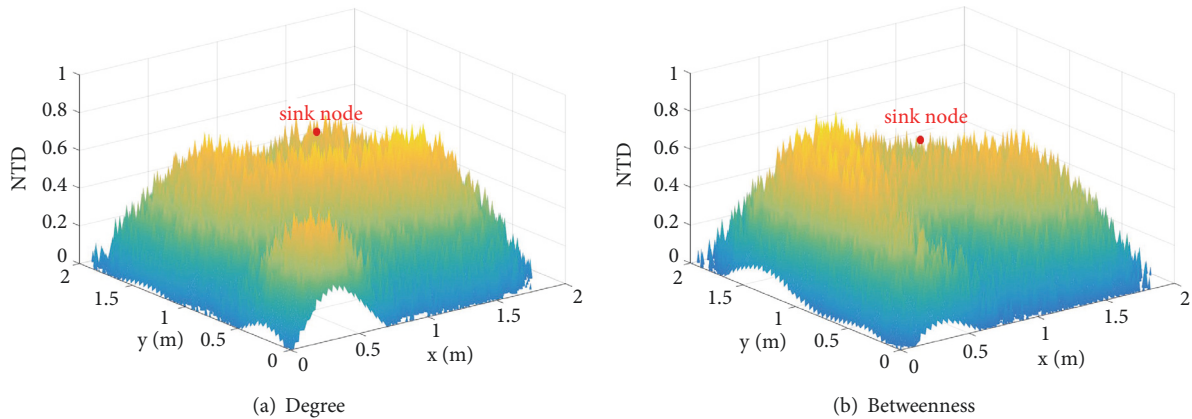


FIGURE 5: Network traffic distribution (NTD) generated by the degree-based and betweenness-based cascading models, respectively.

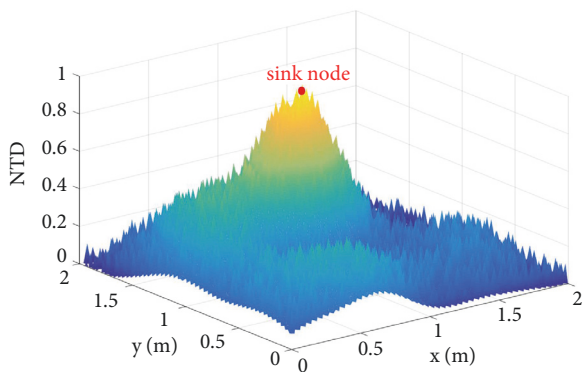


FIGURE 6: Network traffic distribution (NTD) generated by the proposed cascading model ( $\alpha = 1$ ).

node, a phenomenon described by many researchers as the “sink hole” [22–26]. Apparently, the sink hole phenomenon is an important indicator for judging whether the network is characterized by sink convergence.

Figure 5 shows the network traffic distribution created by the degree-based cascading model and the betweenness-based cascading model, respectively. It can easily be observed that there is no significant difference between the load of the nodes around the sink node and the load of the nodes in other areas, so the energy hole phenomenon is not so obvious. Figure 6 shows the network traffic distribution created by the proposed cascading model. We can easily observe a high-load peak around the sink node, so the sink-convergence characteristic is verified.

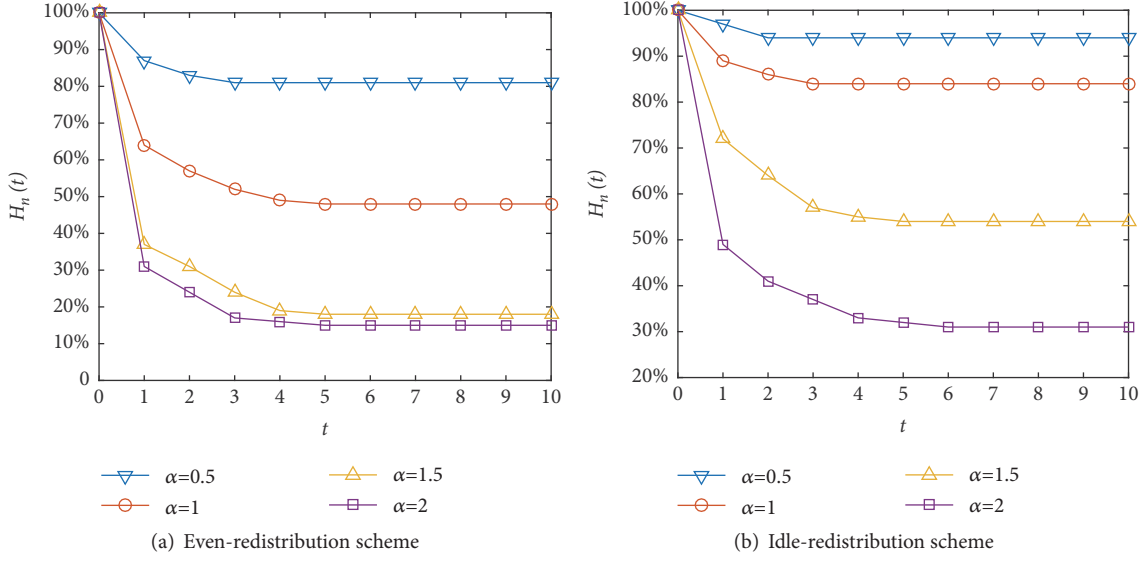
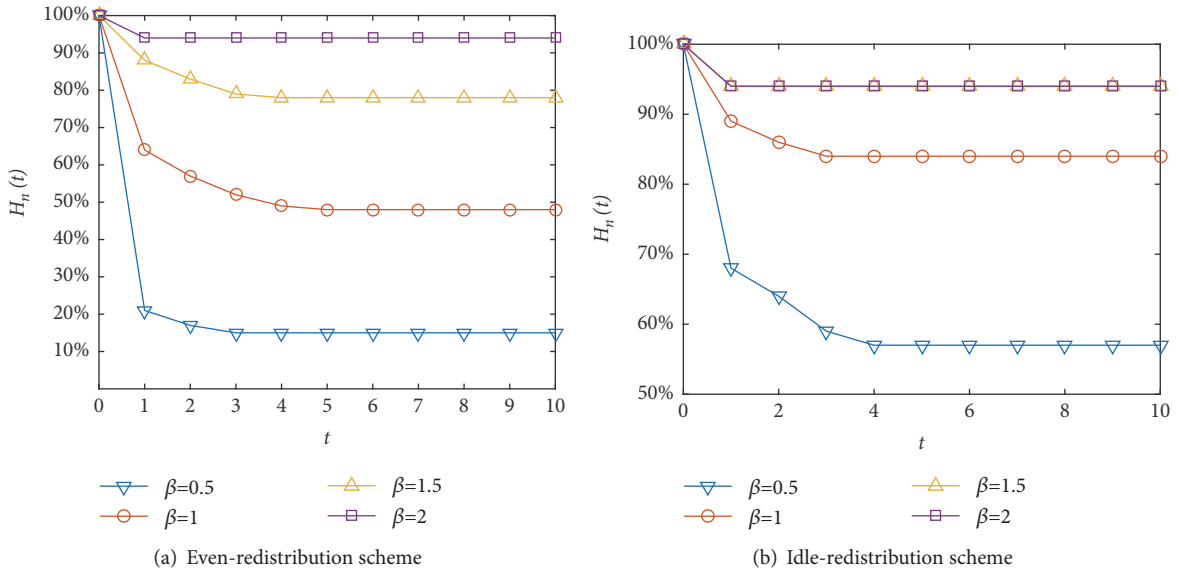
**4.2.2. Impacts of Modeling Parameters.** As is shown in Figure 7, we can easily find that, with the increase of exponential coefficient  $\alpha$ ,  $H_n(t)$  tends to decrease and the cascading process will reach the steady state faster. For example, under the even-redistribution scheme, when  $\alpha = 0.5$ ,  $H_n(t)$  will reach the steady state at 81% at time  $t = 3$ . When  $\alpha$  rises to 2,  $H_n(t)$  will stabilize at 17% at time  $t = 5$ . It is easy to

understand that the load taken by sensor nodes will increase much faster in an exponential way with the growth of  $\alpha$ , which will lead to a more evident gap between low-load nodes and high-load nodes. When the high-load nodes are attacked, the low-load nodes can hardly have enough capability to tackle the extra load transferred from failed high-load nodes. In our model, the configuration of  $\alpha$  is closely related to the data type of WSNs. The above simulation results tell us that for the high-volume data type, the risks and the damage brought by cascading failures will be much higher and the network designer should pay more attention to prevention of cascading failures.

Through the comparison between Figures 7(a) and 7(b), we can also find that the idle-redistribution scheme demonstrates a stronger invulnerability than the even-redistribution scheme when facing cascading failures. In the case that  $\alpha=1$ , under the even-redistribution scheme and idle-redistribution scheme,  $H_n(t)$  will stabilize at 47% and 83%, respectively. This is because under the idle-redistribution scheme, the idle capacity can be fully used, and thus more load can be tackled.

From Figure 8, we can find that the network invulnerability can be significantly improved with the increase of overload-tolerance coefficient  $\beta$ . In our model, a higher  $\beta$  means that sensor nodes can have more capacity to tackle load. Thus, there is surely a threshold  $\beta^*$  that can provide enough capacity for sensor nodes and can protect them from being overloaded. From Figure 8(b), we can find that the cascading process under  $\beta = 1.5$  and  $\beta = 2$  is totally the same. This phenomenon tells us that the threshold  $\beta^*$  should be within  $[1, 1.5]$ .

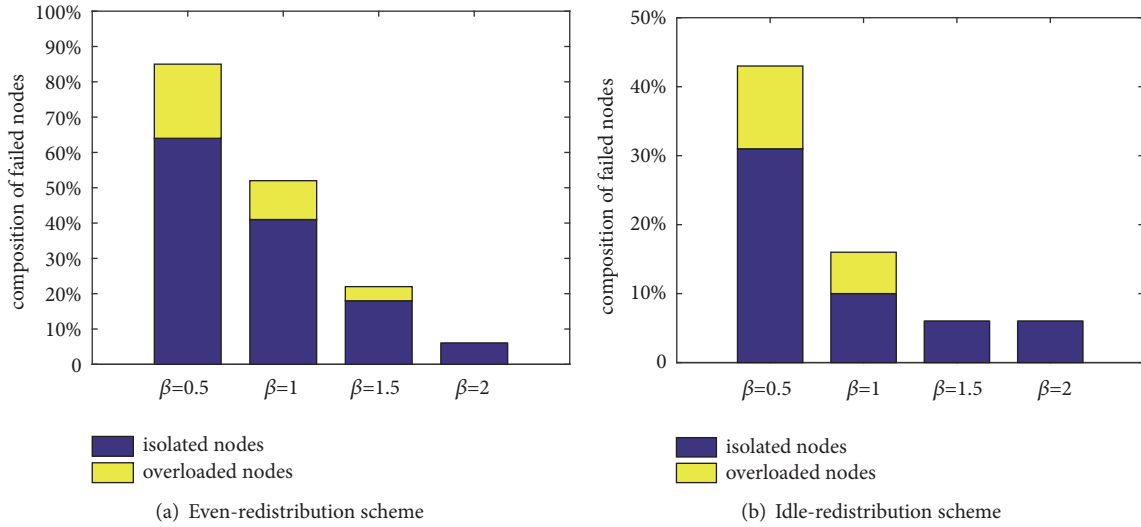
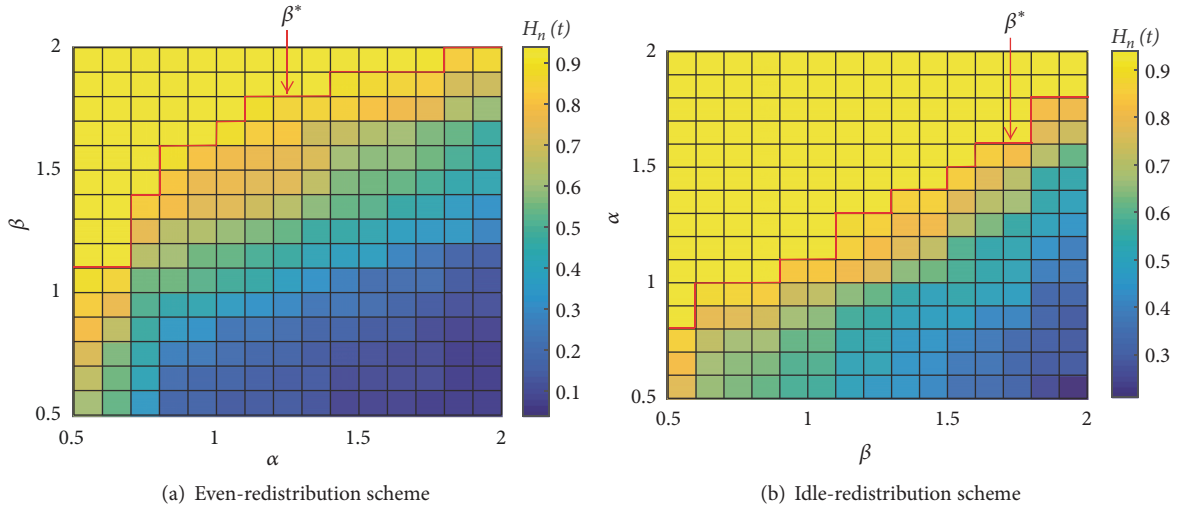
Figure 9 depicts the composition of failed nodes when the network reaches the steady state. We can clearly find that although neighboring nodes being overloaded constitute the major reason that makes nodes isolated, the majority of failed nodes are isolated nodes. Under the even-redistribution scheme, in the case that  $\beta = 0.5$ , isolated nodes are 63% and overloaded nodes are 21%. Moreover, with the increase of  $\beta$ , the ratio of overloaded nodes tends to be smaller. When  $\beta$  reaches a certain value, there will be no overloaded

FIGURE 7: Survival ratio with varying  $\alpha$  ( $\beta = 1, \Delta t = \infty$ ).FIGURE 8: Survival ratio with varying  $\beta$  ( $\alpha = 1, \Delta t = \infty$ ).

nodes in the network, which means the cascading process will not be triggered and the only damage is the isolated nodes caused by initial attacks. The existence of  $\beta^*$  is further verified.

As is shown in Figure 10, it can be easily observed that the threshold  $\beta^*$  will increase with the growth of  $\alpha$ , which means that more capacity resources are required to protect the network from cascading failures. Through comparison between Figures 10(a) and 10(b), we can find that at the same settings  $\beta^*$  will be smaller under the idle-redistribution scheme than under the even-redistribution scheme. The advantages of idle-redistribution scheme are further verified.

Figure 11 depicts the impacts of recovery time  $\Delta t$  on survival ratio  $H_n(t)$ . It can be easily observed that  $H_n(t)$  tends to fluctuate more wildly with the increase of  $\Delta t$ . In the case of  $\Delta t=1$ , when some sensor nodes are overloaded, on the one hand, they will redistribute their load and cause some other nodes to overload in the next time step and, on the other hand, they can recover from overload at the next time step due to the expiry of  $\Delta t$ . Therefore, we can find that at each time step after  $t=2$ , some nodes in the network fall into failure and some nodes return to normal, which makes  $H_n(t)$  demonstrate slight fluctuations with the network cascading process. Although in this case the damage caused by cascading failures can be minimized, it is actually hard to

FIGURE 9: Composition of failed nodes with varying  $\beta$  ( $\alpha = 1, \Delta t = \infty$ ).FIGURE 10: Heatmap of  $H_n(\infty)$  in the parameter space  $[\alpha, \beta]$  (threshold  $\beta^*$  is labelled by the red curve).

achieve because sensor nodes take time to reboot. When  $\Delta t$  increases to 2 or 3, overloaded nodes require more time for recovery, which will make cascading failures spread to a wider range and then lead to more obvious fluctuations of  $H_n(t)$ . When  $\Delta t = \infty$ , sensor nodes lose the recovery ability and  $H_n(t)$  decreases monotonically to a steady-state value.

## 5. Conclusions

In this paper, we developed a more realistic cascading model for WSNs. The most significant advantage of this model is that it can properly reflect the sink-convergence characteristic of WSNs. The simulation results show that (1) the network invulnerability is positively correlated to the overload-tolerance coefficient and negatively correlated to the load-exponential coefficient; (2) under the idle-redistribution scheme, the network can have stronger invulnerability against

cascading failures; and (3) the extension of the recovery time will exacerbate the fluctuation of the cascading process. These results provide us with some meaningful guidelines to build a more invulnerable WSN against cascading failures.

(1) The network with high-volume data type is more vulnerable to cascading failures.

(2) Due to the advantages of the idle-redistribution scheme, congestion-aware routing protocols can tackle more load, thus gaining stronger invulnerability against cascading failures.

In this work, we only discuss the cascading invulnerability of WSNs with deploying only one sink node. In recent years, multisink WSNs are becoming more and more widely used due to their advantages in energy efficiency and load balancing. Therefore, in our future work, we hope to upgrade the proposed model to adapt to the multisink WSNs, and on this basis, study its cascading invulnerability.



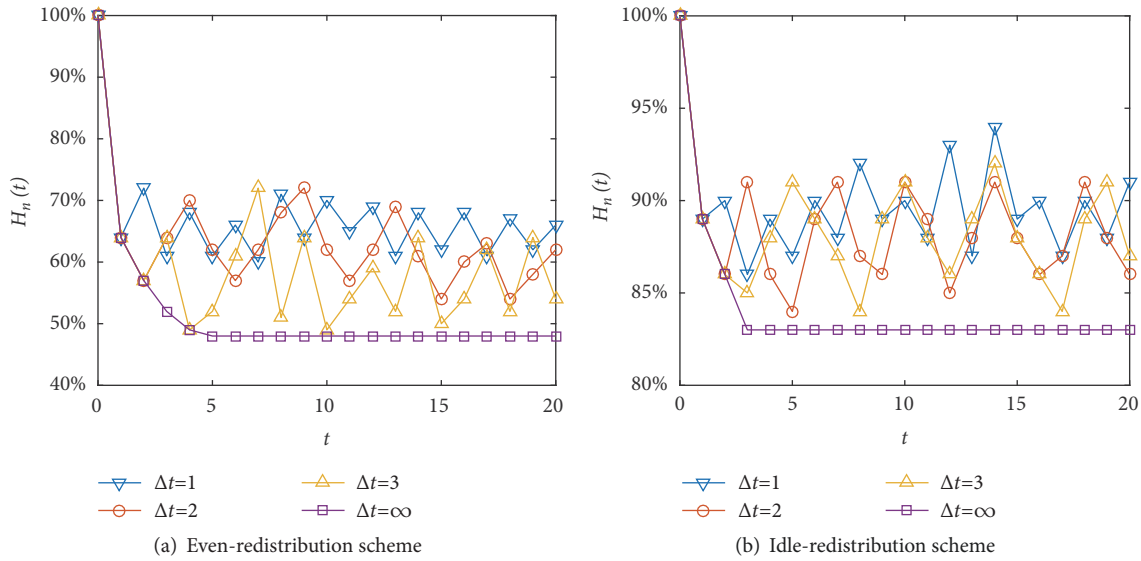


FIGURE 11: Survival ratio with varying  $\Delta t$  ( $\alpha = \beta = 1$ ).

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

## Acknowledgments

This work is supported in part by the National Natural Science Foundation of China (NSFC) under Grant No. 61571336.

## References

- [1] Y. Duan, X. Fu, W. Li, Y. Zhang, and G. Fortino, "Evolution of scale-free wireless sensor networks with feature of small-world networks," *Complexity*, vol. 2017, Article ID 2516742, 15 pages, 2017.
- [2] X. Fu, Y. Yang, and H. Yao, "Analysis on invulnerability of wireless sensor network towards cascading failures based on coupled map lattice," *Complexity*, vol. 2018, 14 pages, 2018.
- [3] X. Fu, H. Yao, O. Postolache, and Y. Yang, "Message forwarding for WSN-Assisted Opportunistic Network in disaster scenarios," *Journal of Network and Computer Applications*, vol. 137, pp. 11–24, 2019.
- [4] X. Fu, G. Fortino, and W. Li, "Environment-cognitive multipath routing protocol in wireless sensor networks," in *Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 2760–2765, IEEE, Miyazaki, Japan, October 2018.
- [5] G. Fortino, W. Russo, C. Savaglio, W. Shen, and M. Zhou, "Agent-oriented cooperative smart objects: from iot system design to implementation," *IEEE Transactions on Systems Man & Cybernetics Systems*, no. 99, pp. 1–18, 2017.
- [6] T. Wu, J. Wu, and W. You, "Optimizing robustness of complex networks with heterogeneous node functions based on the memetic algorithm," *Physica A: Statistical Mechanics and its Applications*, vol. 511, pp. 143–153, 2018.
- [7] Y. Wang and F. Zhang, "Modeling and analysis of under-load-based cascading failures in supply chain networks," *Nonlinear Dynamics*, vol. 92, no. 3, pp. 1403–1417, 2018.
- [8] M. Rohden, D. Jung, S. Tamrakar, and S. Kettemann, "Cascading failures in ac electricity grids," *Physical Review E*, vol. 94, no. 3, Article ID 032209, 2016.
- [9] W. Ren, J. Wu, X. Zhang, R. Lai, and L. Chen, "A stochastic model of cascading failure dynamics in communication networks," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 5, pp. 632–636, 2018.
- [10] Z. Chen, J. Wu, Y. Xia, and X. Zhang, "Robustness of interdependent power grids and communication networks: a complex network perspective," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 1, pp. 115–119, 2018.
- [11] J. Wu, X. Zhang, S. Ma, Z. Rong, and C. K. Tse, "Bifurcation in transmission networks under variation of link capacity," *International Journal of Bifurcation and Chaos*, vol. 28, no. 9, Article ID 1850109, 2018.
- [12] H. Tu, Y. Xia, J. Wu, and X. Zhou, "Robustness assessment of cyber-physical systems with weak interdependency," *Physica A: Statistical Mechanics and its Applications*, vol. 522, pp. 9–17, 2019.
- [13] A. Candelieri, B. G. Galuzzi, I. Giordani, and F. Archetti, "Vulnerability of public transportation networks against directed attacks and cascading failures," *Public Transport*, vol. 11, no. 1, pp. 1–23, 2019.
- [14] H. Liu, L. Zhao, R. Yin, X. Hao, Y. Li, and B. Liu, "A metric of topology fault-tolerance based on cascading failures for wireless sensor networks," *Journal of Information and Computational Science*, vol. 8, no. 14, pp. 3227–3237, 2011.
- [15] R. Yin, B. Liu, and Y. Li, "The critical load of scale-free fault-tolerant topology in wireless sensor networks for cascading

- failures,” *Physica A Statistical Mechanics and Its Applications*, vol. 409, no. 3, pp. 8–16, 2014.
- [16] Y. Li, R. Yin, and B. Liu, “Cascading failure research on scale-free fault tolerant topology in wireless sensor networks,” *Journal of Beijing University of Posts and Telecommunications*, vol. 37, no. 2, pp. 74–78, 2014.
- [17] Z. Ye, T. Wen, Z. Liu, X. Song, and C. Fu, “Fault-Tolerant scheme for cascading failure of scale-free wireless sensor networks,” in *Proceedings of the 2016 IEEE International Conference on Information and Automation, IEEE ICIA 2016*, pp. 2006–2011, China, August 2016.
- [18] X. Hu, W. Li, and X. Fu, “Analysis of cascading failure based on wireless sensor networks,” in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015*, pp. 1279–1284, Hong Kong, October 2015.
- [19] X. Fu, Y. Yang, and O. Postolache, “Invulnerability of clustering wireless sensor networks against cascading failures,” *IEEE Systems Journal*, no. 99, pp. 1–12, 2018.
- [20] X. Fu, H. Yao, and Y. Yang, “Modeling and analyzing cascading dynamics of the clustered wireless sensor network,” *Reliability Engineering and System Safety*, vol. 186, pp. 1–10, 2019.
- [21] X. Fu, H. Yao, and Y. Yang, “Modeling and analyzing the cascading invulnerability of wireless sensor networks,” *IEEE Sensors Journal*, vol. 19, no. 11, pp. 4349–4358, 2019.
- [22] C. Qiu, H. Shen, and K. Chen, “An energy-efficient and distributed cooperation mechanism for k-coverage hole detection and healing in WSNs,” *IEEE Transactions on Mobile Computing*, vol. 17, no. 6, pp. 1247–1259, 2018.
- [23] S. Jannu and P. K. Jana, “A grid based clustering and routing algorithm for solving hot spot problem in wireless sensor networks,” *Wireless Networks*, vol. 22, no. 6, pp. 1901–1916, 2016.
- [24] X. Deng, Z. Tang, L. T. Yang, M. Lin, and B. Wang, “Confident information coverage hole healing in hybrid industrial wireless sensor networks,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 2220–2229, 2018.
- [25] J. Ren, Y. Zhang, K. Zhang, A. Liu, J. Chen, and X. S. Shen, “Lifetime and energy hole evolution analysis in data-gathering wireless sensor networks,” *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 788–800, 2016.
- [26] H. Huang, H. Yin, G. Min, X. Zhang, W. Zhu, and Y. Wu, “Coordinate-assisted routing approach to bypass routing holes in wireless sensor networks,” *IEEE Communications Magazine*, vol. 55, no. 7, pp. 180–185, 2017.

## Research Article

# A Resilience Toolbox and Research Design for Black Sky Hazards to Power Grids

**Dmitry Borisoglebsky** <sup>1</sup> and **Liz Varga** <sup>2</sup>

<sup>1</sup>*School of Management, Cranfield University, Bedford MK43 0AL, UK*

<sup>2</sup>*Department of Civil, Environmental & Geomatic Engineering, UCL, London WC1E 6BT, UK*

Correspondence should be addressed to Liz Varga; [l.varga@ucl.ac.uk](mailto:l.varga@ucl.ac.uk)

Received 25 January 2019; Accepted 12 May 2019; Published 9 June 2019

Guest Editor: Seyedmohsen Hosseini

Copyright © 2019 Dmitry Borisoglebsky and Liz Varga. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A structured collection of tools for engineering resilience and a research approach to improve the resilience of a power grid are described in this paper. The collection is organized by a two-dimensional array formed from typologies of power grid components and business processes. These two dimensions provide physical and operational outlooks, respectively, for a power grid. The approach for resilience research is based on building a simulation model of a power grid which utilizes a resilience assessment equation to assess baseline resilience to a hazards' profile, then iteratively selects a subset of tools from the collection, and introduces these as interventions in the power grid simulation model. Calculating the difference in resilience associated with each subset supports multicriteria decision-making to find the most convenient subset of interventions for a power grid and hazards' profile. Resilience is an emergent quality of a power grid system, and therefore resilience research and interventions must be system-driven. This paper outlines further research required prior to the practical application of this approach.

## 1. Introduction

Power grids play an important role for modern society [1]. A failure in a power grid demonstrates a lack of engineered and engineering resilience to one or more hazards. A failure in a power grid may result in follow-up failures in the grid and other infrastructures [2]. Bo et al. [3] mapped and summarized 23 major blackouts from 1965 to 2012, representing major failures in power grids. Among the surveyed literature, no additional major blackouts were found, which is partially validated by the list of billion-dollar weather and climate disasters [4] in the US. Statistical studies show that major outages happen more often than can be concluded from statistics on minor and intermediate outages [5, 6].

Therefore, according to theoretical distributions and history of disasters, the world will experience major outages in the future. The probability of an outage in a specific power grid may be reduced by the application of resilience research to improve the resilience of power grid infrastructures.

This paper describes an approach for resilience research to improve the resilience of a power grid. It is based on

mapping existing tools for resilience enhancement in a matrix-based classification for the follow-up targeted resilience research. This paper is limited by the tools arising in the literature review and so does not provide the complete list of existing tools. The literature does not provide a way to rank the tools for their effectiveness so no prioritization is available.

## 2. Methods

Literature review was performed to identify academic articles according to two schemas: functional and summary. The functional schema facilitated the search for tools for resilience enhancement, definitions of resilience, and experience of blackouts. The summary schema enabled the identification of summaries and reviews of resilience assessment frameworks.

Search strings for advanced search in Scopus database are shown in Table 1. The primary focus of this article is to identify and classify the tools for resilience enhancement. Other search topics provide context.

TABLE 1: Strings for search queries run in Scopus database.

Schema	Topic	Query
Functional	Resilience	TITLE ( ( resilien* ) AND ( electric* OR power* ) ) OR KEY ( ( resilien* ) AND ( electric* OR power* ) )
Functional	Blackout	TITLE ( ( blackout* OR outage* ) AND ( electric* OR power* ) ) OR KEY ( ( blackout* OR outage* ) AND ( electric* OR power* ) )
Summary	Framework	TITLE ( resilien* AND ( measur* OR assess* OR indicat* OR quant* OR metric* ) AND ( review OR overview OR survey OR state ) )

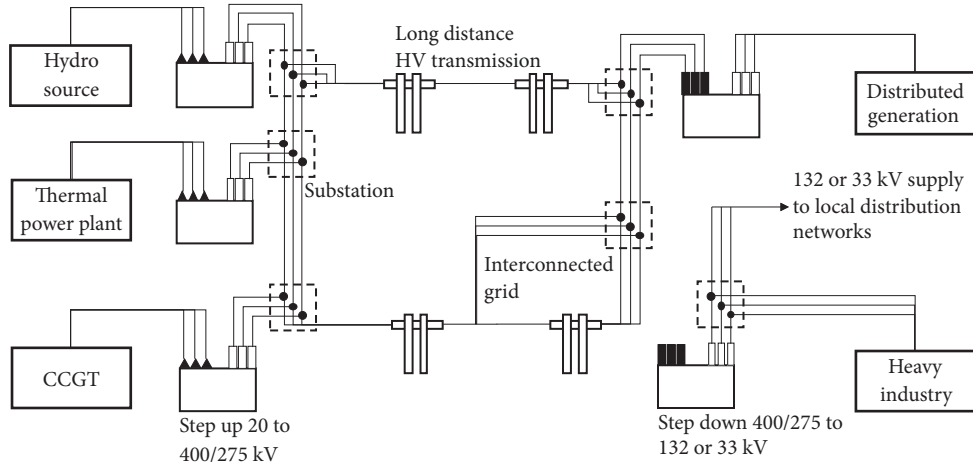


FIGURE 1: Generation and transmission network (UK voltages and practice) [7].

The functional review assumes that the search strings provide source papers for an incomplete yet sufficient study coverage of the topics. The functional search queries selected source papers based either on matching titles (TITLE) or on keyword values (KEY). SCOPUS allows an advanced search producing a search engine results page with a set of tools for filtering on the left panel. Using the database's filtering capabilities, the initial list of results was cleared of non-English (e.g., German or Chinese), non-article (a peer-review is expected for most articles), and irrelevant topics (e.g., medicine or biology). The toolbox cites 54 papers, reduced from 174 papers.

The summaries' review assumes that a reasonably complete review of resilience assessment frameworks, for the purpose of this research, is possible via study of existing review papers on this topic. Papers for review were selected by the relevance of their titles and abstracts.

### 3. Results

**3.1. Power Grid.** A power grid is a system that produces, transports, and consumes electrical energy. Generators convert fuels and other energy sources into electrical power. Step-up and step-down transformation stations border transmission lines with substations in between. The UK power grid uses several voltages with 400/275 kV in the transmission grid and 132/33 kV in distribution grids, similar architecture in other countries. Accompanied by visualization (see Figure 1), components of a power grid are listed below:

- (i) *Producer* is a part of a power grid that produces electrical energy regardless of the amount of energy or stability of production. A hydropower plant, fossil-fuel plant, wind generator farm, and a microgrid are examples of producers of electricity.
- (ii) *Step-up/down substation* is required to transform current, e.g., from a producer's 20 kV to 400/275 kV of transition grid and from 400/275 kV to 132/33 kV of distribution grid.
- (iii) *Power line* is an overhead or underground electricity transmission line.
- (iv) *Substation* is a station that transitions and controls the power flow.
- (v) *Consumer* is a part of a power grid that consumes electrical energy regardless of the amount of energy or stability of consumption. A factory, a distribution grid, and a microgrid are treated as consumers of electricity.
- (vi) *Control* is a hardware, software, or organizational part of a utility company. While control is not shown in Figure 1, it is a critical part of a power grid.

**3.2. Hazards.** Mukherjee et al. [9] analyzed billion-dollar blackouts in the US between 2000 and 2016, while Bie et al. [10] provided statistics on of the major blackouts in the world. While the lists of causes are slightly different, there is a major difference in the distributions of blackouts by causes between these two regions (see Table 2). For both the

TABLE 2: Causes of major blackouts in the USA and the world.

USA		The world	
severe weather	52.9%	equipment failure	47.8%
intentional attack	22.9%	natural disasters	30.7%
system operability disruption	10.3%	malfunctions/ miscellaneous	10.1%
public appeal	4.2%	vandalism	5.7%
equipment failure	4.0%	supply shortage	4.3%
fuel supply emergency	3.1%	cyber attack	1.4%
islanding	2.6%		

US and world the first three causes are responsible for over 85% of blackouts, 86.9%, and 88.67% respectively. Additional research is required to explain the difference in causes and magnitudes.

Type and intensity are the basic characteristics of hazards when considering the resilience of engineered systems such as a power grid. Type, e.g., precipitation, earthquake, and intensity, e.g., 300mm in 24 hours, 4 Richter scale, differ greatly but can have similar effects, e.g., outage of a power grid component. The impact of a hazard is at least infrastructure-, technology-, and asset-specific. Hazards and their likelihoods are specific to geographic regions, so despite both being island nations, the likelihood of earthquakes is much smaller in the UK than in Japan. A country-specific hazards' profile must be addressed during resilience research and for engineering the resilience of a national power grid. For example, *National risk register of civil emergencies, 2017 edition* [11] contains a list of hazards for the UK. An analysis of global risks [12] is annually produced by the World Economic Forum, Black Sky hazards [13] were defined by Electric Infrastructure Security Council, and academia also produces lists of hazards [14].

**3.3. Resilience.** In 2015, Fisher [15] mentioned the existence of 70 definitions for resilience, although a list and analysis of definitions were absent. 42 instances of definitions for resilience were collected during literature review. These were grouped by similarity and the group selected for this black sky study was the one to include definitions from governmental or international organizations likely expressing the need for strategic decision-making, such as the US White House [16], the UK Cabinet Office [17], or the UN Office for Disaster Risk Reduction (UNISDR) [18].

The definition by UNISDR [18] is selected for this paper, *resilience is the ability of a system, community, or society exposed to hazards to resist, absorb, accommodate to, and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions.* The intended meaning is expressed in the analysis below.

This definition lists *abilities*, e.g., *resist, recover*; however it does not provide a rationale for the completeness of this list. This definition states that this ability belongs to an *entity*, e.g., *system*; we treat the power grid infrastructure as a 'system'. This definition states that that resilience is a *function* of hazards. This definition also states that resilience is *conditional*, e.g., *exposed to hazards, effects of a hazard,*

and suggests either a secondary condition or an insight how resilience should be *measured*, via *timely and efficient manner*. Resilience is *manifested* via *preservation and restoration of essential basic structures and functions*, noting that not all the entity's elements need to be restored

From the literature, the following assumptions, abstractions, and explications were made for this paper, namely,

- (i) A system has a normal and a disrupted state; a hazard changes a normal state of a system to a disrupted state. Resilience is the ability of an entity to execute the opposite transition, namely, from a disrupted state to a normal state, or to prevent a transition into a disrupted state.
- (ii) A shorter time period in a disrupted state indicates a more resilient system, as is the ability to make a faster transition from a disrupted state to a normal state. Also, limited disruption to critical parts of the entity and fast recovery of critical parts reflects greater resilience to the scale of disruption caused by hazards.
- (iii) Infrastructure is a complex system constructed from many alternative assets utilizing different technologies that work on different physical principles. For the purposes of this paper we assume a predetermined power grid configuration. Therefore, the manifestation of resilience is the preservation or recovery of basic functions. Basic functions are available to the customers in volume through time in locations.
- (iv) Time, volume and location are possible input data for resilience assessment although location is mostly disregarded in this paper as it is focused on resilience of a whole infrastructure instead of its parts and the lack of theoretical means to address physical and likely multilayered virtual subgrids.

In a normal state, an infrastructure delivers a normal volume of basic functions, while an infrastructure in a disrupted state, in comparison to expected delivery in a normal state, delivers a reduced volume. A normal output is delivered by an infrastructure working in normal conditions, while the same infrastructure under a pressure of a hazard delivers a smaller volume. The nearer the volume to the normal output volume under pressure by a hazard, the greater the resilience, and the smaller the difference between the outputs; see Figure 2.

**3.4. Resilience Assessment.** Hosseini et al. [19] provide a classification of approaches to resilience assessment. The

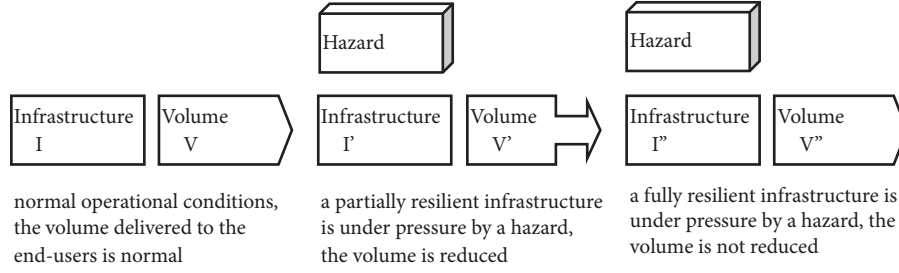


FIGURE 2: Three simplified hypothetical scenarios on the delivered volume in normal and disrupted states.

TABLE 3: Mathematical and conceptual representation of volume, time, and volume-over-time frameworks. ‘R’ means resilience value, ‘f’ and ‘ρ’ mean failure and recovery ratio, ‘v’ and ‘t’ mean volume and time, and ‘D(t)’ and ‘U(t)’ mean disturbed and undisturbed delivery value at a time; encodings for subscripts, ‘d’ for disaster, ‘e’ for the lowest delivery, ‘n’ for normalization, and ‘o’ and ‘c’ mostly technical indicators of the opening and closing the chart.

Branch	Equation	Conceptualization
Volume	$R = \frac{v_n - v_e}{v_n}$	$R = \frac{\text{impact on performance}}{\text{baseline performance}}$
Time	$R = \frac{t_d - t_o + f(t_e - t_d) + \rho(t_n - t_e)}{t_d - t_o + (t_e - t_d) + (t_n - t_e)}$ $f = \frac{\int_{t_d}^{t_e} D(t)dt}{\int_{t_d}^{t_e} U(t)dt}, \quad \rho = \frac{\int_{t_e}^{t_n} D(t)dt}{\int_{t_e}^{t_n} U(t)dt}$	$R = \frac{\text{weighted time period}}{\text{time period}}$
Volume over time	$R = \frac{\int_{t_d}^{t_n} D(t)dt}{\int_{t_d}^{t_n} U(t)dt}$	$R = \frac{\text{disturbed deliverery}}{\text{undisturbed deliverery}}$

general measure approach is used in this study because it is quantitative and generic. Yodo and Wang [20] listed quantitative resilience assessment frameworks for engineering systems, and this is used as the initial source of ‘general measure’ approaches.

A system produces output over time, which can be visualized via an area chart with volume on Y axis and time on X axis. This generates three main branches of resilience assessment, namely, (i) volume-focused that utilizes Y axis, (ii) time-focused that utilizes X axis, and (iii) volume-over-time-focused that utilizes X and Y axes. Most frameworks in Yodo and Wang [20] paper belong to this category.

Quantitative resilience assessment frameworks use mathematical equations. For the ease of understanding the idea behind a branch and an equation, representative examples were selected from Yodo and Wang’s paper. Volume-focused framework is referring to [21], time-focused is referring to [22], and volume-over-time is referring to [23]. The corresponding mathematical equations are listed in Table 3, while Figure 3 provides a supplementary visualization.

**3.5. Business Processes.** One or more organizations maintain a power grid and stable supply of electric energy. Just as a power grid may be described by components (see Figure 1, [7]), an organization may be described by business processes (see Figure 4, [24, 25]). The business processes of an organization influence the abilities of the organization to maintain the resilience of their products, in particular the operate business processes that focus on adding value to the customer. If

operate processes are able to ensure that the power grid (the engineered system) is resilient, then the capabilities of people, technology, policies, and systems (the engineering system) that underpin these processes are resilient. Short descriptions [26] of business processes are provided below.

*Manage* processes consist of a subset of processes formulating organizations’ vision and mission, methods to achieve those, and methods for effective utilization of resources. *Set directions* focus on formulating the vision and mission of the organization. *Formulate strategies* focus on identification and definition of strategies to reach the vision and mission. *Direct business* focuses on use of resources in actions to reach the vision and mission.

*Operate* processes consist of a subset of processes that deliver value to the customer. *Develop product* focuses on designing the product or service that would add value to the customer.

*Get order* focuses on finding customers and setting contracts with the customers. *Fulfil order* focuses on fulfilling the demand of the market. *Support product* process focuses on increasing value of the product or service.

*Support* processes consist of a subset of processes ensuring that operate processes are running. *Manage finance* focuses on having a sufficient cash flow. *Support personnel* focuses on ensuring sufficient human resources for operate and other processes. *Manage technology* focuses on creating an environment without bottlenecks due to a lack of technology.

*Corporate learning* focuses on increasing the quality of human resources.

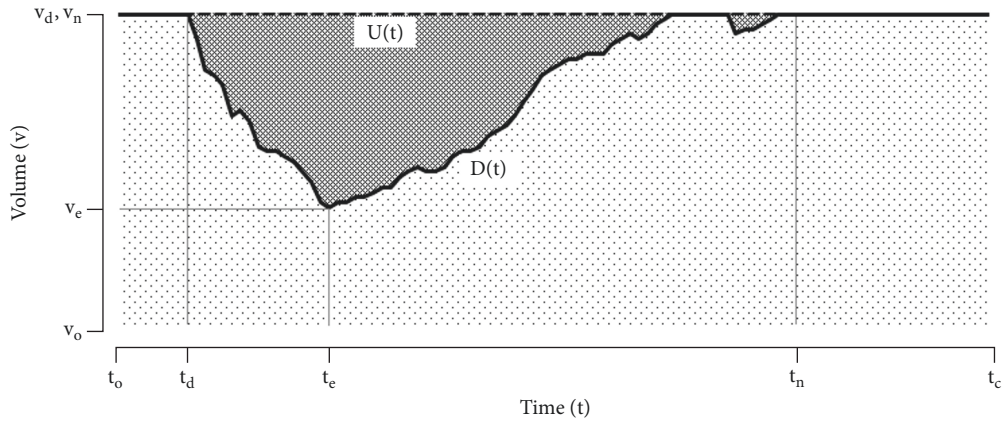


FIGURE 3: A disaster decreases the volume of delivery over time with the dark gray highlighting the impact of ‘under-delivery’ area, the solid black line (D) indicating the actual volume, and the dashed black line (U) indicating that the expected delivery would not make the disaster happen; encodings for subscripts, ‘d’ for disaster, ‘e’ for the lowest delivery, ‘n’ for normalization, and ‘o’ and ‘c’ mostly technical indicators of the opening and closing the chart.

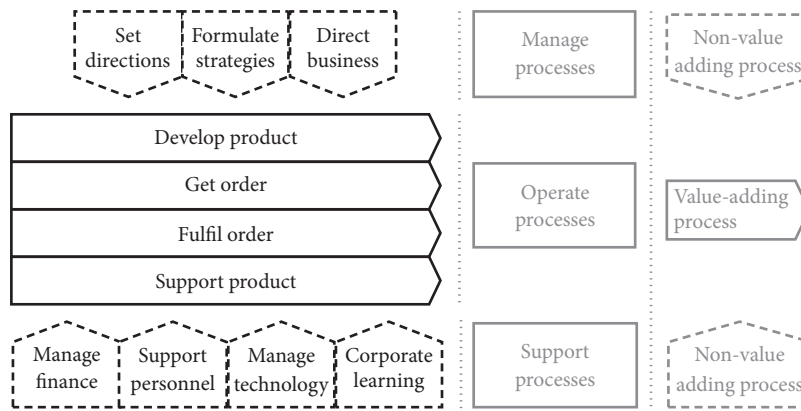


FIGURE 4: Business process architecture, redrawn from Bititci *et al.* [24, 25].

The criticality of the different classes of business process is not addressed; however, ‘value-adding processes’ have direct influence on operations. However some non-value adding processes, for example, *manage technology*, may have a direct bearing on the ability of *operate processes* to function.

3.6. *Toolbox*. Components of a power grid (see Figure 1) and business processes of an organization operating a power grid (see Figure 4) are sufficiently descriptive for high-level resilience modelling of a grid under operation. Together, these categories form a two-dimensional conceptual space. Tools to improve the resilience of a power grid may be assigned to relevant cells in the two-dimensional grid based on which component of the power grid and which business process are targeted by the tool. The tools identified in the literature review have been reviewed and allocated to the grid (see Table 4). A dash in a cell indicates that there is the absence of tools to improve resilience associated with the business process (row) and power grid component (column). Gaps may present opportunities for new tools to improve resilience or may indicate areas where resilience is either not a priority or has not yet been addressed.

Tools for improving resilience of a power grid are listed and briefly described below. These tools were suggested or mentioned in journal papers reviewed for this research. Names of these tools are taken from the papers with minor changes where needed to improve clarity. Tools are alphabetically ordered. The core idea of each tool is taken from its respective paper, and a short description is adjusted for the current paper or created when none is explicitly given. Each tool is described according to the following template: *Name (abbreviation) [source] Short description. Statement on categorization.*

Some of the tools are mentioned in multiple papers, yet usually a single citation is provided below indicating a reasonable reference. Because of the specifics of the literature review, it is assumed that the list of tools is sufficiently representative, yet it cannot be considered complete. Some tools are dependent on or are alternatives to other tools. This has not been addressed in the respective short descriptions. Another research may allocate tools into different categories, as a valid deductive or cogent inductive logical and ontological foundation have not been found nor created, and the current allocation is intuitive.

TABLE 4: Index of tools for resilience enhancement of a power grid; tools are referred as abbreviations of their names.

Process	Component					
	Producer	Substation	Power line	Substation	Consumer	Control
Set directions	-	-	-	-	-	-
Formulate strategies	-	RAT, RET	RAT, RET	RAT, RET	-	CP, OR, RAT, RET
Direct business	IB	-	-	-	-	CP, DA, DP, DE, DMG, MAF, OM, PI, RSA, RM, RP
Get order	-	-	-	-	-	-
Develop product	-	-	-	-	AC	-
Fulfil order	DIG, GG, MEG, MG, PP, RCUC, TESS	MUT	-	-	BP, DSM, LS, MG	CS, CI, DSM, DIA, LS, MG, NR, pMEG, PP, PSS, RCUC, RSA, SG, SM
Support product	-	-	-	-	AC	-
Manage finance	-	-	-	-	-	-
Support personnel	-	-	-	-	-	CRS
Manage technology	GG, MG, RS	RAT, RET, RS, R&D	MV, OS, RAT, RET, RS, UL	MO, OS, RAT, RET, RS	MG, RS	DA, DP, DE, MS, MG, MO, PMU, PI, PSS, RAT, RET, ST, SCADA, SG, SD, SE
Corporate learning	-	-	-	-	-	BD, RD, SE

*Assist customers (AC)* [27] is with survivability features at blackouts, such as backup generators, power storages, or control components. This service is clearly associated with the end-user's part of the power grid, and it might be related to an idea for a new consultation service for the industry. It is associated with the consumer component and product development and product support business processes.

*Backup power (BP)* [27–29] improves resilience, mostly in a form of gas-fired generators. A backup power source is installed on the customer side and is an alternative means to order fulfilment process. It is associated with the consumer component and fulfils order business process.

*Blackout drill (BD)* [28] is a preparatory measure for better disaster management. Blackout drills are associated with the control component and corporate learning business process.

*Contingency plan (CP)* [30] is a tool such as cutting off consumers by a set of criteria to minimize economic loss. Other types of contingency plans might be deduced for grid operators and consumers. Contingency plans are likely to be related to the control component and strategy formulation and directing business processes.

*Control switch (CS)* [29] may improve resilience, e.g., after receiving a signal from a smart meter a control software switches to a backup power or changes typology of a grid. This tool is associated with the control component and order fulfilment business processes.

*Controlled islanding (CI)* [31–33] may improve resilience of a power grid. The core idea is based on splitting a grid into islands of balanced production-consumption, usually to avoid cascade failure, minimize lost load and speed-up restoration. Controlled islanding is related to the control component (though producers and consumers might be affected as well, and redesign of the grid would affect most

of the physical components) and order fulfilment business processes.

*Crew staging (CRS)* [34] is a preparatory measure to improve resilience. Crew staging is likely to be associated with the control component and support personnel business process.

*Damage assessment (DA)* [27, 34–36] provides the size and extent of the damage and resources required. Damage assessment is likely to be associated with the control component and direct business and manage technology business processes.

*Damage prediction (DP)* [34] is a preparatory measure to improve resilience. Damage prediction is likely to be associated with control component and direct business and manage technology business processes.

*Demand-side management (DSM)* [37] is a process of managing energy consumption. Demand-side management is related to control and consumer components and order fulfilment business process.

*Disaster forecasting (DF)* [30], if the industry is notified about the disaster in advance, damage might be minimized. Disaster forecasting is related to the control component and business direction and technology management business processes.

*Disaster management group (DMG)* [38] is used to manage the impacts of power outages. Groups in Germany are formed from local fire brigades, administrative departments, and a disaster management authority [38]. The UK has similar groups in energy [39] and other sectors. Disaster management group is a part of the control component and business directing business process.

*Distributed automation (DIA)* [29] enhances the resilience of distribution system via accurate and in-time control. Distributed automation is associated with the control component and order fulfilment business process.



*Distributed generation (DIG)* [29] enhances the resilience of distribution system via local sources of energy. Distributed generation is associated with the producer component and order fulfilment business process.

*Gas-fired generator (GG)* [40] is used for distributed generation. Distributed generation is associated with the producer component and order fulfilment and technology management business processes.

*Integrate black-start resources (IB)* [41] may improve resilience. It is associated with the producer component and business direction business process.

*Load shedding (LS)* [42] is an emergency control action to avoid cascade failure by cutting a subset of customers. Load shedding is associated with control and consumer components and order fulfilment business process.

*Maintenance scheduling (MS)* [43–45] may improve resilience of the power grid. Maintenance is associated with the control component and technology management business process.

*Manage vegetation (MV)* [41] is mostly by cutting trees near overhead power lines. It is associated with the power line component and technology management business process.

*Microgrid (MG)* [46–52] is a group of interconnected consumers and producers of energy resources that act as a single controllable entity with respect to the grid. Microgrid may be associated with producer, consumer, control components, and order fulfilment and technology management business processes.

*Mobile emergency generator (MEG)* [53] is used for temporal distribution generation during emergencies. It is associated with producer component and fulfils order business process.

*Mobile unit transformer (MUT)* [54] is used during failures of stationary transformers or their inability to process the load. It is related to substation component and fulfils order business process.

*Modelling using IEEE bus test systems (MO)* [33] and many other authors have mentioned that IEEE bus models are useful for evaluating for resilience changes to power grids. IEEE bus is associated with substation and control components and technology management business process.

*Mutual assistance framework (MAF)* [27] has been set up in Europe to aid restoration after major power disasters with access to spare parts and workforces. The framework is associated with the control component and business direction business process.

*Network reconfiguration (NR)* [55–59] is as automatic and dynamic change of topology of the power grid. Network reconfiguration is likely associated with the control component and order fulfilment business process.

*Optimal reactive reserve (OR)* [60] that meets demand spikes under heavily loaded conditions allows the avoidance of voltage instability problems. Optimal reactive reserve is likely associated with the producer component and strategy formulation business processes.

*Outage management system (OM)* [61] can dramatically decrease the durations and sizes of power outages. Outage

management system is associated with the control component and business direction business process.

*Overhead structure reinforcement (OS)* [62] improves resilience of the grid, which can be achieved by use of robust materials, optimization of tower foundations for soil type and weather conditions. Overhead structure reinforcement is associated with the power line component and technology management business process.

*Performance prediction of renewable-based resources (PP)* [63] improves resilience of the grid by overcoming uncertainty and variability of renewable-based production of electricity. Performance prediction is associated with producer and control components and order fulfilment business process.

*Phasor measurement unit (PMU)* [64–68] is a device that provides synchronized, real-time, dense, and highly accurate measurement of current and voltage phasors. PMU is associated with the control component and technology management business process.

*Post-incident investigation practices (PI)* [35] are used to study major outages post factum and improve recovery for the future. These investigation practices are associated with the control component and business direction and technology management business process.

*Power system stabilizer (PSS)* [69] dynamically provides supplementary feedback signals aiding power system control, thus adding the grid's resilience. PSS is associated with the control component and order fulfilment and technology management business process.

*Repositioning of mobile emergency generators (pMEG)* [53] is a technique for optimization of location mobile emergency generators. It is associated with control component and fulfils order business process.

*Real-time statistical analysis (RSA)* [70] may identify signals of an approaching blackout; thus utilization of this analysis followed by blackout preventive actions aids the grid's resilience. The real-time analysis is associated with the control component and business direction and order fulfilment business process.

*Reallocate transmission routes (RAT)* [41, 71]: a grid with elevated or reallocated substations (and power lines) might be less prone to extreme weather and floods thus more resilient. Reallocation of substations is associated with substations and power lines components and strategy formulation and technology management business processes.

*Redundant transmission routes (RED)* [71]: construction of redundant transmission lines improves resilience of the power grid. Transmission routes are associated with substations and power line components and strategy formulation and technology management business processes.

*Reinforce structure (RS)* [41] increases resilience of the grid. It is associated with all power grid components and technology management business process.

*Research and development of transformers for resilience (R&D)* [35]: continuous R&D of transformers for resilience improves resilience of the power grid. This activity is associated with the substation component and technology management business process.

*Resilience-constrained hourly unit commitment (RCUC)* [72]: this technique optimizes the loading of generators. It is associated with the producer and controller components and fulfills order business process.

*Restoration drills (RD)* [35] increase resilience of the power grid. Restoration drills are associated with the control component and corporate learning business process.

*Restoration management (RM)* [34, 73] may be used to increase resilience of the power grid. It is likely to be associated with the control component and business direction business process.

*Restoration priority (RP)* [74] may reduce loss load thus improving resilience of the power grid. Restoration priority is likely to be associated with the control component and business direction business process.

*Satellite technology (ST)* [75] may be used to detect power outages in real time, thus improving resilience of the power grids. It is associated with the control component and technology management business process.

*Supervisory control and data acquisition (SCADA)* [6, 76, 77] is a data collection tool for control and management of the grid. SCADA is associated with the control component and technology management business process.

*Smart grid (SG)* [10, 65, 78–84] is a more resilient version of the power grid. A smart grid utilizes advanced data collection and analysis. It is associated with the control component (yet have physical and virtual sensors in all other components) and order fulfilment and technology management business processes.

*Smart meter (SM)* [10, 40, 55, 61, 84] is a device that feeds data for dynamic control of the grid (e.g., production, load). Smart meters are installed to all components of the grid but are associated with the control component and order fulfilment business process.

*Standards' development (SD)* [35] for utility cyber control systems is a strategy for increasing resilience of the grid; this can be generalized to development of power grid standards for resilience. Standards' development is associated with the control component and technology management and corporate learning business processes.

*State estimation (SE)* [68, 77] provides a real-time state of the power system, which is relevant for reactive and proactive control of the power grid. State estimation is associated with the control component and technology management business process.

*Transportable energy storage system (TESS)* [85] is proposed for post-disaster restoration of large distribution grids via initialization of microgrids and consists of an energy storage, means of transportation, control of transportation, and application scheme.

*Undergrounding power lines (UL)* [41, 71] could enhance the resilience via replacement of overhead power as the underground power lines are less prone to wind-induced damage, lightning, and vegetation contact. The price of higher installation and repair costs could be offset against cost of damage or interruption to service. Undergrounding is associated with the power line component and technology management business process.

## 4. Discussion

*4.1. Simulation Modelling.* Volume- and time-driven resilience assessment frameworks do not provide insights into points of failure, triggers, probabilities of failures or triggers, and asset conditions. These are highly relevant information to link both hazards and tools to power grid components. They are needed for resilience research and if a resilience assessment framework does not address these factors, then it has limited application. However, resilience assessment frameworks that do address these factors are subsector specific and, moreover, asset-specific and therefore have an absence of a systemic and cross-sectoral understanding of resilience. They would also require interdependency studies and cross-sectors resilience-driven projects.

The need for generic and the need for specific resilience assessment frameworks are a contradiction. This contradiction can be removed by the abstraction between two stages of resilience calculation. The system output is calculated at the sector- or asset-specific first stage by the means of real system or a simulation model. These sector-specific outputs are processed by the same function into a generic resilience values at the second stage.

Resilience is an emergent quality of a system to a hazard. Due to the constructional and behavioral complexity of the system of interest, the power grids, this emergent quality cannot be predicted without a system. A hazard and its impact on the power grid from the producers to the consumers increase complexity of the resilience research, while the number of potential tools for resilience engineering and the number of their combinations makes it a challenging research domain.

It is impossible to replicate most of hazards in a real power grid, and for many threats, testing is highly expensive to emulate even on a highly limited scale if legal, e.g., a cyberterrorism or electromagnetic interference, which makes *physical experimentation* as a method for resilience research a very limited application. *Mental experimentation* has limited use for a conceptual study of resilience which attempts to quantify the improvement from alternative interventions and is an unreasonable method for resilience engineering of complex systems using a toolbox. Post-event studies of major outages may provide some important information and insights, and the lack of *observational* capabilities on site during an emergency can be overcome with multitude of sensors logging power grid supply and consumption; however, access to this data is usually limited due to organizational reasons; it is a unique case that cannot be replicated, and it is *ex-post* evaluation limiting its usability for resilience engineering. *Inductive reasoning* is partially applicable to resilience research as it relies on a strong and cogent selection of facts, which is limited for the reasons above, similarly, with *formalization* for the follow-up logical and mathematical reasoning. *Theoretical reasoning* would be a method of predicting results, but the *theory* for engineering resilience does not exist; therefore, it is an inapplicable method at this moment. A *model* may selectively address the constructional and behavioral complexity of a system and provide insights into the resilience value of alternative subsets of tools without

the costs or barriers of real-world experimentation which makes modelling and especially computational modelling, the most suitable method for *ex-post* and *ex-ante* resilience research.

Banks *et al.* [86] classify simulation models as either static or dynamic, deterministic, stochastic, discrete, or continuous. A model for resilience research and engineering must be *dynamic* since static models have very limited capabilities in addressing behavioral aspects; *stochastic* as complex multidomain deterministic rules are very challenging to define; *discrete* because continuous models are mostly based on differential and integral calculus that if applicable has highly specialized application, e.g., for thermodynamics. Dynamic stochastic discrete models appear most suitable; however, multiple formalisms for simulation modelling fit these characteristics, and multiple formalisms might be used within a complex computational model with some submodels being static, deterministic, or continuous.

Some architectural considerations could be made for the modelling- and toolbox-driven resilience research. It can be safely assumed that resilience research would incorporate massive simulation modelling. A colossal number of interactions between elements are expected. A state or a change of state of elements before an element may result a change of state in this element. A high-level and lightweight computation is a preferred option unless required otherwise. It is beneficial to represent a power grid as a *network* of assets producing, transforming, redirecting, or consuming electric energy, because it fits the typology of assets and is efficiently computable within vector, matrix, or tensor algebras and is applicable and partially transferrable to modelling of infrastructures from other sectors that is especially relevant for multisectoral resilience research.

*Interfaces* for control and information exchange model and submodel are critical, as is parallelization of modelling and computation to improve computational durations. A hazard may affect the behavior of consumers, which is one of viable research objectives, and it is commonly addressed with *agent-based modelling*. Toolbox-induced changes in a model assume the same domain language to describe assets within the model and tools from the toolbox, and *ontology* can address both the structure of the language and the database of assets and tools.

According to Kelly *et al.* [87], five modelling approaches are most commonly used for integrated environmental assessment and management. Each of these approaches has been used for modelling complex systems: system dynamics [88, 89], Bayesian networks [90, 91], coupled component models [92], agent-based models [93, 94], and knowledge-based models [95]. In addition to the summary of these approaches (see Tables 1 and 3 in [87]), Kelly *et al.* [87] provided a heuristic for selection of one of these approaches under standard application (see Figure 1 in [87]). However, as is shown in review papers [96–107] a more comprehensive heuristics on a larger spectrum of modelling approaches would benefit the resilience research and engineering.

**4.2. Toolbox.** The toolbox is a collection and description of tools to improve resilience of a power grid. A grid has

a certain level of resilience to a hazard, and by changing the tools available to the grid its resilience may stay the same, and it may be increased or decreased to the same hazard. A change is based on intervention of a subset of tools into the grid, and tools within a subset might be utilized with a different intensity. Simulation modelling is the most convenient method to evaluate the resilience of a changed grid.

Power grid components and organizational business processes are used to structure the toolbox, which currently may be associated with a two-dimensional array, as a subset for a change. A network-based model of a grid may also be represented with arrays. Therefore, this approach on a theoretical level allows utilization of one of the most convenient and efficient methods for mass computation, vector operations. Vector, matrix, and tensor algebras provide a mathematical apparatus for this and more complex computations, e.g., incorporating dynamics or multilayered virtual subgrids; combinatorics and set theory could support scenario design.

Qualitative resilience assessment frameworks [19] which could be used to direct attention to certain ‘cells’ in the toolbox for these frameworks are based on best practices for resilience.

Examples of such frameworks are a semiquantitative approach proposed by Shirali *et al.* [108] qualitative frameworks that could be transformed into a semiquantitative index and subsets of resilience indicators proposed by van der Merwe *et al.* [109].

The toolbox shows the areas currently lacking tools to improve resilience. While this could be partially explained by insufficiently full and systematic literature review, this also indicated the lack of interest in this area, and if the lack of interest is unreasonable, then the toolbox shows gaps in the current state of resilience research and engineering.

While the current two-dimensional classification provides a valuable categorization for the tools to improve resilience, it might be insufficient because it does not address the constructional complexity of a power grid and its components. The component dimension might be replaced or extended with system’s representation (see Figure 2.3 in [110]), systems engineering processes (see Figure 1.1 in [110]), domain ontology [111], or systems holarchy [112], though the latter is more suitable for grid modelling, but a strong relationship between the holarchy and the toolbox must be established beforehand. Other classifications might be used as well.

Similar toolboxes could be created for resilience research and engineering of other sectors of infrastructure. Multisectoral toolboxes would require an additional research, probably resulting in a higher-dimensionality of classification matrices and generic functional-constructional descriptors of tools.

**4.3. Resilience Value.** Most of the general quantitative resilience assessment frameworks in [19, 20] utilize volume, time, or volume-over-time data to calculate resilience value. These variables and different operators from different branches of mathematics form alternative mathematical equations for calculation of resilience value. Each equation is supported by

TABLE 5: A temporal outlook to disruptive events [8]. Relative to a disruptive event, resilience may be improved long before, before, during, after, and long after the event.

Time period	Objectives	System state	System function	Event feature
Long before	Identify possible events, know system weaknesses, plan response and restoration, prepare resources	Normal	Awareness, planning, and preparation	Predisposition
Before	Monitor system, monitor threats, ensure resources are available, switch to alert state if a threat is detected	Alert	Monitor, detect, recognition	Precursors
During	Stop effect propagation, core goals, manage cascade failures	Emergency	Containment, mitigation	Effects
After	Resume normal operations, reverse alert/emergency, manage delayed failures	Restoration	Repair, reconfigure, replacement	Impact
Long after	Understand what happened, know why it happened, plan and implement actions to improve resilience	Recovery	Reflection, learning, improvement	Outcome

at least one line of argumentation which could be probed and developed further. Other ideas may and would likely result from other equations, for example,

- (1) Minimum resilience value is taken as the resilience of the system, and the volume might now be the lowest delivery. Resilience value is calculated for each time value at the first step, and the minimum resilience value is selected at the second step.
- (2) Lowest (minimum) delivery volume defines the resilience of the system. The minimum value is selected for the equation with the associated time and this time is used to select the baseline value.
- (3) This equation can be further extended by incorporating more statistical values:
  - (a) mean (the average) of delivered and baseline volume during at disrupted state,
  - (b) median (the middle value of an ordered list) of volume,
  - (c) mode (the most common value),
  - (d) filtering out outliers with the calculation of averages (or minimum, as above),
  - (e) using other statistical measures and statistical analysis techniques, e.g., quartiles with the associated box-and-whisker plots.
- (4) Statistics may provide additional insights. For example, skewness may indicate whether the system fails or recovers faster, while kurtosis may indicate whether the system has extreme drops of performance or how fast the hardest part of the system is processed. While these can be easily deduced from the visualization, the quantitative statistical technique is instrumental for automatic optimization of the infrastructure.

An equation can be used on natural (e.g., litres) and normalized (e.g., percent) data. Some considerations might

require nonlinear normalization (e.g., onto a logarithmic scale for similar reasons the Richter magnitude scale is a logarithmic scale) of the input data or resilience values, in this case an equation might undergo adjustments. The nonlinear normalization might be the answer to the intensity aspect of the hazards' profile, but this requires additional research.

The significance of each tool in isolation with respect to resilience improvement is questionable unless a sound logical and pragmatic reasoning is provided; usually, papers are lacking in that aspect. In the event the toolbox is applied for resilience research, then one or more equations would be selected, combined, and developed. Simulation modelling would support any tool and related equations, subject to limitations of the platform.

*4.4. Decision-Making.* Several preconditions must be met before application of a resilience assessment framework as resilience values are calculated for strategic decision-making and operationalization of the decisions. If resilience value is insufficient, then a change might be introduced into the system resulting in a new resilience value of the changed system. The difference between the old and new resilience values represents the impact of the change on resilience, and the delta resilience might be compared to delta sustainability or monetary investment.

Overall, resilience engineering is likely a continuous activity and resilience of a complex system can be addressed at different stages to disruptive events; see Table 5.

Multiple methods can be used to define the criteria to consider whether the output of each resilience assessment framework indicates a 'good' level of resilience. If resilience value is used for decision-making, it is multicriteria decision-making, and thus this task is about naming complementary criteria for the multicriteria decision-making. Three approaches to multicriteria decision-making are listed below from the least to the most comprehensive.

*Single criteria indicators:* with resilience value as supply/demand ratio, the criteria are the threshold value for the volume of delivery or capacity margin.

*Multicriteria n-lemmas:* an example of such concept is a well-known *cost–quality–time* triangle with a heuristics rule that for any system it is possible to meet two qualities only. Similar triangle is associated with the resilience of electricity system [113] presenting *sustainability–security of supply–affordability* triangle, where resilience is a part of security and decarbonization is a part of sustainability indicators.

*Evaluation frameworks:* HM treasury provides guidance on how to appraise and evaluate policies, projects, and programmes; see The Green Book [114].

## 5. Conclusions

A structured collection of tools for engineering resilience and an approach for resilience research to improve the resilience of a power grid infrastructure is described in this paper. The collection is a two-dimensional array formed from a classification of components of a power grid and a typology of business processes. These two dimensions provide a physical and operational outlook for a power grid. The approach for resilience research is based on building a simulation model for a power grid which utilizes a resilience assessment equation to assess its baseline resilience to a hazards profile, then iteratively selecting a subset of tools from the collection and introducing these as interventions in the power grid simulation. Calculating the difference in resilience associated with each subset supports multicriteria decision-making to find the most convenient subset of interventions for a power grid and hazards profile.

This highlights the importance of a portfolio which is strategies which should take account of variety of natural disasters relevant to the regions/geographical areas of the power grid (in addition to general hazards), as well as compensating for design decisions which can compromise resilience.

The approach outlined of iterative testing of subsets of tools assumes simulation modelling. The simulation model should be in-line with the structured description of the elements of the collection and the possible paths of impact of hazards. Hazard is a mandatory element of resilience experience; however, resilience of a system to a hazard is less relevant than resilience of a system to a hazards' profile. Matrix algebra, set theory, and combinatorics allow automatic construction of scenarios within this approach. Volume, time, or volume-over-time resilience assessment framework could be selected or designed for the simulation modelling architecture. The resulting resilience value, in combination with other factors, could be used for multicriteria decision-making on the quality of the subset.

Resilience is an emergent quality of a power grid system, and therefore resilience research and interventions must be system-driven. Usually, multiple infrastructures are utilized by social or production systems, and a hazard often affects multiple infrastructures as well-illustrated by assessment of volcanic hazards by Wilson et al. [115], which also states

the importance of hazards' profile for a country. Moreover, interdependencies [1, 2] between infrastructures may impact the recovery and resilience of a single infrastructure. Therefore, simulation-based search/optimization for a resilient infrastructure would benefit from a simultaneous search/optimization for multisectoral resilience; e.g., Najafi et al. [52] described resilience improvement of power-water distribution system. However, simulation-based resilience research and engineering require in-depth single- or multi-topic analysis of hazards, infrastructure components, vulnerabilities, tools, regions, simulation modelling techniques, and search or optimization algorithms for the follow-up model driven resilience research and engineering.

## Data Availability

Published articles provide the data used to support the findings of this study. The selection rules for these articles are included within the article.

## Conflicts of Interest

The manuscript has no conflicts of interest with the source of funding or otherwise.

## Acknowledgments

We kindly acknowledge EPSRC grant funding reference EP/N010019/1 regarding the ENCORE project on engineering complexity resilience.

## References

- [1] J. W. Hall, J. J. Henriques, A. J. Hickford, and R. J. Nicholls, "Systems-of-systems analysis of national infrastructure," *Proceedings of the Institution of Civil Engineers: Engineering Sustainability*, vol. 166, no. 5, pp. 249–257, 2013.
- [2] M. A. Ehlen and V. N. Vargas, "Multi-hazard, multi-infrastructure, economic scenario analysis," *Environment Systems & Decisions*, vol. 33, no. 1, pp. 60–75, 2013.
- [3] Z. Bo, O. Shaojie, Z. Jianhua, S. Hui, W. Geng, and Z. Ming, "An analysis of previous blackouts in the world: lessons for China's power industry," *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 1151–1163, 2015.
- [4] "Billion-Dollar Weather and Climate Disasters: Table of Events, National Oceanic and Atmospheric Administration, 2019," <https://www.ncdc.noaa.gov/billions/events/US/1980-2019>.
- [5] S. N. Talukdar, J. Apt, M. Ilic, L. B. Lave, and M. G. Morgan, "Cascading failures: Survival versus prevention," *The Electricity Journal*, vol. 16, no. 9, pp. 25–31, 2003.
- [6] P. Andriani and B. McKelvey, "Beyond Gaussian averages: Redirecting international business and management research toward extreme events and power laws," *Journal of International Business Studies*, vol. 38, no. 7, pp. 1212–1230, 2007.
- [7] D. F. Warne, *Newnes Electrical Power Engineer's Handbook*, Elsevier, Oxford, UK, 2nd edition, 2005.
- [8] M. Mayfield, G. Punzo, R. Beasley, G. Clarke, N. Holt, and S. Jobbins, "Challenges of complexity and resilience in complex engineering systems," *ENCORE Network+ White Paper*, 2018.

- [9] S. Mukherjee, R. Nateghi, and M. Hastak, "A multi-hazard approach to assess severe weather-induced major power outage risks in the U.S.," *Reliability Engineering & System Safety*, vol. 175, pp. 283–305, 2018.
- [10] Z. Bie, Y. Lin, G. Li, and F. Li, "Battling the extreme: a study on the power system resilience," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1253–1266, 2017.
- [11] "National Risk Register of Civil Emergencies – 2017 Edition," Cabinet Office, National security and intelligence, and The Right Honorable Caroline Nokes MP," <https://www.gov.uk/government/publications/national-risk-register-of-civil-emergencies-2017-edition>, 2017.
- [12] A. Collins, "The Global Risks Report 2018," Tech. Rep., World Economic Forum, Geneva, Switzerland, 2018.
- [13] "Black Sky Hazards, EIS Council," <https://www.eiscouncil.org/blackSky.aspx>, 2019.
- [14] T. Huang, S. L. Voronca, A. A. Purcarea, A. Estebarsari, and E. Bompard, "Analysis of chain of events in major historic power outages," *Advances in Electrical and Computer Engineering*, vol. 14, no. 3, pp. 63–70, 2014.
- [15] L. Fisher, "More than 70 ways to show resilience," *Nature*, vol. 518, p. 35, 2015.
- [16] "Presidential Policy Directive – Critical Infrastructure Security and Resilience," The White House, 2013, <https://obamawhitehouse.archives.gov/the-press-office/2013/02/12/presidential-policy-directive-critical-infrastructure-security-and-resil>.
- [17] "Keeping the country running: natural hazards and infrastructure, Cabinet Office," 21 October 2011, <https://www.gov.uk/government/publications/keeping-the-country-running-natural-hazards-and-infrastructure>.
- [18] UNISDR, 2009 UNISDR Terminology for Disaster Risk Reduction, 2009, <https://www.unisdr.org/we/inform/publications/7817>.
- [19] S. Hosseini, K. Barker, and J. E. Ramirez-Marquez, "A review of definitions and measures of system resilience," *Reliability Engineering & System Safety*, vol. 145, pp. 47–61, 2016.
- [20] N. Yodo and P. Wang, "Engineering resilience quantification and system design implications: A literature survey," *Journal of Mechanical Design*, vol. 138, no. 11, Article ID 111408, 2016.
- [21] M. Omer, A. Mostashari, and R. Nilchiani, "Assessing resilience in a regional road-based transportation network," *International Journal of Industrial and Systems Engineering*, vol. 13, no. 4, pp. 389–408, 2013.
- [22] B. M. Ayyub, "Systems resilience for multihazard environments: Definition, metrics, and valuation for decision making," *Risk Analysis*, vol. 34, no. 2, pp. 340–355, 2014.
- [23] D. Gama Dessavre, J. E. Ramirez-Marquez, and K. Barker, "Multidimensional approach to complex system resilience analysis," *Reliability Engineering & System Safety*, vol. 149, pp. 34–43, 2016.
- [24] U. S. Bititci, K. Mendibil, V. Martinez, and P. Albores, "Measuring and managing performance in extended enterprises," *International Journal of Operations and Production Management*, vol. 25, no. 4, pp. 333–353, 2005.
- [25] U. S. Bititci, V. Martinez, P. Albores, and K. Mendibil, "Creating and sustaining competitive advantage in collaborative systems: The what and the how," *Production Planning and Control*, vol. 14, no. 5, pp. 410–424, 2003.
- [26] S. Oktaufanus, "Improvement CIMOSA," <http://shutsatria.blogspot.com/2012/01/v-behaviorurldefaultvml-o.html>, 2012.
- [27] C. Chen, J. Wang, and D. Ton, "Modernizing distribution system restoration to achieve grid resiliency against extreme weather events: an integrated solution," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1267–1288, 2017.
- [28] S. B. Miles, H. Gallagher, and C. J. Huxford, "Restoration and impacts from the september 8, 2011, san diego power outage," *Journal of Infrastructure Systems*, vol. 20, no. 2, Article ID 05014002, 2014.
- [29] J. Najafi, A. Peiravi, and J. M. Guerrero, "Power distribution system improvement planning under hurricanes based on a new resilience index," *Sustainable Cities and Society*, vol. 39, pp. 592–604, 2018.
- [30] K. Kim and Y. Cho, "Estimation of power outage costs in the industrial sector of South Korea," *Energy Policy*, vol. 101, pp. 236–245, 2017.
- [31] M. Panteli, D. N. Trakas, P. Mancarella, and N. D. Hatzigiorgi, "Boosting the power grid resilience to extreme weather events using defensive islanding," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2913–2922, 2016.
- [32] S. Kamali and T. Amraee, "Blackout prediction in interconnected electric energy systems considering generation re-dispatch and energy curtailment," *Applied Energy*, vol. 187, pp. 50–61, 2017.
- [33] V. Thirugnanasambandam and T. Jain, "Placement of synchronized measurements for power system observability during cascaded outages," *International Journal of Emerging Electric Power Systems*, vol. 18, no. 6, 2017.
- [34] R. Nateghi, S. D. Guikema, and S. M. Quiring, "Forecasting hurricane-induced power outage durations," *Natural Hazards*, vol. 74, no. 3, pp. 1795–1811, 2014.
- [35] B. A. Wender, M. G. Morgan, and K. J. Holmes, "Enhancing the resilience of electricity systems," *Engineering Journal*, vol. 3, no. 5, pp. 580–582, 2017.
- [36] K. Forssén, K. Mäki, M. Räikkönen, and R. Molarius, "Resilience of electricity distribution networks against extreme weather conditions," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 3, no. 2, Article ID 021005, 2017.
- [37] H. Syadli, M. P. Abdullah, M. Y. Hassan, and F. Hussin, "Demand side management for reducing rolling blackouts due to power supply deficit in sumatra," *Jurnal Teknologi*, vol. 69, no. 5, pp. 39–43, 2014.
- [38] T. Münzberg, M. Wiens, and F. Schultmann, "A spatial-temporal vulnerability assessment to support the building of community resilience against power outage impacts," *Technological Forecasting & Social Change*, vol. 121, pp. 99–118, 2017.
- [39] Department for Business, Energy, and Industrial Strategy, "Electricity Supply Emergency Code, Department for Business, Energy & Industrial Strategy," <https://www.gov.uk/government/publications/electricity-supply-emergency-code>, 2018.
- [40] S. Baik, A. L. Davis, and M. G. Morgan, "Assessing the cost of large-scale power outages to residential customers," *Risk Analysis*, vol. 38, no. 2, pp. 283–296, 2018.
- [41] S. Ma, B. Chen, and Z. Wang, "Resilience enhancement strategy for distribution systems under extreme weather events," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 1442–1451, 2018.
- [42] R. Mageshvaran and T. Jayabarathi, "GSO based optimization of steady state load shedding in power systems to mitigate blackout during generation contingencies," *Ain Shams Engineering Journal*, vol. 6, no. 1, pp. 145–160, 2014.
- [43] M. Abirami, S. Ganesan, S. Subramanian, and R. Anandhakumar, "Source and transmission line maintenance outage scheduling in a power system using teaching learning based

- optimization algorithm,” *Applied Soft Computing*, vol. 21, pp. 72–83, 2014.
- [44] S. Tang, C. Hale, and H. Thaker, “Reliability modeling of power transformers with maintenance outage,” *Systems Science & Control Engineering*, vol. 2, no. 1, pp. 316–324, 2014.
- [45] G. Ji, W. Wu, and B. Zhang, “Robust generation maintenance scheduling considering wind power and forced outages,” *IET Renewable Power Generation*, vol. 10, no. 5, pp. 634–641, 2016.
- [46] C. Stefanovic, M. Angelichinoski, P. Danzi, and P. Popovski, “Resilient and secure low-rate connectivity for smart energy applications through power talk in DC microgrids,” *IEEE Communications Magazine*, vol. 55, no. 10, pp. 83–89, 2017.
- [47] X. Liu, M. Shahidehpour, Z. Li, X. Liu, Y. Cao, and Z. Bie, “Microgrids for enhancing the power grid resilience in extreme conditions,” *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 589–597, 2017.
- [48] H. Farzin, M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie, “Enhancing power system resilience through hierarchical outage management in multi-microgrids,” *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2869–2879, 2016.
- [49] S. Bird and C. Hotaling, “Multi-stakeholder microgrids for resilience and sustainability,” *Environmental Hazards*, vol. 16, no. 2, pp. 116–132, 2017.
- [50] Y. Koraz and A. Gabbar, “Risk analysis and self-healing approach for resilient interconnect micro energy grids,” *Sustainable Cities and Society*, vol. 32, pp. 638–653, 2017.
- [51] K. B. Jones, M. James, and R.-A. Mastor, “Securing our energy future: three international perspectives on microgrids and distributed renewables as a path toward resilient communities,” *Environmental Hazards*, vol. 16, no. 2, pp. 99–115, 2017.
- [52] J. Najafi, A. Peiravi, A. Anvari-Moghaddam, and J. M. Guerrero, “Resilience improvement planning of power-water distribution systems with multiple microgrids against hurricanes using clean strategies,” *Journal of Cleaner Production*, vol. 223, pp. 109–126, 2019.
- [53] S. Lei, J. Wang, C. Chen, and Y. Hou, “Mobile emergency generator pre-positioning and real-time allocation for resilient response to natural disasters,” *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 2030–2041, 2018.
- [54] G. A. Hamoud, “Use of mobile unit transformers in high voltage load stations,” in *Proceedings of the 10th International Conference on Probabilistic Methods Applied to Power Systems*, Rincon, Ga, USA, 2008.
- [55] S. Baik, M. G. Morgan, and A. L. Davis, “Providing limited local electric service during a major grid outage: a first assessment based on customer willingness to pay,” *Risk Analysis*, vol. 38, no. 2, pp. 272–282, 2018.
- [56] J. Wei and D. Kundur, “GOAliE: goal-seeking obstacle and collision evasion for resilient multicast routing in smart grid,” *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 567–579, 2016.
- [57] Y. Fang and G. Sansavini, “Optimizing power system investments and resilience against attacks,” *Reliability Engineering & System Safety*, vol. 159, pp. 161–173, 2017.
- [58] H. Lin, C. Chen, J. Wang et al., “Self-healing attack-resilient PMU network for power system operation,” *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1551–1565, 2018.
- [59] A. Nuhanović, J. Hivziefendić, and A. Hadžimehmedović, “Distribution network reconfiguration considering power losses and outages costs using genetic algorithm,” *Journal of Electrical Engineering*, vol. 64, no. 5, pp. 265–271, 2013.
- [60] S. Padaiyatchi and M. Daniel, “OPF-based reactive power planning and voltage stability limit improvement under single line outage contingency condition through evolutionary algorithms,” *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 21, no. 4, pp. 1092–1106, 2013.
- [61] H. Sun, Z. Wang, J. Wang, Z. Huang, N. Carrington, and J. Liao, “Data-driven power outage detection by social sensors,” *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2516–2524, 2016.
- [62] D. Kyung, Y. Choi, S. Jeong, and J. Lee, “Improved performance of electrical transmission tower structure using connected foundation in soft ground,” *Energies*, vol. 8, no. 6, pp. 4963–4982, 2015.
- [63] D. Apostolopoulou, Z. De Grève, and M. McCulloch, “Robust optimization for hydroelectric system operation under uncertainty,” *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3337–3348, 2018.
- [64] J.-C. Chen, W.-T. Li, C.-K. Wen, J.-H. Teng, and P. Ting, “Efficient identification method for power line outages in the smart power grid,” *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1788–1800, 2014.
- [65] A. Ahmed, M. Awais, M. Naem et al., “Multiple power line outage detection in smart grids: probabilistic bayesian approach,” *IEEE Access*, vol. 6, pp. 10650–10661, 2018.
- [66] A. Arabali, M. Majidi, M. S. Fadali, and M. Etezadi-Amoli, “Line outage identification-based state estimation in a power system with multiple line outages,” *Electric Power Systems Research*, vol. 133, pp. 79–86, 2016.
- [67] O. Gomez, M. A. Rios, and G. Anders, “Reliability-based phasor measurement unit placement in power systems considering transmission line outages and channel limits,” *IET Generation, Transmission & Distribution*, vol. 8, no. 1, pp. 121–130, 2014.
- [68] X. Tai, D. Marelli, E. Rohr, and M. Fu, “Optimal PMU placement for power system state estimation with random component outages,” *International Journal of Electrical Power & Energy Systems*, vol. 51, pp. 35–42, 2013.
- [69] M. Soliman, “Robust non-fragile power system stabilizer,” *International Journal of Electrical Power & Energy Systems*, vol. 64, pp. 626–634, 2015.
- [70] M. Moazzami, R.-A. Hooshmand, A. Khodabakhshian, and M. Yazdanpanah, “Blackout prevention in power system using flexible AC transmission system devices and combined corrective actions,” *Electric Power Components and Systems*, vol. 41, no. 15, pp. 1433–1455, 2013.
- [71] M. Panteli and P. Mancarella, “Modeling and evaluating the resilience of critical electrical power infrastructure to extreme weather events,” *IEEE Systems Journal*, vol. 11, no. 3, pp. 1733–1742, 2017.
- [72] Y. Wang, L. Huang, M. Shahidehpour, L. L. Lai, H. Yuan, and F. Y. Xu, “Resilience-Constrained hourly unit commitment in electricity grids,” *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5604–5614, 2018.
- [73] M. Figueroa-Candia, F. A. Felder, and D. W. Coit, “Resiliency-based optimization of restoration policies for electric power distribution systems,” *Electric Power Systems Research*, vol. 161, pp. 188–198, 2018.
- [74] R. S. Liévanos and C. Horne, “Unequal resilience: The duration of electricity outages,” *Energy Policy*, vol. 108, pp. 201–211, 2017.
- [75] T. A. Cole, D. W. Wanik, A. L. Molthan, M. O. Román, and R. E. Griffin, “Synergistic use of nighttime satellite data, electric utility infrastructure, and ambient population to improve power outage detections in urban areas,” *Remote Sensing*, vol. 9, no. 3, p. 286, 2017.

- [76] C. Yang, Z.-H. Guan, Z.-W. Liu, J. Chen, M. Chi, and G.-L. Zheng, "Wide-area multiple line-outages detection in power complex networks," *International Journal of Electrical Power & Energy Systems*, vol. 79, pp. 132–141, 2016.
- [77] A. Ashok, M. Govindarasu, and J. Wang, "Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1389–1407, 2017.
- [78] C. Chen, J. Wang, F. Qiu, and D. Zhao, "Resilient distribution system by microgrids formation after natural disasters," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 958–966, 2016.
- [79] A. Dabrowski, J. Ullrich, and E. R. Weippl, "Botnets causing blackouts: how coordinated load attacks can destabilize the power grid," *Elektrotechnik und Informationstechnik*, vol. 135, no. 3, pp. 250–255, 2018.
- [80] L. Zhao and W.-Z. Song, "Distributed power-line outage detection based on wide area measurement system," *Sensors*, vol. 14, no. 7, pp. 13114–13133, 2014.
- [81] M. M. Othman, N. A. Salim, and I. Musirin, "Sustainability from the occurrence of critical dynamic power system blackout determined by using the stochastic event tree technique," *Sustainability*, vol. 9, no. 6, p. 941, 2017.
- [82] A. AlMajali, A. Viswanathan, and C. Neuman, "Resilience evaluation of demand response as spinning reserve under cyber-physical threats," *Electronics*, vol. 6, no. 1, p. 2, 2017.
- [83] E. D. Dongmo, P. Colet, and P. Woafu, "Power grid enhanced resilience using proportional and derivative control with delayed feedback," *The European Physical Journal B*, vol. 90, no. 1, p. 6, 2017.
- [84] D. Tan, D. Baxter, S. Foroozan, and S. Crane, "A first resilient DC-dominated microgrid for mission-critical space applications," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 4, no. 4, pp. 1147–1157, 2016.
- [85] S. Yao, P. Wang, and T. Zhao, "Transportable energy storage for more resilient distribution systems with multiple microgrids," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, 2018.
- [86] J. Banks, J. S. Carson II, B. L. Nelson, and D. M. Nicol, "Introduction to simulation," in *Discrete-Event System Simulation*, pp. 1–22, Pearson Education Limited, Harlow, UK, 5th edition, 2014.
- [87] R. A. Kelly, A. J. Jakeman, O. Barretheau et al., "Selecting among five common modelling approaches for integrated environmental assessment and management," *Environmental Modeling and Software*, vol. 47, pp. 159–181, 2013.
- [88] C. K. Gotangco, J. See, J. P. Dalupang et al., "Quantifying resilience to flooding among households and local government units using system dynamics: a case study in Metro Manila," *Journal of Flood Risk Management*, vol. 9, no. 3, pp. 196–207, 2016.
- [89] Y. X. He, J. Jiao, R. J. Chen, and H. Shu, "The optimization of Chinese power grid investment based on transmission and distribution tariff policy: A system dynamics approach," *Energy Policy*, vol. 113, pp. 112–122, 2018.
- [90] S. Hosseini and K. Barker, "Modeling infrastructure resilience using Bayesian networks: A case study of inland waterway ports," *Computers & Industrial Engineering*, vol. 93, pp. 252–266, 2016.
- [91] S. Hosseini, A. Al Khaled, and M. D. Sarder, "A general framework for assessing system resilience using Bayesian networks: A case study of sulfuric acid manufacturer," *Journal of Manufacturing Systems*, vol. 41, pp. 211–227, 2016.
- [92] S. Peña-Alzate and J. E. C. Barriga, "Approaching the concepts of ecosystems resilience and stability through spatiotemporal system dynamics and agent-based modelling," *Revista Facultad de Ingeniería Universidad de Antioquia*, vol. 84, pp. 84–96, 2017.
- [93] J. M. Gonzalez de Durana, O. Barambones, E. Kremers, and L. Varga, "Agent based modeling of energy networks," *Energy Conversion and Management*, vol. 82, pp. 308–319, 2014.
- [94] G. Pumpuni-Lenss, T. Blackburn, and A. Garstenauer, "Resilience in complex systems: an agent-based Approach," *Systems Engineering*, vol. 20, no. 2, pp. 158–172, 2017.
- [95] A. John, Z. Yang, R. Riahi, and J. Wang, "Application of a collaborative modelling and strategic fuzzy decision support system for selecting appropriate resilience strategies for seaport operations," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 1, no. 3, pp. 159–179, 2014.
- [96] R. W. Scholz, J. Gallati, Q. B. Le, and R. Seidl, "Integrated systems modeling of complex human-environment systems," in *Environmental Literacy in Science and Society: from Knowledge to Decisions*, pp. 341–372, Cambridge University Press, New York, NY, USA, 2011.
- [97] A. S. R. Subramanian and T. A. A. I. Gundersen, "Modeling and simulation of energy systems: a review," *Processes*, vol. 6, no. 12, p. 238, 2018.
- [98] R. H. E. M. Koppelaar, J. Keirstead, N. Shah, and J. Woods, "A review of policy analysis purpose and capabilities of electricity system models," *Renewable & Sustainable Energy Reviews*, vol. 59, pp. 1531–1544, 2016.
- [99] L. Shi, Q. Dai, and Y. Ni, "Cyber-physical interactions in power systems: A review of models, methods, and applications," *Electric Power Systems Research*, vol. 163, pp. 396–412, 2018.
- [100] P. Lopion, P. Markewitz, M. Robinius, and D. Stolten, "A review of current challenges and trends in energy systems modeling," *Renewable & Sustainable Energy Reviews*, vol. 96, pp. 156–166, 2018.
- [101] W. L. Theo, J. S. Lim, W. S. Ho, H. Hashim, and C. T. Lee, "Review of distributed generation (DG) system planning and optimisation techniques: Comparison of numerical and mathematical modelling methods," *Renewable & Sustainable Energy Reviews*, vol. 67, pp. 531–573, 2017.
- [102] J.-H. Cho, Y. Wang, I.-R. Chen, K. S. Chan, and A. Swami, "A survey on modeling and optimizing multi-objective systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1867–1901, 2017.
- [103] Z. Ding, W. Gong, S. Li, and Z. Wu, "System dynamics versus agent-based modeling: a review of complexity simulation in construction waste management," *Sustainability*, vol. 10, no. 7, p. 2484, 2018.
- [104] H.-K. Ringkjøb, P. M. Haugan, and I. M. Solbrenke, "A review of modelling tools for energy and electricity systems with large shares of variable renewables," *Renewable & Sustainable Energy Reviews*, vol. 96, pp. 440–459, 2018.
- [105] D. Olsthoorn, F. Haghghat, and P. A. Mirzaei, "Integration of storage and renewable energy into district heating systems: A review of modelling and optimization," *Integration of Storage and Renewable Energy into District Heating*, vol. 136, pp. 49–64, 2016.
- [106] P. S. Georgilakis and N. D. Hatziaargyriou, "A review of power distribution planning in the modern power systems era: models, methods and future research," *Electric Power Systems Research*, vol. 121, pp. 89–100, 2015.



- [107] L. M. H. Hall and A. R. Buckley, "A review of energy systems models in the UK: Prevalent usage and categorisation," *Applied Energy*, vol. 169, pp. 607–628, 2016.
- [108] G. A. Shirali, I. Mohammadfam, and V. Ebrahimipour, "A new method for quantitative assessment of resilience engineering by PCA and NT approach: A case study in a process industry," *Reliability Engineering & System Safety*, vol. 119, pp. 88–94, 2013.
- [109] S. E. van der Merwe, R. Biggs, and R. Preiser, "A framework for conceptualizing and assessing the resilience of essential services produced by socio-technical systems," *Ecology and Society*, vol. 23, no. 2, p. 12, 2018.
- [110] T. M. Shortell, Ed., *INCOSE Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities*, John Wiley & Sons, 4th edition, 2015.
- [111] N. M. El-Gohary and T. E. El-Diraby, "Domain ontology for processes in infrastructure and construction," *Journal of Construction Engineering and Management*, vol. 136, no. 7, pp. 730–744, 2010.
- [112] A. Levenchuk, "Sistemnaya kholarkhiya," *Sistemnoye myshleniye, Izdatel'skiye resheniya*, pp. 89–117, 2018.
- [113] "The Resilience of the Electricity System, Science and Technology Committee of the House of Lords," <https://publications.parliament.uk/pa/ld201415/ldselect/ldsctech/121/12102.htm>, 2015.
- [114] H. Treasury, *The Green Book: Central Government Guidance on Appraisal and Evaluation*, HM Treasury, London, UK, 2018.
- [115] G. Wilson, T. M. Wilson, N. I. Deligne, and J. W. Cole, "Volcanic hazard impacts to critical infrastructure: A review," *Journal of Volcanology and Geothermal Research*, vol. 286, pp. 148–182, 2014.

## Research Article

# Simulating Environmental Innovation Behavior of Private Enterprise with Innovation Subsidies

Hongjun Guan,<sup>1</sup> Zhen Zhang,<sup>1</sup> Aiwu Zhao ,<sup>2</sup> and Shuang Guan<sup>3</sup>

<sup>1</sup>School of Management Science and Engineering, Shandong University of Finance and Economics, Jinan 250014, China

<sup>2</sup>School of Management, Jiangsu University, Zhenjiang 212013, China

<sup>3</sup>Courant Institute of Mathematical Sciences, New York University, USA

Correspondence should be addressed to Aiwu Zhao; aiwuzh@ujs.edu.cn

Received 4 January 2019; Revised 11 April 2019; Accepted 24 April 2019; Published 13 May 2019

Guest Editor: Seyedmohsen Hosseini

Copyright © 2019 Hongjun Guan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

By reducing innovation costs, innovation subsidies can help private enterprises convert their production modes to green production. Based on method of computational experiment in social science, we construct a dynamic model for environmental innovation behaviors of private enterprises to simulate their evolution process in different market mechanisms, product competitions, and innovation subsidies and explore the impact of different subsidy modes on environmental technological innovation behaviors. The experimental results show that, under actions of multiagents, the combination of market subsidy and technology transformation subsidy can achieve the highest utilization efficiency of subsidy funds. However, when level of innovation technology is low, the innovation process should be subsidized at the same time to improve the competitiveness of innovative products. Besides, according to the level of innovation technology, flexible innovation subsidy combinations can be adopted to optimize subsidy in the different stages. The experimental results are of great significance for increasing efficiency of innovation subsidy funds and promoting green sustainable development of private enterprises.

## 1. Introduction

Due to the uncertainty and dual externalities of environmental technological innovation, for enterprises, the private return is less than the social return. Therefore, the environmental technological innovation levels of enterprises purely guided by market mechanism are bounded to be lower than the optimal level of society. Therefore, besides the two-standard demand-driven and technology-driven factors [1], the government's innovation policies have become an important driving force to stimulate the innovation vitality of enterprises [2]. It is a common policy for innovative countries to give some direct subsidies or tax incentives to technological innovation practices encouraging enterprises to develop new technology researches and developments [3, 4]. As a technology catching-up country, China government has also used R&D subsidy as a major policy to encourage enterprises innovate independently [5].

Compared with state-owned enterprises or foreign-funded enterprises with strong capital and technical strength,

the private enterprises, mainly small and medium-sized enterprises, have weak flexibility in choosing alternative production modes. Li Hongxia (2014) [6] constructed a green technology preference model for private enterprises with comparison of the average innovation intensities of private enterprise, state-owned enterprises, and foreign-funded enterprises. It was found that, due to high cost of machine renovation and technological innovation, the small-scale enterprises had weak impetus to change to green growth mode. Although the practices of European countries and the US have proved that environmental tax can promote enterprises to abandon the traditional production modes through continuous technological innovation and realize the switching to green production mode, tax may increase the burdens of private small and medium enterprises. It is even more unfavorable for private enterprises to adopt green innovation mode. Rather, the financial subsidy for green technological innovation can reduce the innovation cost of private enterprises, which is of great significance for the sustainable development of private enterprises.

By analyzing the data of the first national economic census and comparing the efficiencies of supports on direct subsidy and tax preference to enterprise innovation, Jiang Jing (2011) [7] found that the direct subsidy policy can significantly improve the R&D intensity of domestic enterprises. At the same time, based on panel data of 28 provinces in China, Fan Qi and Han Minchun (2011) [8] found that government innovation subsidy has a significant impact on improving national and regional independent innovations and the effect of subsidy on innovations in relatively developed regions is higher than that in relatively undeveloped regions. As for the impacts of different subsidy modes on innovations, Sheng Yanchao (2008) [9] used a three-stage game model to find that the mode of innovation subsidy on products is more effective than the mode of innovation subsidy on inputs under the government intervenes into the innovation system of technology alliance. However, through the study of asymmetric Cournot game model, Chen Lin and Zhu Weiping (2008) [10] found that innovation input subsidy, represented by “three fees for scientific research,” did not significantly stimulate the growth of innovation output of the whole society; thus the effect of subsidy policies was somehow uncertain.

Although subsidy policies play an important role in technological innovation of enterprises, their actual performance is affected by many factors such as subsidy modes and external atmosphere. In fact, in addition to the external atmosphere impact, the heterogeneity in microlevel of enterprises also affects their attitudes towards environmental innovation technology. However, most of the existing studies are based on panel data to carry out empirical research or build a mathematical model with government and enterprises as the two sides of the game, which only analyze the relationship between government subsidies and technological innovation performance from the macrolevel. Enterprise environmental technological innovation is a complex process involving policy support, market mechanism, and product competition. Enterprise innovation decision-making depends on many internal and external factors, such as enterprise nature, capital situation, risk attitude, market expectation, and innovation policies. Especially, for private enterprises with small scale and flexible business model, macrolevel analysis is difficult to reveal the impact mechanism of policy changes on the microlevel of private enterprises, while grasping the motivation of environmental technological innovation in microlevel has a far-reaching significance for private enterprises to turn to green technology.

Based on empirical data, this paper constructs a dynamic simulation model of environmental technological innovation behaviors of private enterprises and uses social science computational experiments to dynamically simulate the interaction mechanism between environmental technological innovation behavior of private enterprises and external driving forces [11, 12]. The purpose of this paper is to clarify the motivation and influence mechanism of private enterprise's environmental technological innovation behaviors, provide microtheoretical support for environmental policy makers, promote private enterprises to consciously adopt environmental innovation technology, and adopt green sustainable development.

## 2. Model Construction

*2.1. Scenario Description.* The actual prototype of this model is a kind of private chemical enterprises. In the initial stage of the system, all enterprises adopt traditional technology and use an organic solvent in the production process. The production process is mature, cost is low, and the quality is good, but the VOC emission level is high. Although the emission reduction can be achieved by means of end-treatment, it is limited by technical means and the emission reduction effect is unsatisfactory. Environmental innovation technology uses a certain green solvent resulting from the low VOC emission, but the use of green solvent requires certain equipment input and production process reform. Because the production process is not yet mature, the production cost is higher and the product quality is not as good as traditional ones. But, through technological R&D and reformation, the production process can be continuously improved to reduce production costs.

The purpose of innovation subsidy is to induce enterprises to consciously choose environmental innovation technology. Sun Xiao Hua et al. (2014) [13] found that consumers' heterogeneity preference provides niche market for new products and plays an important role in industrial evolution. Therefore, the study of different policy efficiency needs to combine many complex self-correlation evolution mechanisms, such as enterprises competition and market choices. Because of the reasons that empirical methods are difficult to find comparative samples of different policy backgrounds, this paper adopts computing experimental method in social science, refers to some designed ideas of multiagent model constructed by Afaroui (2014) [14] and Liu Xiao Feng (2013) [15], and builds computational experimental model based on real prototype to simulate environmental technological innovation processes in complex environment of innovation subsidy, market mechanism, and enterprise competition. The proposed model emphasizes on dynamics and disequilibrium processes from an evolutionary perspective. It draws on basic principles of the evolutionary theory of technological change [16, 17].

The model mainly includes two kinds of subjects: production enterprises and consumers. Consumers choose products from different production enterprises based on their preferences, while enterprises can freely choose product technology routes according to their decision rules. In order to better observe the trajectories of enterprises' environmental technological innovation under different innovation subsidies, this model takes into account the complexity of the real system as far as possible with abstraction and simplification.

*2.2. Basic Hypotheses.* According to the scenario description in Section 2.1, the basic hypotheses of the system are as follows.

(1)  $T_1$  represents the traditional technology of using organic solvents;  $T_2$  represents the environmental innovation technology of using green solvents. The products produced by the two technologies are technology products  $T_1$  and technology products  $T_2$ . According to Lancaster (1971) [18], each product is described by three attributes: price (related

to production costs), quality (representing technical performance), and VOC emissions (representing environment performance). The prices and qualities of the two products are different. The VOC emission per unit in the production process is also different. Moreover, after technological transformation, the lowest production cost, the highest product quality, and the minimum VOC emissions are also different.

(2) There are  $m$  manufacturing enterprises in the system, which are represented by enterprise number  $i$  ( $i=1,2,\dots,m$ ). At the initial stage, all enterprises adopt technology  $T_1$ , but, because technology  $T_2$  represents the direction of future development, according to the level of enterprise  $i$ 's attention to technology  $T_2$ , a certain proportion of R&D input is applied to the early stage researches and developments of technology  $T_2$ . When certain conditions are met, enterprises may begin to adopt technology  $T_2$  for formal production, but technology  $T_1$  can coexist at the same time, until a certain condition is reached, and enterprises would abandon technology  $T_1$ . The adoption threshold of technology  $T_2$  is different from the elimination threshold of technology  $T_1$ . In order to ensure stable total number of production enterprises in the system and maintain the free competition pattern, it is assumed that when the loss of production enterprises reaches a certain level, they will withdraw from the market, while new entrants will enter the market, regardless of the situation of enterprises obtaining loans and other external funds.

(3) There are  $n$  consumers in the system, which are represented by number  $j$  ( $j=1,2,\dots,n$ ) of the consumer agent. Assuming that the product is a nondurable necessity, consumers need to buy one product every cycle. Because consumers do not know the details of production process, according to model of Zeppini et al. (2014) [19], consumers would mainly consider product price, performance, consumption habits, and product reputation in the diffusion of innovative products and have certain social imitation ability. Therefore, it is assumed that consumers choose products of different technologies according to product price and quality and different consumers may have different preferences for product price and quality. In addition, purchase decisions are influenced by other consumers. At the same time, considering that consumers have certain path dependence attributes under the influence of consumption habits, consumers would still opt to the original enterprise when the price and quality of the current production enterprises are within the tolerable range. Consumers have different tolerances for product price and quality.

(4) The product is a constant reward type. That is to say, the production efficiency will not increase with the expansion of production scale. Only by technological transformation can the production cost be reduced [20]. Therefore, it is assumed that product pricing is based on production cost,  $P = C(1 + \mu)$ , where  $P$  is the product price,  $\mu$  is the producer's satisfactory profit level (considering the producer's bounded rationality), and  $C$  is the production cost. We assume that all production enterprises have the same satisfactory profit level and can reflect the product price level through the production efficiency level of the production enterprises.

### 2.3. Rules of Agent's Behavior

**2.3.1. Rules of Consumer Behavior.** The decision model of consumers is based on previous theoretical works on evolutionary demand [21]. Bounded rationality characteristic of customer is embodied by the comparison of specific threshold and imperfect information and some routines when they decide to purchase a product and to keep or leave original enterprise product [22]. Consumers consider both price and quality factors when choosing products and only when they reach the highest price affordability and the lowest quality requirements will they make out purchase decision. At the same time, different consumers may have different preferences for price and quality and are influenced by other consumers' choices [23]. Therefore, the effect function of consumers' choice of products is as follows:

$$U_{k,i,t}^j = \left[ (A - P_{k,i,t}) \times (Ms_{i,t-1} + \mu(0, 0.1)) \right]^{\text{inf}} p_j^p \times \left[ (X_{k,i,t} - B) \times (Ms_{i,t-1} + \mu(0, 0.1)) \right]^{\text{inf}} p_j^x \quad (1)$$

Among them,  $j$  is the consumer number,  $k$  ( $k=1,2,3,\dots$ ) is the product type (traditional technology products or environmental innovation technology products),  $i$  is the manufacturer number, and  $t$  is the simulation cycle. Assume that the total evolution cycles of the system are  $T$ ; then each cycle is expressed by  $t = 1, 2, 3$ , respectively.  $A$  is the highest price that consumers can afford,  $B$  is the lowest quality requirement that consumers can accept, and  $P_{k,i,t}$  and  $X_{k,i,t}$  are the price and quality level of product  $k$  of enterprise  $i$  in  $t$  cycle, respectively.  $Ms_{i,t-1}$  is the market share of enterprise  $i$  in the previous cycle and  $u(0, 0.1)$  is a random number between 0 and 0.1. It reflects the influence of other uncertain factors in the market and avoids the situation that the effect is zero when market share is empty.  $\text{inf}$  is interpreted as group psychology effect [24], reflecting consumers' imitation behavior;  $P_{k,i,t}$  and  $X_{k,i,t}$ , respectively, reflect consumers' preferences for product price and quality;  $p_j^p + p_j^x = 1$ .

When choosing a product at the first time, consumers may determine their choice probability according to the effect function of each product and randomly select the product. In the follow-up cycle, consumers firstly observe the lowest price and the highest quality of all current enterprise products according to the principle of path dependence. When the quality-price ratio of the original producer's products is within the tolerance of consumers, the consumers would select the original production enterprise; otherwise, the product selection probability is determined according to its product effect function and the product is randomly selected.

**2.3.2. Rules of Conduct for Manufacturing Enterprises.** The decision model of enterprises is based on the combination of economic theory and evolutionary theory and observation [14]: budget, mark-up pricing, R&D allocation, technology portfolio, and innovation process. Meanwhile, enterprises are bounded rationality [22]. They choose their technology portfolio by considering specific thresholds. According to the hypotheses, in the initial stage of the system, all enterprises

have the same scale and capital status. In each simulation cycle, enterprises gain profits from production sales and carry out technology transformation to improve product competitiveness. This model does not consider other ways to obtain funds, such as loans. The total disposable capital of the producer in each cycle is expressed as

$$BG_{i,t} = BG_{i,t-1} + \Pi_{i,t-1} - RD_{i,t-1} \quad (2a)$$

Among them,  $\Pi_{i,t-1}$  and  $RD_{i,t-1}$  are profit and R&D expenditure of enterprises in the last period.

For new technology  $T_2$  adopters, additional technology switch cost (such as equipment investment, and staff training cost) is required.

$$BG_{i,t} = BG_{i,t-1} + \Pi_{i,t-1} - RD_{i,t-1} - SC_{i,t} \quad (2b)$$

$SC$  is the related switching cost for the first adoption of technology  $T_2$ .

As mentioned before, product price can be deduced from production cost by applying a producer's satisfactory profit level  $\mu$  (also called as mark-up rate) as  $P = C(1+\mu)$ . Therefore, the profit formula for each production cycle is as follows:

$$\Pi_{i,t} = (\mu \times C_{i,t} \times Q_{i,t}) - FC \quad (3)$$

Among them,  $\mu$  is the producer's satisfactory profit level,  $C_{i,t}$  is the production cost,  $Q_{i,t}$  is the sales of products, and  $FC$  is the fixed cost.

(1) *Entry/Exit Rules for Manufacturing Enterprises.* When the disposable capital of a manufacturing enterprise is less than a certain level, the enterprise declares bankruptcy and withdraws from the market. At the same time, a new manufacturing enterprise enters the market as Van der and Brouillat (2015) [25] suggested. This is to maintain a constant number of enterprises over the whole time period. The technological route adopted by new production enterprises would imitate an existing enterprise in the system and the probability of choosing the imitated target is based on the market share of each enterprise. New enterprises imitate the technological route of target enterprises and the learning absorptive capacity is described as a random number between 0.8 and 1.2. This enables the new entrant to underperform or overperform in comparison with the imitated firm at a reasonable degree. Sensitive tests found that excessively low learning absorptive capacity would lead to exit from market faster than existing enterprises while excessively high setting would lead to outstanding performance. The price, quality, and VOC emission of products are multiplied or divided by random numbers (multiply positive index and divide negative index) on the basis of imitated enterprises' product indicators, so they can be lower or higher than those of the imitated enterprises. The initial disposable capital  $BG$  and fixed cost  $FC$  of the new enterprise are similar to those of other enterprises at the beginning. The switch cost  $SC$  of technology  $T_2$  and knowledge  $K$  would take the industries' averages.

(2) *Technical Route Selection Rules for Manufacturing Enterprises.* Innovation is an endogenous and uncertain process.

In fact, enterprises cannot know perfectly the results of their R&D activity. Therefore, the proposed model considers a stochastic process of innovation: most behavioral parameters are randomly drawn, the accumulation of knowledge that results from technology watch on  $T_2$  is stochastic, etc. Firstly, each enterprise calculates its perception of technology  $T_2$  maturity in a given cycle:

$$ADindex_{i,t}^{T_2} = K_{i,t-1} \times Ms_{t-1}^{T_2} \quad (4)$$

$Ms^{T_2}$  represents the total market share of technology products  $T_2$ , and  $K$  represents the knowledge accumulation of technology  $T_2$  acquired by enterprises through technology research and development. It is obvious that the possibility of adopting technology  $T_2$  depends on the knowledge accumulation of technology  $T_2$  and market diffusion of technology products  $T_2$ . When the enterprise considers the fact that the maturity of technology  $T_2$  is greater than a certain degree (one of the attributes of the enterprise: technology  $T_2$  adoption threshold), the enterprise checks whether there is enough disposable capital to support the new technology transformation. When  $BG_{i,t} \geq SC_{i,t}$ , the enterprise would formally adopt technology  $T_2$  for production.

Adopting technology  $T_2$  does not necessarily mean abandoning technology  $T_1$ . It is assumed that technologies  $T_1$  and  $T_2$  can coexist in the same enterprise. Whether or not to abandon technology  $T_1$  depends on the proportion of product income of technology  $T_2$  in total enterprise income:

$$Share_{i,t}^{T_2} = \frac{P_{i,t-1}^{T_2} \times Q_{i,t}^{T_2}}{\sum_{k=1}^2 (P_{i,t-1}^{T_k} \times Q_{i,t}^{T_k})} \quad (5)$$

When the proportion of technology  $T_2$  reaches the threshold of enterprises abandoning technology  $T_1$ , enterprises will abandon technology  $T_1$  and adopt technology  $T_2$  exclusively. This threshold also reflects the enterprise's risk attitude towards technology  $T_2$ . The higher the threshold value is, the more conservative the enterprise is and the possibility of technology  $T_1$  being abandoned is less and vice versa.

(3) *Rules of Enterprise R&D Activities.* Each enterprise improves the performance of products through R&D activities every cycle. The investment amount of R&D is as follows:

$$RD_{i,t} = \delta \times BG_{i,t} \quad (6)$$

Among them,  $\delta$  is the investment ratio of R&D, on the premise that the enterprise's current disposable capital  $BG_{i,t} > 0$ .

The R&D investment of enterprises is proportionally applied to the R&D of technology  $T_1$  and  $T_2$ :

$$RD_{i,t}^{T_1} = \delta_1 \times RD_{i,t} \quad (7a)$$

$$RD_{i,t}^{T_2} = (1 - \delta_1) \times RD_{i,t} \quad (7b)$$

Among them,  $\delta_1 \in [0, 1]$  for enterprises that only adopt technology  $T_2$ ,  $\delta_1 = 0$  for enterprises that only adopt

technology  $T_1$ , and  $(1-\delta_1)$  represents enterprises' initial R&D and learning of technology  $T_2$ . That is,  $RD_{i,t}^{T_2} = RDwatch_{i,t}$ . Because the progress in R&D and learning of technology  $T_2$  is uncertain and is related to funding, therefore, the knowledge accumulation needs to meet the following conditions:

$$1 - e^{-\alpha_w \times RDwatch_{i,t}} \geq u(0, 1) \quad (8)$$

where  $\alpha_w$  is a model parameter, which determines the speed of knowledge accumulation for the current technology.  $u$  is evenly and randomly distributed in  $[0,1]$ , reflecting the uncertainty of innovation activities in the real world. The closer to 1, the more difficult it is to satisfy condition (8). If the conditions are met, it means that R&D activities have achieved phased results, the knowledge accumulation of technology  $T_2$  increases, and the switching cost of technology  $T_2$  decreases. This is shown as

$$K_{i,t} = K_{i,t-1} + \alpha_k \times u(0, 1) \times (K_{\max} - K_{i,t-1}) \quad (9a)$$

$$SC_{i,t} = SC_{i,t-1} - \alpha_{SC} \times u(0, 1) \times (SC_{i-1} - SC_{\min}) \quad (9b)$$

$\alpha_k$  and  $\alpha_{SC}$  are model parameters;  $K_{\max}$  and  $SC_{\min}$  are the extreme values of knowledge accumulation  $K$  and switching cost  $SC$ .

The process of technology transformation in production activities is similar to that of predevelopment and learning of  $T_2$  technology. The success of technology transformation depends on whether conditions are met or not:

$$1 - e^{-\alpha_1 \times RD_{k,i,t}} \geq u(0, 1) \quad (10)$$

Among them,  $\alpha_1$  represents the speed of technology transformation and  $u$  reflects the uncertainty of innovation activities. If the technology transformation is successful, the attributes of the product will be updated as

$$X_{k,i,t} = X_{k,i,t-1} + \beta_1 \times u(0, 1) \times (X_{\max}^k - X_{k,i,t-1}) \quad (11a)$$

$$\begin{aligned} Cost_{k,i,t} &= Cost_{k,i,t-1} + \beta_2 \times u(0, 1) \\ &\times (Cost_{\max}^k - Cost_{k,i,t-1}) \end{aligned} \quad (11b)$$

$$\begin{aligned} Voc_{k,i,t} &= Voc_{k,i,t-1} + \beta_3 \times u(0, 1) \\ &\times (Voc_{\max}^k - Voc_{k,i,t-1}) \end{aligned} \quad (11c)$$

where  $u$  is a uniform random number,  $\beta_1$  is the product quality,  $\beta_2$  is the production cost, and  $\beta_3$  is the improvement efficiency of VOC emission.  $X_{\max}^k$ ,  $Cost_{\max}^k$ , and  $Voc_{\max}^k$  are the extreme value that technology  $K$  can reach in all aspects of product performance. When the actual level of a given product property comes closer to the limit of what is achievable with specific technology, a given R&D expenditure will achieve less and less progress. The workflow of each production cycle is shown in Figure 1. This workflow reflects the basic principle of evolutionary theory of technological change, such as path-dependency, incremental versus radical innovation, and innovation risk, being cumulative and localized in a certain direction. Innovation is firm-specific which leads to technological diversity and heterogeneous performances. Consumer choices lead to heterogeneous demand, coevolution of firm strategies, and market structure.

**2.4. Parameter Settings.** To set the public parameters and individualized parameters for system simulation, we should consider the objective and realistic prototypes as far as possible. For the parameters that are difficult to quantify in reality (such as consumers' preference for product price and quality) [26], we would use the basic model (no innovation subsidy situation) to carry out "virtual-real linkage." By adjusting the parameters repeatedly, we observe the intermediate results and the final result to make the outcomes tally with the reality. After determining the parameters of the basic model, the policy parameters are introduced to observe the impact of the policies on the simulation results.

The main variables and their initial assignment rules in the system are shown in Table 1.

Empirically based parameters in Table 1 are determined basically in [14] and data from <http://www.cefic.org/Facts-and-Figures/>. Some parameters are adjusted according to the questionnaire survey of private enterprises in China (e.g., reduce initial disposable capital from 15 to 12). The questionnaire is related to enterprise environmental innovation behavior and innovation performance, including "R&D investment of enterprise environmental innovation," "factors affecting enterprise environmental innovation," "factors affecting enterprise environmental innovation," "economic performance of enterprise environmental innovation." Simulation training parameters are adjusted by the comparison result of benchmark model and empirical investigation. For example, although most enterprises agree that environmental innovation is the best way to overcome current environmental barriers and improve their competitiveness, only about 10% of enterprises with strong technical strength are willing to carry out environmental innovation considering a long-term interest. The reason of low adoption comes from immature technology, inadequate product competitiveness of environmental technology, and high expenditure of additional equipment investment, personnel training and market development costs, and so on. Therefore, related parameters are trained to coincide with this empirical result.

### 3. Simulation Experiment and Result Analysis

Based on the situation of without innovation subsidy, this model mainly analyses the effects of three subsidy policies on environmental technological innovation behavior of private enterprises: process subsidy, technology transformation subsidy, and market subsidy of environmental innovation products. In order to compare the effects of each subsidy policy, this paper designs the following scenarios for comparative experiments:

O: no innovation subsidy

P: subsidies for enterprise in the process of environmental technological innovation.

T: subsidies for enterprises to adopt environmental innovation technology when requiring switching costs, such as investment in new equipment

M: price subsidies for environmental innovative products

TABLE 1: Main variables and initial assignment rules.

Variables/parameters	Implication	Assignment	Assignment rules
$m$	Number of enterprises	10	Fixed
$n$	Number of consumers	200	Fixed
$A/B$	Consumers' maximum affordable Price/minimum Acceptable quality	6/20	Simulation training
$p^p$	Consumer price preference	[0.1,0.9]	Consumer random
inf	Consumer conformity effect	0.05	Simulation training
$Tol$	Consumer tolerance	[0.8,1.2]	Simulation training
$X_{max}^1/Cost_{min}^1/Voc_{min}^1$	Technology product T1 maximum Quality/minimum Cost/minimum VOC emission	100/0.4/2	Based on the empirical analysis
$X_{max}^2/Cost_{min}^2/Voc_{min}^2$	Technology product T2 maximum Quality/minimum Cost/minimum VOC emission	100/0.4/0	Based on the empirical analysis
$X_0^1/Cost_0^1/Voc_0^1$	Technology product T1 Initial quality/initial cost/initial VOC emission	100/0.5/23	Based on the empirical analysis
$X_0^2/Cost_0^2/Voc_0^2$	Technology product T2 Initial quality/initial cost/initial VOC emission	32/3/0.05	Based on the empirical analysis
$B_0$	Initial disposable capital of enterprise	12	Based on the empirical analysis
$\mu$	Satisfied profit level	0.5	Based on the empirical analysis
$FC$	Fixed cost	2	Based on the empirical analysis
$Ad_{min}$	Threshold of adoption technology T2	[0,2]	Enterprise random
$SC$	Initial switch cost	20	Based on the empirical analysis
$Aband_{min}$	Threshold for abandoning technology T1	[0.5,1]	Enterprise random
$\delta$	R&D investment ratio	0.2	Based on the empirical analysis
$\delta_1$	R&D investment ratio of technology T1	[0,1]	Enterprise random
$\alpha_w$	Knowledge speed	0.2	Simulation training
$d_1$	Speed of technology transformation	0.25	Simulation training
$K_{max}$	Maximum cumulative value of knowledge	1	Based on the empirical analysis
$SC_{min}$	Minimum switching cost	10	Based on the empirical analysis
$\beta_1/\beta_2/\beta_3$	Product quality/production cost/improvement efficiency in VOC emissions	0.05/0.08/0.05	Simulation training

PT, PTM, PM, and TM: the combination subsidy schemes of the above three subsidies accordingly.

According to the above definition, we simulate the processes of enterprise environmental technological innovation in different scenarios and further analyze the efficiencies and effects of various innovation subsidy schemes.

**3.1. Path Analysis of Enterprise Environmental Technological Innovation on Single Subsidy Scenario.** In the scenario P, on the basis of the basic model, the government implements a 1:1 matching approach to subsidize the R&D investment of enterprises for environmental technological innovation. Therefore, the subsidy of enterprises  $i$  in the cycle  $T$  is described as follows:

$$Sub_{i,t}^{T_2} = (1 - \delta_1) \times RD_{i,t} \quad (12a)$$

In the scenario T, the government subsidizes switching costs of enterprises to adopt environmental innovation technology. Therefore, the subsidy of enterprises  $i$  in the cycle  $T$  is described as follows:

$$Sub_{i,t} = SC_{i,t} \quad (12b)$$

In the scenario M, the government subsidizes environmental innovative products in markets. After subsidization, the market average price of technology products T2 is equal to the market average price of technology products T1. Therefore, the subsidy of enterprises  $i$  in the cycle  $T$  is described as follows:

$$Sub_{i,t} = (P_{2,i,t} - avg(P_{1,i,t})) \times Q_{i,t}^{T_2} \quad (12c)$$

We set the total simulation cycle number to 300 and repeat simulation 50 times in the different scenarios and

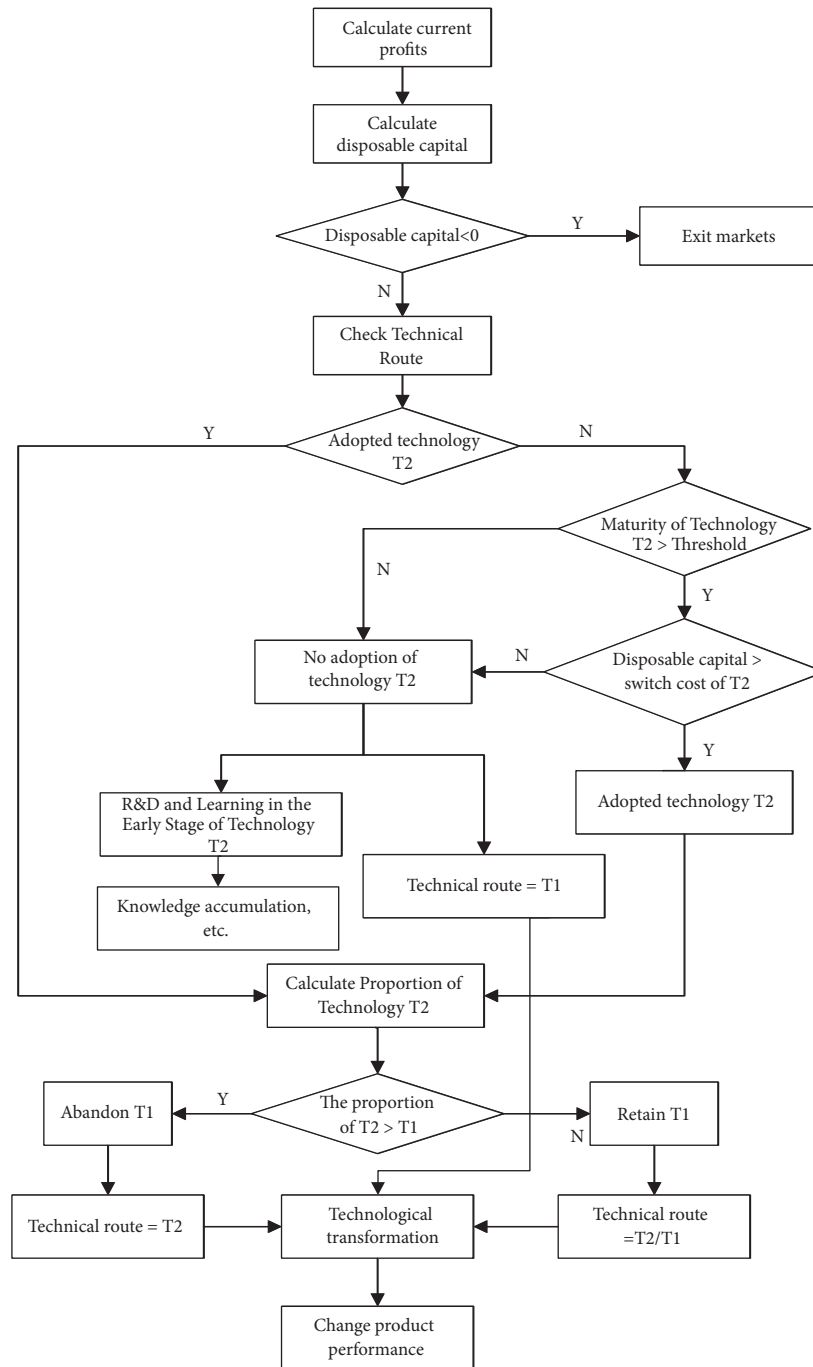


FIGURE 1: The workflow of each production cycle.

then take the average of each simulation results. In a single subsidy scenario, the adoption of environmental innovation technology by enterprises in each cycle is shown in Figure 2.

Figure 2 shows that different subsidy policies have different impacts on the technological route of enterprises: (1) compared with the P scenario and the O scenario, the number of enterprises adopting combination technology has increased significantly, but no enterprises adopt technology T2 completely; (2) in the scenario T, from 120<sup>th</sup> cycle on,

some enterprises have completely switched to technology T2, and after 150 cycles although most enterprises still retain technology T1, they have partially or completely adopted technology T2; (3) in the scenario M, although some enterprises begin to completely switch to technology T2 after 120 cycles, 80% of them still do not adopt technology T2. From the perspective of enterprise technology route only, subsidy policy in the scenario T is the most effective mode to diffuse environmental innovation technology.



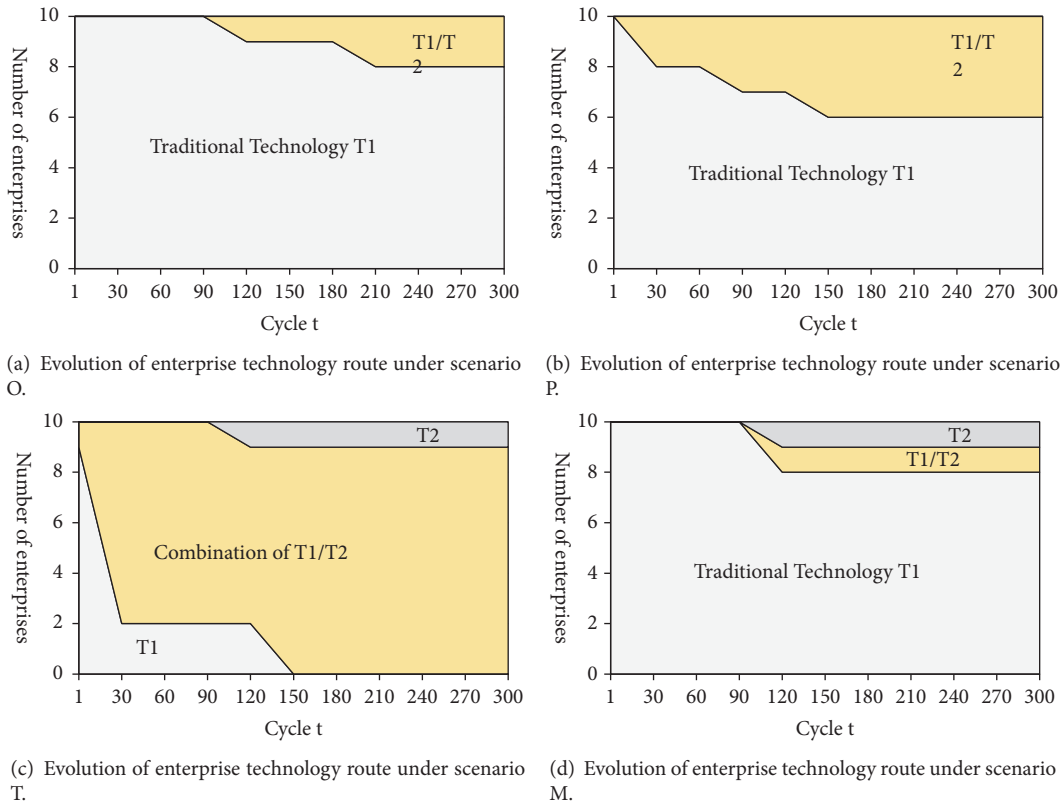


FIGURE 2: Evolution of enterprise technology route under single subsidy policy scenarios.

The market share and competitiveness of products T2 in the different scenarios are shown in Figure 3.

As can be seen from Figure 3(a), although there are still a large number of enterprises using only technology T1 in the scenario M of Figure 2(d), the market share of technology products T2 is still larger than those of other scenarios. In the scenario T, although all enterprises adopt environmental innovation technology, the lack of competitiveness in product T2 regarding quality and price (Figure 3(b)) leads to a low market share (lower than P and M scenarios) because the government only subsidized the technology switching cost of enterprises. Thus, from Figure 2(c), most enterprises still retain technology T1. In the scenario P, with the support of enterprise innovation process subsidy, it can be seen from Figure 3(b) that the competitiveness of technology products T2 is rising faster. Therefore, although the adoption rate of technology T2 in Figure 2(b) is low, the market share is higher than that in the scenario T. Overall, the three subsidy policies all have positive impacts on the increase of market share of technology products T2, but the overall competitiveness of products T2 is less than that of products T1 (less than 50). The market diffusion effect of products T2 in cycle 300 is not ideal; even with the highest market share in the scenario M, the diffusion effect in cycle 300 is only about 12%. The reason is in the scenario M; although the price of products T2 has been subsidized, the competitiveness of products T2 is weak because the quality of products T2 is not as good as that of technology products T1. In

the scenario P, although the performance of products T2 improves rapidly, the actual adoption rate of technology T2 is low due to the lack of sufficient technology transformation funds. In the scenario T, although technology transformation is subsidized, it is difficult for technology products T2 to win the market due to their slow performance improvement and weak competitiveness.

*3.2. Path Analysis of Enterprise Environmental Technological Innovation on Combination Subsidies Scenario.* In the single subsidy scenario, the market diffusion effect of technology products T2 is not ideal. Therefore, various combinations subsidies are considered: PT, PTM, PM, and TM. In the combination subsidy scenario, the innovation subsidy obtained by enterprises in each cycle is accumulated with individual subsidy schemes in each scenario. The adoption of environmental innovation technology by enterprises in each cycle is shown in Figure 4.

From Figure 4, it can be seen that the proportion of adopting technology T2 in the combination subsidy scenarios is higher than those in the single subsidy scenarios. Especially in the scenario PTM, all enterprises adopt technology T2 partially or completely after 30 cycles, and 70% of them adopt technology T2 completely. In the scenario TM, although 10% of enterprises did not adopt technology T2 in the 300<sup>th</sup> cycle, 60% of enterprises completely adopt technology T2. In the scenario PM, the number of enterprises adopting technology T2 is also significantly higher than that in the scenario P or

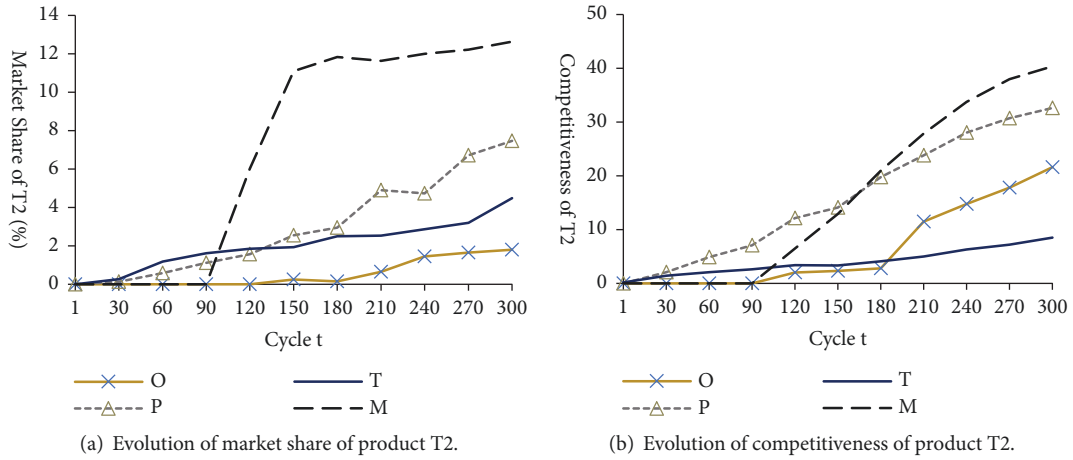


FIGURE 3: Evolution of market share and competitiveness of product T2 under single subsidy scenario. *Note.* Product competitiveness is calculated according to consumer effect function (1). The calculation results are standardized with the maximum value of current consumer effect as 100. The competitiveness of product T2 is the average value of all standardized products T2 in the market.

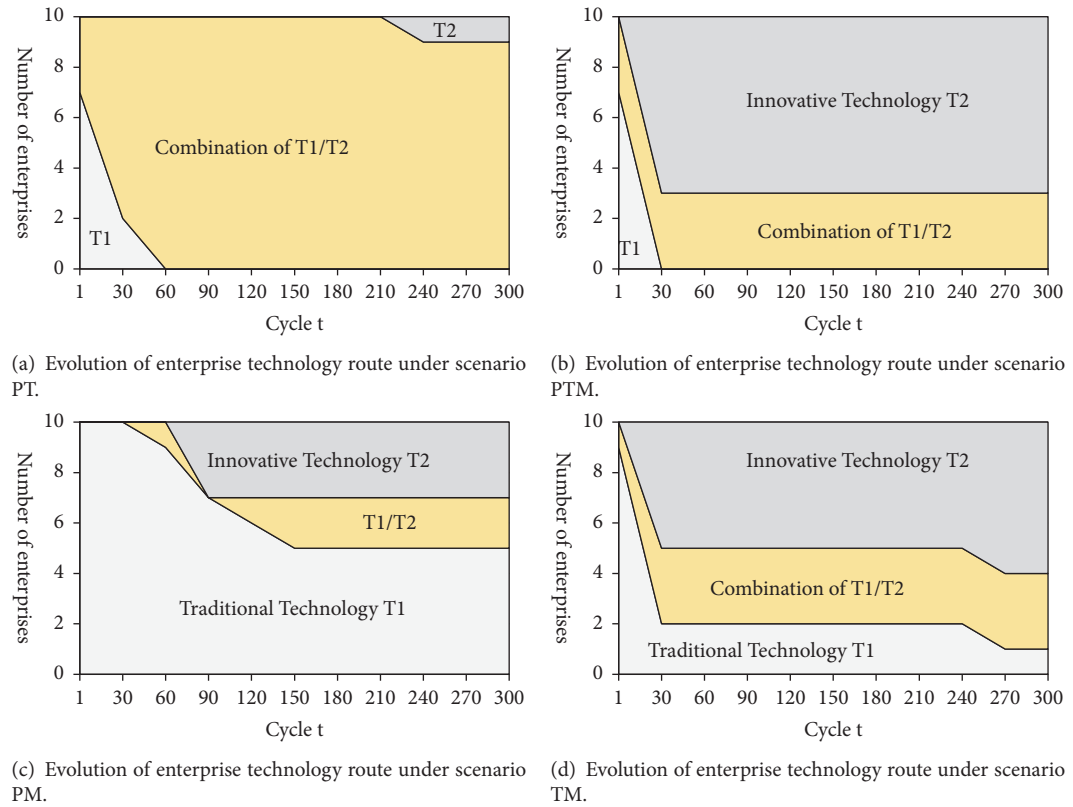


FIGURE 4: Evolution of enterprise technology route under combination subsidy scenarios.

M alone. In the scenario PT, although all enterprises begin to adopt technology T2 partly or completely after 60 cycles, enterprises which adopt technology T2 completely only appear after the 240<sup>th</sup> cycle and are only composed of 10%. Compared with other combination policies, the adoption rate of enterprise environmental innovation technology in the scenario PT is the lowest.

The market share and competitiveness of products T2 in the different combination subsidies scenarios are shown in Figure 5.

Figure 5(a) shows that the market share of technology products T2 is ordered as PTM > TM > PM > PT, which corresponds to the technological route adopted by enterprises in the combination subsidy policies in Figure 4. However,

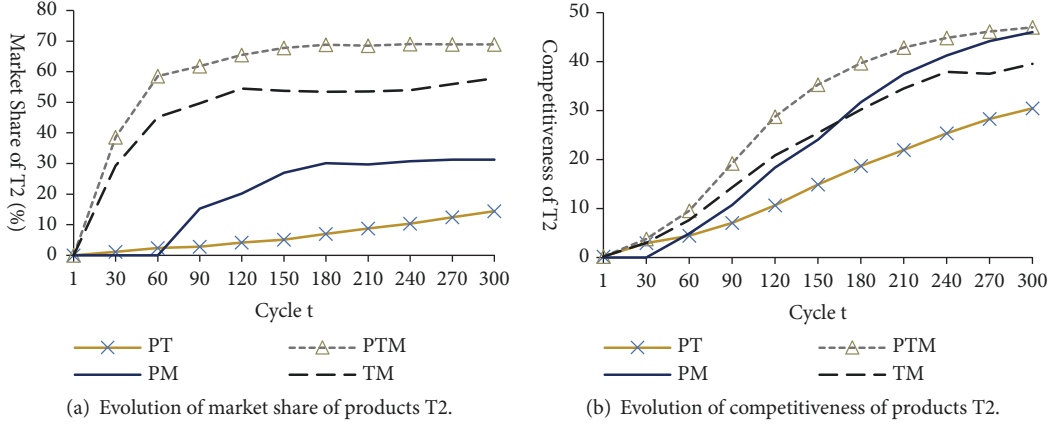


FIGURE 5: Evolution of market share and competitiveness of products T2 under combination subsidy scenarios.

as shown in Figure 5(b), the competitiveness of technology products T2 in the 300<sup>th</sup> cycle is ordered as PTM > PM > TM > PT; that is to say, the subsidies are cancelled after 300 cycles, because the competitiveness of technology products T2 in the scenario PM is greater than that in the scenario TM. The market share of technology products T2 will be PM > TM, which is due to the fact that PM subsidy mode in technology T2 innovation process is more beneficial for the improvement of product competitiveness than that of TM subsidy mode. In addition, because TM subsidy mode helps more enterprises to adopt technology T2 in advance by subsidizing enterprises to complete technology transformation, the market share of technology products T2 is higher than that of PM subsidy mode. In the scenario PT, although the innovation process subsidy benefits the competitiveness of products T2, the competitiveness of products T2 is still weaker than that of products T1. In the absence of market subsidy incentives, products T2 would have the lowest market share.

It can be seen that the combination of M policy (market subsidies for technological products T2), P policy (subsidies for enterprise environmental innovation process), and T policy (technology transformation subsidies for enterprises) can better promote the diffusion of environmental innovation technology.

**3.3. Analysis of Innovation Subsidy Efficiency in Different Scenarios.** From the above analysis, it can be seen that different subsidy policies have different effects on the adoption of technology T2 and the diffusion of products T2. Assume that subsidies needed to increase a unit share of products T2 in the different scenarios are described as follows:

$$U\_subsidy_t^{T_2} = \frac{T\_subsidy_t^{T_2}}{Sup_t^{T_2}} \quad (13)$$

$T\_subsidy_t^{T_2}$  is the total investment of innovation subsidy in the cycle T;  $Sup_t^{T_2}$ , based on the market share of products T2 in each cycle under no innovation subsidy, is the growth rate of market share of products T2 in different scenarios.  $U\_subsidy_t^{T_2}$  is the subsidy needed to increase the unit share

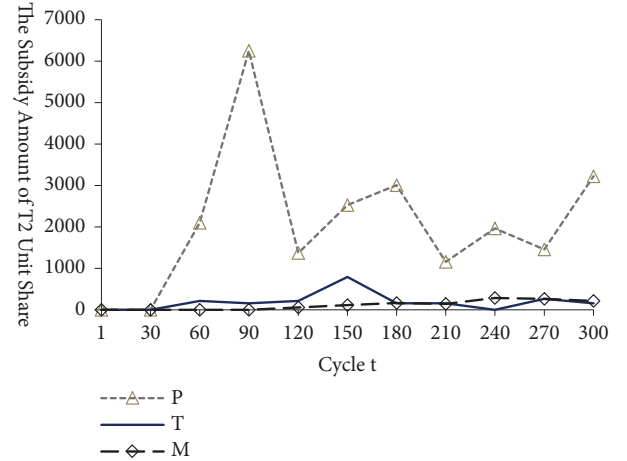


FIGURE 6: Evolution of innovation subsidy efficiency under single subsidy scenarios.

of products T2. Without considering the indirect effects of subsidy, such as improvement of the competitiveness of products t2, the reciprocal of innovation subsidy efficiency can be regarded as the work efficiency of innovation subsidy in different scenarios. The reciprocal of innovation subsidy efficiency in the single subsidy scenario is shown in Figure 6.

As can be seen from Figure 6, in the single subsidy scenario, the efficiency of innovation subsidy in the scenarios M and T is the highest, while that in the scenario P is the lowest.

The reciprocal of innovation subsidy efficiency in the combination subsidy scenario is shown in Figure 7.

As shown in Figure 7, in the combination subsidy scenario, the efficiency of innovation subsidy is the highest in the scenario TM and the lowest in the scenario PTM at the 300<sup>th</sup> cycle. Figures 4 and 5 show that TM subsidy effect is only second to PTM. Although PTM subsidy effect is the best, the efficiency of capital utilization is low; especially after the 180<sup>th</sup> cycle, it becomes the lowest. We may consider optimizing the PTM subsidy by stages. PTM can improve the

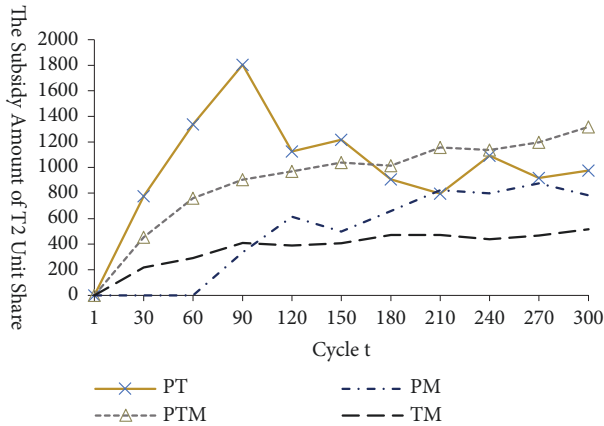


FIGURE 7: Evolution of innovation subsidy efficiency under combination subsidy scenarios.

competitiveness of products T2 in the early stage and improve the efficiency of innovation subsidy in the later stage.

#### 4. Conclusions

Taking the environmental technology innovation process of a private enterprise in a chemical industry as an example, through social science computational experiment method, this paper constructs a model of environmental technological innovation of private enterprises and simulates the processes of environmental technological innovation of enterprises in the different innovation subsidies scenarios. The proposed model can help to closely observe the dynamic innovation process and technology transformation process from a microperspective.

As we discussed, under the increasingly severe pressure of environment problem, more and more enterprises begin to pay attention to environmental innovation technology. Enterprises' original motivation of carrying out environmental innovation is the improvement of product competitiveness to obtain higher profits. However, the survey of enterprises' willingness to adopt environmental innovation technology found that although most enterprises agree that environmental innovation is the best way to overcome current environmental barriers and improve their competitiveness, only about 10% of enterprises with strong technical strength are willing to carry out environmental innovation considering a long-term interest. The reason of low adoption comes from immature technology, inadequate product competitiveness of environmental technology, and high expenditure of additional equipment investment, personnel training and market development costs, and so on. The price subsidy of environmental innovative products directly improves the price competitiveness of environmental innovative products. The process subsidy of environmental innovation helps to improve the performance of all dimensions of innovative products. Subsidizing the conversion costs of enterprises when they adopt environmental innovative technologies can help enterprises break through the bottleneck of capital during new conversion. Therefore, these subsidy modes are conducive

to the improvement of environmental innovation adoption. The proposed model reveals the influence mechanism of different subsidy modes on enterprises' environmental innovation behavior and the limitation of single subsidy mode from the microperspective. For example, when government provides market price subsidy for environmental innovative products, the market for environmental innovative products is growing, and more enterprises are encouraged to increase the R&D investment and actual production of environmental innovative technology. However, high cost of environmental technology conversion and the difficulty of private enterprise financing have become the bottleneck of transformation and upgrading of most private SMEs. Therefore, the combination subsidy would be the better way to improve the effectiveness of subsidy.

Simulation experiments under different scenarios show that efficiency of subsidized funds is related to the level of innovation technology. When innovation technology is not mature enough, policy should focus on innovation process subsidy to improve innovation technology as far as possible. On the contrary, the combination of product market subsidy and innovation technology conversion subsidy would achieve the highest efficiency of capital utilization. Therefore, based on different levels of innovation technology, flexible combination of innovative subsidy modes can be applied in different stages of technology development to optimize the efficiency and effect of subsidy funds. For example, in the early stage of innovation technology promotion, PTM subsidy portfolio policy can be used to improve the market competitiveness of T2 technology products. When innovation technology becomes more mature, TM subsidy portfolio should be used to help enterprises complete the replacement of new and old technologies and increase the market share of innovative products.

In practice, in order to ensure the flow and efficiency of the use of subsidized funds, more flexible specific subsidy modes can be adopted. For example, the combination of R&D input plus deduction policy and reward for innovative achievements can be employed as innovation process subsidies, green financial financing mode for innovation technology transformation can be used as technology conversion subsidy, and the combination of consumption guidance and market subsidy can be applied as environmental innovation product subsidy, and so on.

While designing and analyzing the model, several interesting ideas arise which nevertheless are neglected for the sake of clarity and simplicity. Some of these ideas deserve further development as they might develop into fertile new lines of research: (1) the model assumes that the scale and capital conditions of enterprises are the same at the initial stage of the system, without considering the different enterprise scales. (2) The model does not further discuss the multistage combination of subsidy policy, nor does it discuss the evolution of enterprises' environmental technological innovation behaviors after innovation subsidy policy. (3) Consumers' purchasing decisions only involve product prices, qualities, consumption habits, and conformity effects, without considering their preferences for product environmental attributes.

## Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

The authors are indebted to anonymous reviewers for their very insightful comments and constructive suggestions, which help ameliorate the quality of this paper. This work was supported in part by the National Natural Science Foundation of China (71704066).

## References

- [1] K. Rennings, "Redefining innovation - Eco-innovation research and the contribution from ecological economics," *Ecological Economics*, vol. 32, no. 2, pp. 319–332, 2000.
- [2] L. C. Z. Yong'an, "An empirical research on the impact of regional innovation policy on enterprise innovation efficiency," *Science Research Management*, vol. 35, no. 9, pp. 25–35, 2014.
- [3] C.-L. Lin and W.-M. Xu, "The best way of innovation subsidies with technology externality," *Studies in Science of Science*, vol. 30, no. 5, pp. 766–781, 2012.
- [4] Zhang J.-L. and L.-L. Li, "The study on the relationship between different R&D funding ways and the enterprise technology innovation stage," *Studies in Science of Science*, vol. 32, no. 11, pp. 1740–1746, 2014.
- [5] A. Tong-Liang, Z. Xing-Cai, and P. Jian-Cai, "The stimulating effects of R&D subsidies on independent innovation of chinese enterprises," *Economic Research Journal*, vol. 10, pp. 87–97, 2009.
- [6] H.-X. Li, "Green technological innovation and environmental tax policy of private enterprises," *Taxation Research*, vol. 3, pp. 12–15, 2014.
- [7] J. Jing, "Do public policies support firm's innovation? a comparative analysis on the subsidies and tax credit," *Science Research Management*, vol. 32, no. 4, pp. 1–8, 2011.
- [8] Q. Fan and M.-C. Han, "The impact of government R&D subsidy on the performance of national and regional independent innovations," *Journal of Industrial Engineering and Engineering Management*, vol. 25, no. 3, pp. 183–188, 2011.
- [9] Y.-C. Sheng, "Innovation subsidies or product subsidies: the government strategy choice of technical alliance," *Chinese Journal of Management Science*, vol. 16, no. 6, pp. 184–192, 2008.
- [10] L. Chen and W.-P. Zhu, "Policy effect of rebating export tax and subsidizing innovation in innovation incentive," *Economics Research*, vol. 11, pp. 74–87, 2008.
- [11] L. David, P. Alex, A. Lada et al., "Computational Social Science," *Science*, vol. 323, no. 5915, pp. 721–723, 2009.
- [12] Z.-H. Sheng and W. Zhang, "Computational experiments in management science and research," *Journal of Management Sciences in China*, vol. 14, no. 5, pp. 1–10, 2011.
- [13] S. Xiaohua and G. Shaorong, "Heterogeneity preferences, new technology and evolution of industry," *Journal of Systems Engineering*, vol. 29, no. 3, pp. 334–342, 2014.
- [14] N. Arfaoui, E. Brouillat, and M. S. Jean, "Policy design and technological substitution: investigating the reach regulation in an agent-based model," *Ecological Economics*, vol. 107, pp. 347–365, 2014.
- [15] L. Xiaofeng, S. Zhaohan, and D. Jianguo, "Evolutionary model of cleaner production technologies under product competition and consumer choice," *Journal of Management Science*, vol. 26, no. 6, pp. 25–34, 2013.
- [16] R. Nelson and S. G. Winter, *An Evolutionary Theory of Economic Change*, The Belknap Press of Harvard University, London, UK, 1982.
- [17] F. Malerba, R. Nelson, L. Orsenigo, and S. Winter, "'History-friendly' models of industry evolution: the computer industry," *Industrial and Corporate Change*, vol. 8, no. 1, pp. 3–40, 1999.
- [18] K. Lancaster, *Consumer Demand: A New Approach*. *Columbia Studies in Economics* 5, Columbia University Press, New York, NY, USA, 1971.
- [19] P. Zeppini, K. Frenken, and R. Kupers, "Thresholds models of technological transitions," *Environmental Innovation and Societal Transitions*, vol. 11, pp. 54–70, 2014.
- [20] Z. Aiwu, D. Jianguo, and G. Hongjun, "Simulation of enterprise environmental technology innovation given the context of environmental tax," *Journal of Management Science*, vol. 29, no. 1, pp. 40–52, 2016.
- [21] M. Valente, "Evolutionary demand: a model for boundedly rational consumers," *Journal of Evolutionary Economics*, vol. 22, no. 5, pp. 1029–1080, 2012.
- [22] R. Ariel, *Modeling Bounded Rationality*, The MIT Press, Cambridge, Mass, USA, 1997.
- [23] R. Cowan, W. Cowan, and P. Swann, "A model of demand with interactions among consumers," *International Journal of Industrial Organization*, vol. 15, no. 6, pp. 711–732, 1997.
- [24] H. Leibenstein, "Bandwagon, snob, and veblen effects in the theory of consumers' demand," *The Quarterly Journal of Economics*, vol. 64, no. 2, pp. 183–207, 1950.
- [25] A. Van Der Vooren and E. Brouillat, "Evaluating CO2 reduction policy mixes in the automotive sector," *Environmental Innovation and Societal Transitions*, vol. 14, pp. 60–83, 2015.
- [26] H. Guan, A. Zhao, and J. Du, *Research on Enterprise's Green Technology Innovation Behavior*, Economic Science Press, Pei King, China, 2017.

## Research Article

# Measuring Component Importance for Network System Using Cellular Automata

Li He , Qiyao Cao , and Fengjun Shang 

College of Computer Science, Chongqing University of Posts and Telecommunication, Chongqing 400065, China

Correspondence should be addressed to Li He; 372529646@qq.com

Received 13 December 2018; Revised 26 March 2019; Accepted 2 April 2019; Published 2 May 2019

Guest Editor: Md Sarder

Copyright © 2019 Li He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper concentrates on the component importance measure of a network whose arc failure rates are not deterministic and imprecise ones. Conventionally, a computing method of component importance and a measure method of reliability stability are proposed. Three metrics are analyzed first: Birnbaum measurement, component importance, and component risk growth factor. Based on them, the latter can measure the impact of the component importance on the reliability stability of a system. Examples in some typical structures illustrate how to calculate component importance and reliability stability, including uncertain random series, parallel, parallel-series, series-parallel, and bridge systems. The comprehensive numerical experiments demonstrate that both of these methods can efficiently and accurately evaluate the impact of an arc failure on the reliability of a network system.

## 1. Introduction

As a quantitative measure, reliability can be broadly interpreted as the ability of a system to perform its intended function. During the past ten years, a significant amount of research has been conducted to address reliability evaluation. Network reliability can be estimated using Bayesian approach [1], Monte Carlo simulation [2, 3], genetic algorithm [4], fault-tree analysis [5], etc. Obviously, all those methods apply numerical reliability or boundary value to indicate the reliability of network systems. However, two main questions must be answered for designing a network system. Question 1: Which component is the most important? Question 2: How does the importance of component impact the system reliability stability? For answering such questions, component importance measures must show the effect and rank in system design and preventive maintenance.

Determining the importance of components in complex networks is crucial. Several importance measure methods have been introduced in [6, 7], including Birnbaum measure, criticality importance, improvement potential, risk achievement worth, and risk reduction worth. Based on the fundamental component importance theory initially proposed by Birnbaum [6], there have been a number of approaches used to show component importance. Generally speaking,

the traditional component importance evaluation methods are classified into two kinds, one is on the components failure rates, and the other is without taking components characteristic into account. The first category method is mainly based on the graph theory, including reliability Boolean polynomial [7], minimum trees and their number [8], and minimal cut set. In [9], criticality importance measures for components with respect to system failure intensity and the total system failure count are presented. To evaluate reliability importance of components in a network system, Zio et al. [10] present generalized importance measures based on Monte Carlo simulation. Meanwhile, Wang et al. [11] introduce the failure critically index, restoring critical index and operational index. In [12], importance measures with respect to system failure intensity are developed and it also points out that the Barlow [11] importance only measures the contribution of a component as the last failure in a minimal cut set, not the total contribution. Contini et al. [13] also evaluate network system importance with respect to the system failure. However, an obvious shortcoming is that the impact of the component characteristic on the failure rate of network is not considered. Some examples, including the reliability of Boolean polynomial [7], minimum spanning tree [14, 15], minimum cut set and minimum path set [16], and fault-tree analysis [9], attempting to incorporate

more features of network topology consisting of multiple terminals and dependency between topology are researched. Meanwhile, simulation based on Monte Carlo method [10] often depends more on the convergence of probability than the number of network components; statistical error during reliability analysis may result in slow convergence for achieving acceptable accuracy in low probability estimations. Therefore, these methods depend on the model to decompose the network topology and calculate the reliability of network. And the complexity of calculation will increase by index level as the size of network grows. Although these methods have adequately considered the characteristic of component in network system, how to improve the efficiency of calculation to strength the practical of importance measurements is still a focus.

Recently, importance measure to estimate the effect of a component residing at certain states on the performance is proposed in [17]. Importance measure of components when the system may be reconfigured is designed [18]. Liu [19] presents a chance theory, which contains some basic concepts including chance measure, uncertain random variable, and chance distribution. Then, Gao and Yao [20] research the importance index of components in uncertain random systems; a concept of importance index on a component in uncertain random variable and Boolean system is proposed. At the same time, link component importance is analyzed in [21]. Component maintenance priority is used to select component for preventive maintenance. And a Monte Carlo-based method to generate probability distributions of the two metrics for all of the components of the network is proposed and a stochastic ranking approach based on the Copeland's pairwise aggregation is used to rank components importance in [22]. In addition, a strategy for solving the component placement problem by maximizing the information gain in terms of users' choices in [23] is proposed. At last, Zhu et al. [24] present a nonlinear binary programming model, which focuses on embedding the Birnbaum importance in heuristics and the method of dealing with more than one type of components.

However, to the best of our knowledge, none of the exiting classical importance methods based on Cellular Automata are directly applied to measure the impact of the component importance on the reliability of system. The Birnbaum importance, risk growth factor, and reliability stability to measure the importance of a component or a group of components are defined. A computing method of component importance (NEA) based on Cellular Automata is designed; in addition, a new measure method of reliability stability (NSA) is proposed in this paper. At last, the validities of NSA and NEA are proved by experiments. And it is also proved that the proposed algorithm NSA is more accurate in calculating the importance of the system components compared to the classical algorithm in [25].

The main contributions are as follows:

- (i) defining the Birnbaum importance, risk growth factor, and reliability stability;
- (ii) designing a model for measuring the network system component importance;

- (iii) designing a method for measuring the reliability stability of network system.

The rest of this paper is organized as follows. Three metrics of component importance are introduced in Section 2, and the validity of these measurements is verified in this section. In Section 3, we put forward our system models. In Section 4, we evaluate the component importance and describe our algorithms. In Section 5, the parameters of experiment are given and the performance of the algorithms is analyzed. The conclusion is drawn in Section 6.

## 2. Preliminaries

Let  $G = (V, E)$  be a network system, where  $V$  is the set of  $n$  nodes,  $V = \{v_i \mid i = 1, 2, \dots, n\}$ .  $E \subseteq V \times V$  is the set of  $m$  arcs,  $E = \{e_i \mid i = 1, 2, \dots, m\}$ .  $G$  is a directed, connected, acyclic graph which contains an initial node  $s$  and a terminal node  $t$ . In order to study the issue of this paper, there are some assumptions.

- (i) The state of each node and arc is statistically independent.
- (ii) The evaluation of network reliability  $R$  is considered so that the probability of the initial node  $s$  successfully connects to the terminal node  $t$ .

*2.1. Birnbaum Measure.* The significance of network system component importance is the influence degree of network system components (nodes or links) on network system connectivity, which can be expressed by the Birnbaum measure [9], denoting the importance of network system decided by the reliability of network system and component. The nature of this definition mathematically denotes the partial derivative on the reliability of component to the reliability of network system; that is, when the component reliability is changed, the network system reliability will be accordingly changed. For a network system with  $m$  components, its Birnbaum measure can be defined as

$$I_i^B(t) = \frac{\partial F(t)}{\partial F_i(t)}, \quad 1 \leq i \leq m. \quad (1)$$

Here  $F(t)$  is the function of network system failure,  $F_i(t)$  is the failure function of network system component  $i$ , and the relation of failure function and reliability function is  $F(t) = 1 - R(t)$ .

**Theorem 1.** For a network system with  $m$  components, the mathematical expression of Birnbaum measure is

$$I_i^B(t) = \frac{1 - \partial R(t)}{1 - \partial R_i(t)} = \frac{\partial R(t)}{\partial R_i(t)}, \quad 1 \leq i \leq m. \quad (2)$$

Here,  $I_i^B(t)$  is the Birnbaum measure of component  $i$ ,  $R(t)$  is the function of network system reliability, and  $R_i(t)$  is the reliability function of component  $i$ . If  $I_i^B(t)$  is larger, the impact on network system reliability  $R(t)$  is greater when the reliability of component  $i$  is changed.

Equation (2) shows the changes of network system reliability in the case of component  $i$  from normal state to failure state. Thus, Birnbaum measure can be defined as

$$I_i^B(t) = \frac{\partial R(t)}{\partial R_i(t)} = Pr(R_{e_i=1}(t)) - Pr(R_{e_i=0}(t)), \quad (3)$$

$$1 \leq i \leq m.$$

Here  $R_{e_i=1}(t)$  is the network system reliability when component  $i$  is normal and  $R_{e_i=0}(t)$  is the network system reliability when component  $i$  is failure.

**2.2. Critical Importance.** In network system, the failure rate of each component is different, so Lambert [9] proposes a critical importance to describe the probability of network system failure caused by network system component  $i$ . It can also be functioned as

$$I_i^{CMF}(t) = \frac{\partial F(t)}{\partial F_i(t)} \times \frac{\partial F_i(t)}{\partial F(t)}, \quad 1 \leq i \leq m. \quad (4)$$

From the perspective of the whole system, (4) can be further converted to the following one:

$$I_i^{CMF}(t) = I_i^B(t) \times \frac{\partial F_i(t)}{\partial F(t)}, \quad 1 \leq i \leq m. \quad (5)$$

Equation (5) shows that the reliability of network system is the product of the Birnbaum measure of component  $i$  and the ratio of system failure rate, when the state of component  $i$  is from the normal to failure.

**2.3. Network System Reliability Stability.** Assuming that a network system  $G$  contains  $m$  components, the risk growth factor [26] of component  $i$  can be defined as

$$\begin{aligned} RGI_i(t) &= F_{e_i=0}(t) - F(t) \\ &= (1 - R_{e_i=0}(t)) - (1 - R(t)) \\ &= R(t) - R_{e_i=0}(t). \end{aligned} \quad (6)$$

Here  $F_{e_i=0}(t)$  is the network system failure rate when component  $i$  is in malfunction. This formula describes the impact of the failure of component  $i$  on system reliability. In addition, based on (6), the other two reliability metrics, average risk and reliability stability, can be concluded, which measure the impact of single component malfunction on network reliability. The average risk growth factor can be expressed as follows by its own definition:

$$ARGI(t) = \frac{\sum_{i=1}^m RGI_i(t)}{n} = \frac{\sum_{i=1}^m R(t) - R_{e_i=0}(t)}{n}, \quad (7)$$

where is the average impact of all components failure individually on the network system reliability. On the basis of (7), the network reliability stability [26] can be formulated as

$$RS(t) = \frac{R(t) - ARG I(t)}{R(t)}. \quad (8)$$

It can be known, by the definition of network system reliability stability, that the network system reliability stability and network system reliability are greatly related to the network system average risk growth factor. When  $RS(t) \rightarrow 1$ , the network system component failure has little impact on the network system reliability, and vice versa.

**2.4. Experimental Analysis.** For any network system, it is noted that the complexity of network topology can make network components decomposed as combination of series and parallel system, and the complexity needed by optimal solution grows exponentially with the network size [27]. Next we will verify the validity of the previous measurements for different network structures using typical data recommended by [5, 14].

**(1) Series System.** Assuming that a system has  $m$  components connected in series configuration, the system will operate as long as all components are working. For Figure 1(a), the failure rates of components 1, 2, 3 in the network are  $f_1 \doteq 0.001$ ,  $f_2 \doteq 0.003$ ,  $f_3 \doteq 0.004$ . When  $t = 50$ , their reliabilities are, respectively,  $R_1 = 0.95$ ,  $R_2 = 0.86$ , and  $R_3 = 0.82$ . By [27], the reliability of network system is  $R(t) = R_1 \times R_2 \times R_3 = 0.6703$ , so the Birnbaum measurements of three components are as follows.

$$\begin{aligned} I_1^B(t) &= \frac{\partial R}{\partial R_1} = R_2 \times R_3 = 0.7046 \\ I_2^B(t) &= \frac{\partial R}{\partial R_2} = R_1 \times R_3 = 0.7788 \\ I_3^B(t) &= \frac{\partial R}{\partial R_3} = R_1 \times R_2 = 0.8187 \end{aligned} \quad (9)$$

Obviously,  $I_1^B(t) < I_2^B(t) < I_3^B(t)$ . Component 1 has more impact on the system. Increasing or decreasing the failure of component 3 will be the biggest change to the reliability of the system, so component 3 is the most important component of the system. In addition, the critical importance of components can be computed based on the Birnbaum measurements and (5).

$$\begin{aligned} I_1^{CMF}(t) &= I_1^B(t) \times \frac{1 - \partial R_1(t)}{1 - \partial R(t)} = I_1^B(t) \times \frac{1 - R_1}{1 - R} \\ &= 0.1407 \\ I_2^{CMF}(t) &= I_2^B(t) \times \frac{1 - \partial R_2(t)}{1 - \partial R(t)} = I_2^B(t) \times \frac{1 - R_2}{1 - R} \\ &= 0.3636 \\ I_3^{CMF}(t) &= I_3^B(t) \times \frac{1 - \partial R_3(t)}{1 - \partial R(t)} = I_3^B(t) \times \frac{1 - R_3}{1 - R} \\ &= 0.4501 \end{aligned} \quad (10)$$

By  $I_1^{CMF}(t) < I_2^{CMF}(t) < I_3^{CMF}(t)$  and the definition of critical importance, the probability of component 3 leads to the malfunction when the system is failure.



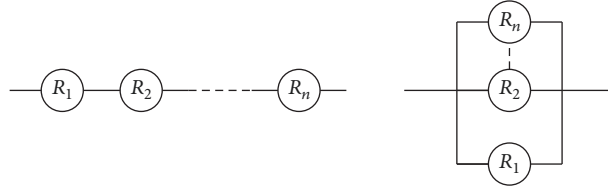


FIGURE 1: Series system and parallel system.

Based on (6), the risk growth fact of each component in system can be computed.

$$\begin{aligned} RGI_1(t) &= R(t) - R_{e_1=0}(t) = R(t) = 0.6703 \\ RGI_2(t) &= R(t) - R_{e_2=0}(t) = R(t) = 0.6703 \quad (11) \\ RGI_3(t) &= R(t) - R_{e_3=0}(t) = R(t) = 0.6703 \end{aligned}$$

The reliability stability depends on the risk growth fact of component, so the reliability stability can be calculated.

$$\begin{aligned} RS(t) &= \frac{R(t) - (RGI_1(t) + RGI_2(t) + RGI_3(t))/3}{R(t)} \quad (12) \\ &= 0 \end{aligned}$$

The result of risk growth fact is not the same as Birnbaum measurement. For each component,  $RGI_1(t) = RGI_2(t) = RGI_3(t)$ . According to physical significance, the failure of any network component in the series network system will lead to the failure of the network system, so the normal operation of series network requires no failure of all components. However, the reliability stability of network system is 0; on the basis of the definition of reliability stability, when  $RS(t) \rightarrow 0$ , the failure of component has an increasing influence on the reliability of network system. therefore, if there is a single component failure, the network must be failure.

(2) *Parallel System.* For Figure 1(b), the failure rates of components 1, 2, 3 in the network are  $f_1 = 0.001$ ,  $f_2 = 0.003$ ,  $f_3 = 0.004$ . When  $t = 50$ , their reliabilities are, respectively,  $R_1 = 0.95$ ,  $R_2 = 0.86$ , and  $R_3 = 0.82$ . Based on [14], the reliability of network system is  $R(t) = 1 - (1 - R_1) \times (1 - R_2) \times (1 - R_3) = 0.9892$ , so the Birnbaum measurements of three components are as follows.

$$\begin{aligned} I_1^B(t) &= \frac{\partial R}{\partial R_1} = (1 - R_2) \times (1 - R_3) = 0.0252 \\ I_2^B(t) &= \frac{\partial R}{\partial R_2} = (1 - R_1) \times (1 - R_3) = 0.0088 \quad (13) \\ I_3^B(t) &= \frac{\partial R}{\partial R_3} = (1 - R_1) \times (1 - R_2) = 0.0067 \end{aligned}$$

Obviously,  $I_1^B(t) > I_2^B(t) > I_3^B(t)$ . Component 1 has more impact on the system. Increasing or decreasing the failure of component 1 will be the biggest change to the reliability of the system, so component 1 is the most important

component in the system. In addition, the critical importance of components can be computed based on the Birnbaum measurements and (5).

$$\begin{aligned} I_1^{CMF}(t) &= I_1^B(t) \times \frac{1 - \partial R_1(t)}{1 - \partial R(t)} \\ &= I_1^B(t) \times \frac{1 - R_1}{(1 - R_1) \times (1 - R_2) \times (1 - R_3)} \\ &= 0.9907 \\ I_2^{CMF}(t) &= I_2^B(t) \times \frac{1 - \partial R_2(t)}{1 - \partial R(t)} \\ &= I_2^B(t) \times \frac{1 - R_2}{(1 - R_1) \times (1 - R_2) \times (1 - R_3)} \quad (14) \\ &= 0.9923 \\ I_3^{CMF}(t) &= I_3^B(t) \times \frac{1 - \partial R_3(t)}{1 - \partial R(t)} \\ &= I_3^B(t) \times \frac{1 - R_3}{(1 - R_1) \times (1 - R_2) \times (1 - R_3)} \\ &= 0.9945 \end{aligned}$$

Based on  $I_1^{CMF}(t) > I_2^{CMF}(t) > I_3^{CMF}(t)$  and the definition of critical importance, if network system is failure, the probability of component 3 leading to the failure of system is the biggest.

According to (6), the risk growth facts of components are as follows.

$$\begin{aligned} RGI_1(t) &= R(t) - R_{e_1=0}(t) \\ &= R_1 \times (1 - R_2) \times (1 - R_3) = 0.0240 \\ RGI_2(t) &= R(t) - R_{e_2=0}(t) \quad (15) \\ &= R_2 \times (1 - R_1) \times (1 - R_3) = 0.0076 \\ RGI_3(t) &= R(t) - R_{e_3=0}(t) \\ &= R_3 \times (1 - R_1) \times (1 - R_2) = 0.0055 \end{aligned}$$

However, the reliability stability depends on the risk growth fact of each component, so the reliability stability of the system is as follows.

$$RS(t) = \frac{R(t) - (RGI_1(t) + RGI_2(t) + RGI_3(t)) / 3}{R(t)} \quad (16)$$

$$= 0.9875$$

The result of risk growth fact is not the same as Birnbaum measurement; for each component,  $RGI_1(t) = RGI_2(t) = RGI_3(t)$ . According to physical significance, the failure of any network component in the parallel network system has little influence on the reliability of network system. However, the reliability stability of network system is 0.9875; on the basis of the definition of reliability stability, when  $RS(t) \rightarrow 1$ , the failure of component has less and less influence on network system reliability. Therefore, in the parallel network, the probability of the failure of single component leading to system's failure is very small, so the parallel network is more stable than the serial network.

### 3. Network Topology Decomposing Model Based on Cellular Automata

For large-scale network system, decomposing algorithm makes it simpler subsystems from scratch. In this paper, we use DB-CA algorithm based on cellular automaton (CA) to decompose network in [2]. In network  $G$ , let  $IV_i$  and  $OV_i$  be neighborhood of node  $i$ , with each node  $i$  mapping to a cell whose neighborhood is represented by two sets of nodes connected to it by its input arcs and output arcs, respectively.  $IV_i = \{e_j \mid e_j = \langle v_k, v_i \rangle, j \in (1, m), v_k \in V, e_j \in E\}$ ,  $OV_i = \{e_j \mid e_j = \langle v_i, v_k \rangle, j \in (1, m), v_k \in V, e_j \in E\}$ . The state of each node at time  $t_s$  is represented by a set  $IS_i(t_s)$ . The elements appearing in set  $IS_i(t_s)$  can be expressed as  $e_{i1}^* e_{i2}^* \cdots e_{ik}^*$ ,  $e_{ij}^* \in \{1, e_{ij}, \bar{e}_{ij}\}$ ,  $j = 1, 2, \dots, k$ . Let  $e_{ij}^* 1 = 1e_{ij}^* = e_{ij}^*$ .

**Theorem 2** (see [2]). *For any  $p_i, p_j \in SP$ , there exists item  $e_k$ , such that both  $P_i$  and  $P_j$  contain the same items, and the items in  $P_i$  and  $P_j$  are complementary; then  $P_i$  and  $P_j$  are disjoint. Otherwise  $P_i$  will intersect with  $P_j$ .*

$$I_i^B(t) = \begin{cases} \Pr(f_{e_i=1}(G) = 1) - \Pr(f_{e_i=0}(G) = 1) & j = i, \\ \Pr(e_j = 1) \times \Pr(f_{e_i=1}(G) = 1) - \Pr(e_j = 0) \times \Pr(f_{e_i=1}(G) = 1) & j \neq i. \end{cases} \quad (18)$$

**Lemma 5.** *The network topology decomposing method based on Cellular Automata in Section 3 decomposes network  $G$ , a disjoint set denoted as  $DB(G)$  can be calculated, and the Birnbaum measure of any component  $i$  in the network can be obtained by the following recursion formula:*

$$I_i^B(t) = Rel(DB(G), i)$$

$$= \begin{cases} +Rel(DB(G)|_{e_i=1}, i) & j = i \\ -1 \times Rel(DB(G)|_{e_i=0}, i) & j = i \\ +Pr(e_j = 1) \times Rel(DB(G)|_{e_i=1}, i) & j \neq i \\ -1 \times Pr(e_j = 0) \times Rel(DB(G)|_{e_i=0}, i) & j \neq i \end{cases} \quad (19)$$

Collecting all output paths generated by Algorithm 1, path set  $SP$  of the network is obtained, it is clear from Algorithm 1 and Theorem 2 that  $SP$  is a disjoint path set. Let  $p$  be the number of items in  $SP$ ,  $P$  be the items constituting the  $SP$ , and  $c$  be the number of arcs in  $P_i$ . By using Theorem 2 together with (5) in [2], the network fuzzy reliability expression at time  $t$  takes the following form:

$$\tilde{R}(t) = 1 - \prod_{i=1}^p \left( 1 - \prod_{j=1}^c R_{ij} \right) \quad (17)$$

$$\text{where } \begin{cases} R_{ij} = e^{-\bar{h}_k t}, & \text{if } (e_{ij}^* = e_k); \\ R_{ij} = 1 - e^{-h_k t}, & \text{if } (e_{ij}^* = \bar{e}_k). \end{cases}$$

## 4. Component Importance Estimation

**4.1. Component Importance Evaluation Model Based on DB-CA.** To simplify the importance evaluation method discussed in Section 3, an importance measure model based on two hypotheses is proposed; the network node reliability and the link failures are independent of each other.

**Theorem 3.** *For a network  $G$  with  $m$  components, the importance evaluation index of its any component  $i$  is as follows:*

$$\text{importance: } I_i^B(t) = \partial R(t) / \partial R_i t = \Pr(f_{e_i=1}(G) = 1) - \Pr(f_{e_i=1}(G) = 0);$$

$$\text{critical importance: } I_i^{CMF}(t) = I_i^B(t) \times F_i t / F_i;$$

$$\text{risk growth factor: } RGI_i(t) = F_{e_i=0}(t) - F(t) = R(t) - R_{e_i=0}(t).$$

Here  $f_{e_i=1}(G)$  denotes the reliability function of  $G$  when network component  $i$  is normal and  $f_{e_i=0}(G)$  represents the reliability function of  $G$  when network component  $i$  is failure.

**Lemma 4.** *In a network  $G$ , the Birnbaum measure of any component  $i$  in the network can be obtained by the following recursion formula:*

where  $DB(G)|_{e_i=1}$  is the remaining links in a path when component  $i$  works normally,  $DB(G)|_{e_i=0}$  is the remaining links in a path when component  $i$  is failure. If  $DB(G)|_{e_i=1}$  does not contain component  $i$ , then  $Rel(DB(G)|_{e_i=1}, i) = 1$ ; if  $DB(G)|_{e_i=0}$  contains component  $i$ , then  $Rel(DB(G)|_{e_i=0}, i) = 0$ .

Consequently, the network system component importance algorithm (NEA) can be designed as in Algorithm 2.

**4.2. Network System Reliability Stability Evaluation Based on DB-CA.** The stability describes the impact of component failure on the reliability of network system: the greater the stability value is, the less the impact of component failure on the reliability of network system is; otherwise the impact will be greater.

**Input:** a network  $G(V, E)$  with source node  $s$  and terminal node  $t$  are represented by node  $v_1$  and  $v_n$  respectively;  
**Output:** decomposed path set  $SP$ ;

- (1) Let  $IS_1(t_s) = \{1\}$ ,  $t_s = 0$ ;
- (2) **for**  $v_i \in V$  and  $v_i \neq v_1$  **do**
- (3)  $IS_i(t_s) = \emptyset$ ;
- (4) **end for**
- (5) **for**  $v_i \in V$  and  $v_i \neq v_n$  **do**
- (6)  $IS_i(t_s + 1) = \emptyset$ ,  $\forall m \in IS_i(t_s)$ , call Enumeration Procedure (m)[2];
- (7) **end for**
- (8) **if**  $\exists v_i \in V$ ,  $v_i \neq v_n$  and  $IS_i(t_s + 1) \neq \emptyset$  **then**
- (9)  $t_s = t_s + 1$ , go to (5);
- (10) **end if**
- (11)  $SP = \bigcup_{i=0}^{t_s} (IS_n(i))$ ;
- (12) **return**  $SP$ ;

ALGORITHM 1: DB-CA.

**Input:** Network  $G$  with source node  $s$  and terminal node  $t$ ;  
**Output:**  $Rel(DB(G), i)$  and  $Rel(DB(G))$ ;

- (1) Use DB-CA to decompose network  $G$ ;
- (2) Let  $Rel(DB(G), i) = 0$ ;
- (3) Assume  $DB_i(G) = \emptyset$  and  $Rel(DB(G)) = 0$ ;
- (4) Let  $Rel(p_j) = 1$  for any  $p_j$  in  $DB(G)$ ;
- (5) **for** any path  $k$  in  $P_j$ , **do**
- (6)  $Rel(p_j) = Rel(p_j) \times Pr(e_l)$ , where  $l \in [1, k]$ ;
- (7) **if**  $i == k$ , **then**
- (8) put path  $p_j$  into  $DB_i(G)$ ;
- (9) **end if**
- (10) **end for**
- (11) Let  $Rel(DB(G)) = Rel(DB(G)) + Rel(p_j)$ ;
- (12) **if** there are any path not been preprocessed **then**
- (13) go to (5);
- (14) **end if**
- (15) **return**  $DB(G)$  and  $Rel(DB(G))$ ;
- (16) **for** any path  $p_j$  in  $DB_i(G)$ ,  $Rel(p_j) = 1$  **do**
- (17) **while**  $k \neq i$  **do**
- (18)  $Rel(p_j) = Pr(e_k = State) \times Rel(p_j | e_k = State)$ , where  $State \in \{0, 1\}$ ;
- (19) **end while**
- (20) **while**  $k == i$  **do**
- (21)  $Rel(p_j) = Pr(e_k) \times Rel(p_j | e_k)$ ;
- (22) **if**  $e_k = 0$  **then**
- (23)  $Rel(p_j) = -1 \times Rel(p_j)$ ;
- (24) **else**  $\{e_k = 0\}$
- (25)  $Rel(p_j) = 1 \times Rel(p_j)$ ;
- (26) **end if**
- (27) **end while**
- (28) **end for**
- (29) Let  $Rel(DB(G), i) = Rel(DB(G), i) + Rel(p_j)$
- (30) **if** there are any path in  $DB(G)$  not given processed **then**
- (31) go to (16)
- (32) **end if**
- (33) **return**  $Rel(DB(G), i)$  and  $Rel(DB(G))$ .

ALGORITHM 2: NEA.

**Input:** Network  $G$  with source node  $s$  and terminal node  $t$ ;  
**Output:**  $Rel(DB(G), i)$  and  $Rel(DB(G), i)$ ;  
(1) Use DB-CA to decompose network  $G$ ;  
(2) Let  $Rel(DB(G)|_{e_j=0}) = 0$ ;  
(3) **for** any not been processed path  $P_j$  in  $DB(G)$ , **do**  
(4)  $Rel(p_j) = 1$ ;  
(5) **for** any not been processed link  $e_k$  in  $p_j$ , **do**  
(6)  $Rel(p_j) = Rel(p_j) \times Pr(e_l)$ , where  $l \in [1, k], l \neq i$   
(7) **if**  $l = i$ , **then**  
(8)  $Rel(p_j|_{e_l=1}) = 0$  **or**  $Rel(p_j|_{e_l=0}) = 1 \times Rel(p_j)$   
(9) **end if**  
(10) **end for**  
(11) **end for**  
(12) **if** there are any link in  $p_j$  not been processed **then**  
(13) go to (5);  
(14) **end if**  
(15) Let  $Rel(DB(G)|_{e_j=0}) = Rel(DB(G)|_{e_j=0}) + Rel(p_j)$ ;  
(16) **if** there are any path not been processed **then**  
(17) go to (3);  
(18) **end if**  
(19) **return**  $Rel(DB(G)|_{e_j=0})$ ;  
(20) **if** component  $c$  is present in the network  $G$ , its  $Rel(p_j|_{e_c=0})$  is not been computed, **then**  
(21) go to (2);  
(22) **end if**  
(23) **return**  $Rel(DB(G))$  and  $Rel(DB(G)|_{e_j=0}), j = 1, 2, \dots, m$ .

ALGORITHM 3: NSA.

**Lemma 6.** For the network  $G$  with  $m$  components, the reliability stability can be expressed as follows.

$$RS(t) = \frac{(1/m) \times \sum_{j=1}^m Rel(DB(G)|_{e_j=0})}{Rel(DB(G))} \quad (20)$$

*Proof.* By (8), the following result is

$$RS(t) = \frac{R(t) - ARG I(t)}{R(t)} = 1 - \frac{ARG I(t)}{Rel(DB(G))} \quad (21)$$

and based on (7)

$$\begin{aligned} ARG I(t) &= \frac{\sum_{i=1}^m R(t) - R_{e_i=0}(t)}{n} \\ &= \frac{n \times Rel(DB(G)) - \sum_{j=1}^m Rel(DB(G)|_{e_j=0})}{m} \end{aligned} \quad (22)$$

so

$$RS(t) = \frac{(1/m) \times \sum_{j=1}^m Rel(DB(G)|_{e_j=0})}{Rel(DB(G))} \quad (23)$$

and this is the end of the proof.  $\square$

**Lemma 7.** With Theorem 1, the risk growth factor is  $RGI_i(t) = R(t) - R_{e_i=0}(t)$  for any component  $i$ . Thus the risk growth factor of component  $i$  and the network reliability are related to the network reliability when component  $i$  is failure; the network

reliability and the network reliability when component  $i$  fails must be calculated. The computing process of network system is as follows.

$$RS(t) = Rel(DB(G))$$

$$= \begin{cases} +Pr(e_i = 1) \times Rel(DB(G)|_{e_i=1}) & 1 \leq j \leq n \\ -1 \times Pr(e_i = 0) \times Rel(DB(G)|_{e_i=0}) & 1 \leq j \leq n \end{cases} \quad (24)$$

Based on Lemma 5, the recurrence formula of network system reliability under component  $i$  failure can be proposed.

$$R_{e_i=0}(t) = Rel(DB(G)|_{e_i=0}, i)$$

$$= \begin{cases} +0 \times Rel(DB(G)|_{e_i=1}, i) & j = i \\ -1 \times Rel(DB(G)|_{e_i=0}, i) & j = i \\ +Pr(e_i = 1) \times Rel(DB(G)|_{e_i=1}, i) & j \neq i \\ -1 \times Pr(e_i = 0) \times Rel(DB(G)|_{e_i=0}, i) & j \neq i \end{cases} \quad (25)$$

On the basis of this, the evaluation algorithm of reliability stability evaluation algorithm is as shown in Algorithm 3.

## 5. Numerical Examples

In this section, NEA and NSA will be applied to the sample networks, shown in Figure 2; thus the Birnbaum measure of component and network reliability stability of each sample

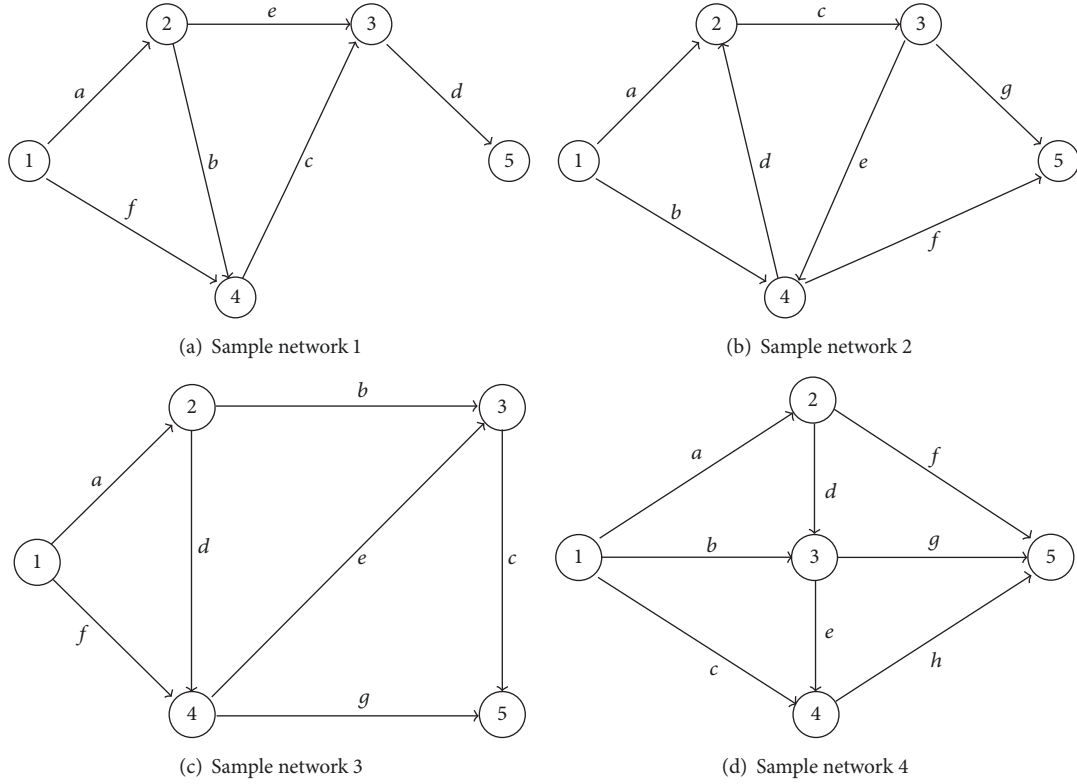


FIGURE 2: Network topology.

network can be calculated. By comparing the importance of the components in some network systems with different network topologies in Figure 1, the correctness of algorithms can be verified.

For all the sample networks in Figure 2, assuming that the failure rate of each component recommended by [5, 14] is 0.003 and  $t = 50$ , by DB-CA, the disjoint paths set of each sample network is obtained as follows:

$$SP_1 = \{aed, \bar{a}fcd, \bar{a}\bar{e}bcd, \bar{a}\bar{e}fcd\};$$

$$SP_2 = \{bf, \bar{b}acg, b\bar{f}dcg, \bar{b}ac\bar{g}ef, b\bar{f}dacg\};$$

$$SP_3 = \{\bar{a}fg, abc, \bar{a}\bar{b}dg, \bar{a}f\bar{g}ec, \bar{a}\bar{b}d\bar{g}ec, ab\bar{c}eg, \bar{a}\bar{b}d\bar{f}g, \bar{a}\bar{b}d\bar{f}\bar{g}ec, \bar{a}b\bar{c}edg, \bar{a}b\bar{c}d\bar{e}fg\};$$

$$SP_4 = \{af, \bar{a}ch\bar{f}, \bar{a}b\bar{c}f\bar{b}g, \bar{a}h\bar{c}e\bar{f}b\bar{g}, \bar{a}h\bar{b}c\bar{f}d\bar{e}, \bar{a}h\bar{b}c\bar{e}f\bar{d}g, \bar{a}h\bar{f}b\bar{g}, \bar{a}h\bar{b}f\bar{d}g, \bar{a}h\bar{c}, \bar{a}c\bar{b}e\bar{h}, \bar{a}c\bar{e}b\bar{g}h, \bar{a}h\bar{b}g\}.$$

As a result, the corresponding components Birnbaum measure and risk growth factor RGI of all sample networks in Figure 1 are shown in Table 1. And according to (1), the results are as follows:

- (i) the network reliability of sample network 1 is 0.7496, and the network reliability stability is 0.7038;
- (ii) the network reliability of sample network 2 is 0.9273, and the network reliability stability is 0.8683;
- (iii) the network reliability of sample network 3 is 0.9550, and the network reliability stability is 0.9132;

- (iv) the network reliability of sample network 4 is 0.9892, and the network reliability stability is 0.9759.

**5.1. Test Analysis I.** Through the study of all sample networks in Figure 2, with corresponding components Birnbaum measure and risk growth factor  $RGI$ , the importance of link  $d$  in Figure 2(a) is 0.9388, which is higher than those of link  $f$ ,  $g$  in Figure 2(b). Combining with the topology of Figure 2(a) and Figure 2(b), we can conclude that the fewer the links to a node, the more the importance to the reliability of the node. Correspondingly, the same results can be concluded from the Birnbaum measures of link  $c$ ,  $g$  in Figure 2(c) and link  $f$ ,  $g$ ,  $h$  in Figure 2(d).

Analyzing the importance of network system components, the next step is reliability calculation. Therefore, the computing time can be used as an important factor to evaluate the efficiency of algorithm. Accordingly, we compare the computational efficiency of network system component importance analysis method based on Cellular Automata (NSA) proposed in this paper and importance analysis method based on Monte Carlo (MCC) [27]. Since the MCC will spend more time on the sample collection process to calculate the reliability and component importance, the time of MCC is almost twice as much as that of NSA for the sample networks 1, 2, 3, 4. Thus, the efficiency of NSA proposed in this paper is better. From the time cost of each sample network in Table 2, combined with its own topology, we can conclude that, for the same number of nodes, the network topology is more complex and the rise of the computing

TABLE 1: Birnbaum measure and risk growth factor RGI.

link	Sample Network 1		Sample Network 2		Sample Network 3		Sample Network 4	
	Birnbaum Measure	RGI	Birnbaum Measure	RGI	Birnbaum Measure	RGI	Birnbaum Measure	RGI
<i>a</i>	0.3115	0.2749	0.1108	0.1101	0.1457	0.1254	0.0511	0.0440
<i>b</i>	0.0131	0.0098	0.2758	0.2133	0.0301	0.0259	0.0370	0.0318
<i>c</i>	0.1125	0.0959	0.1916	0.1865	0.1333	0.1148	0.0226	0.0195
<i>d</i>	0.9388	0.7496	0.0062	0.0106	0.0178	0.0153	0.0039	0.0034
<i>e</i>	0.1492	0.1384	0.0292	0.0106	0.0301	0.0259	0.0039	0.0034
<i>f</i>	0.0907	0.0632	0.4210	0.2133	0.1520	0.1308	0.0226	0.0195
<i>g</i>	–	–	0.1308	0.1101	0.1644	0.1415	0.0370	0.0318
<i>h</i>	–	–	–	–	–	–	0.0511	0.0440

TABLE 2: Comparative experiment.

Sample Network	NSA			MCC		
	reliability	stability	Computing time/s	reliability	stability	Computing time/s
1	0.7496	0.7038	0.853	0.7550	0.6996	1.429
2	0.9273	0.8683	1.235	0.9342	0.8635	2.013
3	0.9550	0.9132	1.573	0.9597	0.9106	2.987
4	0.9892	0.9759	1.823	0.9206	0.9699	3.581

time cost is almost doubled, which is related to the time cost increasing with the complexity of network when the network topology is decomposed by DB-CA.

**5.2. Test Analysis II.** In all sample networks in Figure 2, we assume that the failure rate of each component in networks recommended by [5, 14] is 0.003 and the interval time  $t=50$ ; the reliability of each sample network and the corresponding network reliability stability are shown in Figure 3. Firstly, we compare sample network 1 and sample network 2 with 5 nodes, and in sample network 1, there is a single link to a node; based on practical engineering experience, compared with sample network 2 the impact on the network reliability is more serious when the link is failure in sample network 1, which conforms to the result shown in Figure 3. Next, the sample network 2 and sample network 4 with 5 nodes are compared; the reliability and reliability stability of sample network 4 are higher than those of sample network 2. Because, with the same nodes, when some link fails in a network, the one with more links has more choices to achieve network connectivity, the survivability of the network is also improved. Finally, the sample network 2 and sample network 3 with 5 nodes and 7 links are compared; the reliability and reliability stability of sample network 3 are higher than those of sample network 2, because of the difference of network topology. In the case where origin node 1 and destination node 5 are connected, sample network 2 has longer path, which increases the failure probability of path and leads the reliability and reliability stability of sample network 2 to be lower than those of sample network 3. Through the above comparison results, it will be found that network survivability

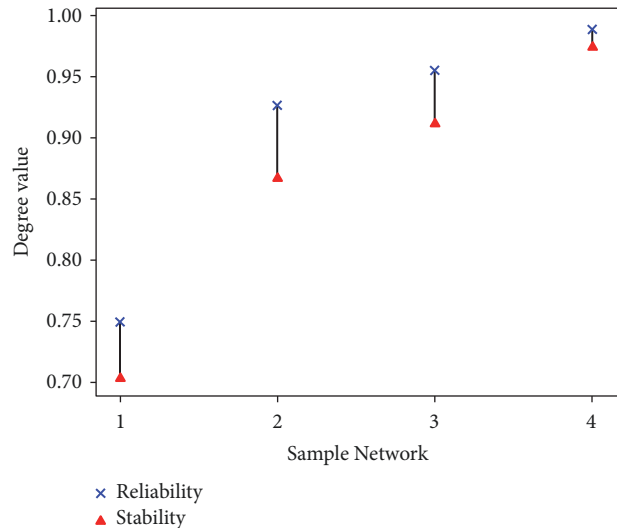


FIGURE 3: The reliability stability of sample networks.

can be described by the network reliability stability in the case of the random failure of network system components.

In the sample network in Figure 2(b), according to the measure method of link importance proposed in the literature [25], the Hash map is as shown in Figure 4. The comparison of the Birnbaum measure of sample network 2 in Table 1 and the Hash map can be concluded: the most important component of sample network 2 by 2 methods is link *b*; when the network link importance is calculated by

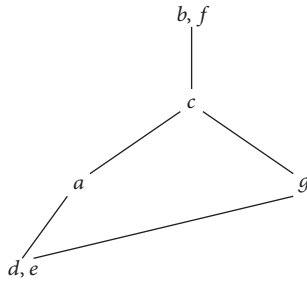


FIGURE 4: The topology of a network system.

NSA, the partial order relationship of Birnbaum measures and the Hash map of each link are consistent, which proves the correctness of the methods proposed in the paper.

However, it is found that the importance on some links cannot be ordered, by analysis in Hash map, for example, link  $b$  and  $f$ , link  $a$  and  $g$ , and link  $d$  and  $e$  in sample network 2. However, based on proposed methods in this paper, it can be concluded that the importance of link  $f$  is higher than that of link  $b$ , the importance of link  $g$  is higher than that of link  $a$ , and the importance of link  $e$  is higher than that of link  $d$ .

## 6. Conclusion

Evaluating the importance of components for complex networks is of great significance to the research of survivability and robusticity of networks. A component importance algorithm (NSA) and a reliability stability measure algorithm (NEA) based on DB-CA are proposed in this paper. NSA can well identify the importance of component with lower computational complexity, NEA can evaluate the influence of component failure on the reliability of network system, and the impact value can mostly affect the reliability of network system. The experimental results show that the proposed methods can effectively and precisely evaluate the impact of an arc failure in a network system. Future work is to implement our methods in various types of real network systems, for example, cloud computing system, social network system, and traffic network system. In particular, the  $k$ -out-of- $n$  system will be the next topic, in addition to correspondingly perfect and optimized measure methods according to system usage pattern.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 61602073 and 61672004) and partly

funded by Chongqing Basic and Frontier Research Project (under Grant No. cstc2017jcyjA0818).

## References

- [1] M. O. Ball, *Network Models*, Elsevier, 1995.
- [2] L. He and X. Zhang, "Fuzzy reliability analysis using cellular automata for network systems," *Information Sciences*, vol. 348, pp. 322–336, 2016.
- [3] S. Neumayer and E. Modiano, "Network reliability under geographically correlated line and disk failure models," *Computer Networks*, vol. 94, pp. 14–28, 2016.
- [4] D. R. Shier, *Network Reliability and Algebraic Structures*, Clarendon Press, 1991.
- [5] N. Nezamoddin and S. S. Lam, "Reliability and topology based network design using pattern mining guided genetic algorithm," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7483–7492, 2015.
- [6] C. J. Colbourn, *The Combinatorics of Network Reliability*, Oxford University Press, New York, NY, USA, 1987.
- [7] Z. Liu, C. Jiang, J. Wang, and H. Yu, "The node importance in actual complex networks based on a multi-attribute ranking method," *Knowledge-Based Systems*, vol. 84, pp. 56–66, 2015.
- [8] F. S. P. Tsen, T. Y. Sung, M. Y. Lin et al., "Finding the most vital edges with respect to the number of spanning trees," *IEEE Transactions on Reliability*, vol. 43, no. 4, pp. 600–603, 1994.
- [9] A. Hoyland and M. Rausand, *System Reliability Theory: Models and Statistical Methods*, John Wiley and Sons, 2009.
- [10] R. N. Allan, "Reliability Evaluation of Power Systems," Springer Science & Business Media, 2013.
- [11] A. Behr, L. Camarinopoulos, and G. Pampoukis, "Domination of  $k$ -out-of- $n$  Systems," *IEEE Transactions on Reliability*, vol. 44, no. 4, pp. 705–708, 1995.
- [12] L. H. Zhou, X. Wu, Y. C. Wang et al., "An algorithm and analysis of availability of fiber optic systems," *Acta Electronica Sinica*, vol. 29, no. 12, pp. 1713–1716, 2001.
- [13] S. Chanas and P. Zieliski, "The computational complexity of the criticality problems in a network with interval activity times," *European Journal of Operational Research*, vol. 136, no. 3, pp. 541–550, 2002.
- [14] J. Zhu and M. Collette, "A dynamic discretization method for reliability inference in dynamic bayesian networks," *Reliability Engineering & System Safety*, vol. 138, pp. 242–252, 2015.
- [15] H. Cancela and M. El Khadiri, "The recursive variance-reduction simulation algorithm for network reliability evaluation," *IEEE Transactions on Reliability*, vol. 52, no. 2, pp. 207–212, 2003.
- [16] B. Elshqeir, S. Soh, S. Rai, and M. Lazarescu, "Topology design with minimal cost subject to network reliability constraint," *IEEE Transactions on Reliability*, vol. 64, no. 1, pp. 118–131, 2015.
- [17] A. Jalali, A. G. Hawkes, L. R. Cui, and F. K. Hwang, "The optimal consecutive  $k$ -out-of- $n$ : G line for  $n \leq 2k$ ," *Journal of Plan Inference*, vol. 128, pp. 281–287, 2005.
- [18] W. Kuo, W. Zhang, and M. Zuo, "A consecutive  $k$ -out-of- $n$ :G system: the mirror image of a consecutive  $k$ -out-of- $n$ : F system," *IEEE Transactions on Reliability*, vol. 39, no. 2, pp. 244–253, 1990.
- [19] Y. Liu, "Uncertain random variables: a mixture of uncertainty and randomness," *Software Computer*, vol. 17, no. 4, pp. 625–634, 2013.

- [20] R. Gao and K. Yao, "Importance index of components in uncertain random systems," *Knowledge-Based Systems*, vol. 109, pp. 208–217, 2016.
- [21] S. Wu, Y. Chen, Q. Wu, and Z. Wang, "Linking component importance to optimisation of preventive maintenance policy," *Reliability Engineering & System Safety*, vol. 146, pp. 26–32, 2016.
- [22] Y.-P. Fang, N. Pedroni, and E. Zio, "Resilience-based component importance measures for critical infrastructure network systems," *IEEE Transactions on Reliability*, vol. 65, no. 2, pp. 502–512, 2016.
- [23] J. Ivanchev, H. Aydt, and A. Knoll, "Information maximizing optimal sensor placement robust against variations of traffic demand based on importance of nodes," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 3, pp. 714–725, 2016.
- [24] X. Zhu, Y. Fu, T. Yuan, and X. Wu, "Birnbbaum importance based heuristics for multi-type component assignment problems," *Reliability Engineering & System Safety*, vol. 165, pp. 209–221, 2017.
- [25] K. Laurio, F. Linker, and A. Narayanan, "Regular biosequence pattern matching with cellular automata," *Information Sciences*, vol. 146, no. 1, pp. 89–101, 2002.
- [26] L. Tian and L. Jie, "Seismic reliability analysis of large scale electric power network," *Journal of Zhengzhou University of Technology*, vol. 21, no. 4, pp. 12–15, 2000.
- [27] Z. Liang and P. Huazhi, "The monte-carlo-method-based analysis of the importance of the road network units in different sections of a harbor," *Traffic Engineering and Technology for National Defence*, vol. 13, no. 6, pp. 33–36, 2015 (Chinese).



## Research Article

# Metrics for Assessing Overall Performance of Inland Waterway Ports: A Bayesian Network Based Approach

Niamat Ullah Ibne Hossain,<sup>1</sup> Farjana Nur,<sup>1</sup> Raed Jaradat,<sup>1</sup> Seyedmohsen Hosseini ,<sup>2</sup> Mohammad Marufuzzaman,<sup>1</sup> Stephen M. Puryear,<sup>3</sup> and Randy K. Buchanan<sup>4</sup>

<sup>1</sup>Department of Industrial and Systems Engineering, Mississippi State University, P.O. Box 9542, Mississippi State, MS 39762, USA

<sup>2</sup>Industrial Engineering Technology, University of Southern Mississippi, Long Beach, MS 39560, USA

<sup>3</sup>Center for Advanced Vehicular Systems Extension (CAVSE), Mississippi State University, 153 Mississippi Parkway, Canton, MS 39046, USA

<sup>4</sup>Institute of Systems Engineering Research (ISER), U.S. Army Engineer Research Development Center (ERDC), 3909 Halls Ferry Rd, Vicksburg, MS 39180, USA

Correspondence should be addressed to Seyedmohsen Hosseini; [mohsen.hosseini@usm.edu](mailto:mohsen.hosseini@usm.edu)

Received 4 December 2018; Accepted 31 March 2019; Published 2 May 2019

Academic Editor: Lucia Valentina Gambuzza

Copyright © 2019 Niamat Ullah Ibne Hossain et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Because ports are considered to be the heart of the maritime transportation system, thereby assessing port performance is necessary for a nation's development and economic success. This study proposes a novel metric, namely, “*port performance index (PPI)*”, to determine the overall performance and utilization of inland waterway ports based on six criteria, *port facility*, *port availability*, *port economics*, *port service*, *port connectivity*, and *port environment*. Unlike existing literature, which mainly ranks ports based on quantitative factors, this study utilizes a Bayesian Network (BN) model that focuses on both quantitative and qualitative factors to rank a port. The assessment of inland waterway port performance is further analyzed based on different advanced techniques such as sensitivity analysis and belief propagation. Insights drawn from the study show that all the six criteria are necessary to predict PPI. The study also showed that port service has the highest impact while port economics has the lowest impact among the six criteria on PPI for inland waterway ports.

## 1. Introduction

With the aid of technology, multiple transportation modes such as rail, water, road, and air are used to transfer goods from one destination to another in a timely fashion. Certain important goods, such as heavy load items or bulk cargos (e.g., ore, grains, and coal), machinery, bulk liquids and oils, automobiles, containers, and perishable refrigerated items require safe shipping to the desired destination. Research showed that ground or air transport is not recommended for these types of goods and the preferred transportation option is by maritime [1]. Maritime transportation is more economic, safe, and environmentally friendly.

Ports are mainstay on maritime transportation system as they play a major role in the global and domestic freight transportation. Ports are generally categorized into two major

classes: *seaports* and *inland waterway ports*. Inland ports, known as coastal gateways for global trade, contribute to the rural, industrial, and agricultural development [2, 3]. Statistics show that 41 US states are being directly served by inland and intracoastal waterways for freight and passenger transportation. Inland waterway ports are located near a navigable river connected by a series of major canals and operated by lock and dam mechanism [4]. Unlike seaports, inland waterway ports do not have a deep draft; thus they cannot handle barges drafting more than 9 feet. Inland ports serve as a principal media for bulk transportation of the agriculture, mining, and manufacturing sectors with the connection of other intermodal facilities such as railroads and highways [3, 4]. A high number of current US ports are still underperforming due to the lack of proper management plans and decisive operational strategies [5]. To improve the

overall ports' performance, port authorities should advance their operational strategies by integrating cutting-edge technologies and agile planning. Port performance measurement is quite complex, due to the different port activities ranging from economic to technical to environmental. The overall ports' performance can be assessed through calculations of various performance activities [6, 7].

Due to the rapid advancement of global supply chain, inland waterway transport has become one of the important transportation modes (Weigmans et al., 2014). Thus, there is a need to employ a more "systemic" approach to better understand and manage any kind of undesirable consequences emanated from this complex system [8–10]. A major issue germane to inland port is the selection of ports based on performance indicators where these indicators determine the ranking of ports. Over the last decade, many port-related researches including sea and inland waterway are conducted on performance management and site selection. For example, Wiegman et al. (2014) conducted a detailed statistical analysis on the performance of the Dutch inland ports. They measured the performance of inland ports through transshipment level and growth in transshipment influenced by economic factors. Results indicated that the presence of a robust container terminal is necessary for a better port performance. Shetty and Dwarakish [7] identified a correlation between different port performance parameters such as loading/unloading rate, container dwell time, and terminal storage with overall productivity. The productivity is measured based on the number of vessels handled by the port. Along the same line, Kutin et al. [11] analyzed the relative efficiencies of fifty ASEAN ports and rank the ports efficiencies based on inland or sea type and supportive yard equipment. Alamoush [12] used a quantitative approach to study the impact of hinterland transport, specifically land transport (trucks) on the operational performance of the Jordanian inland port system. The findings from this study indicate that efficient hinterland transport system improves the operational performance of the inland waterway port. Oliveira and Cariou [13] developed a truncated regression model to explore the influence of interport competition on port efficacy and to investigate how the interrelationship between interport competition and efficacy can be varied if the assessment is performed at different geographic level. They suggested that interport competition has a reverse relationship with port efficacy and this negative relationship becomes more widespread when the competition occurs at a regional level compared to global levels. Bichou and Gray [14] proposed a conceptual framework of port performance through the lens of logistic and supply chain perspectives.

The current body of the literature is replete with other theoretical and empirical studies that focus on the subject of port performance and different types of port efficiencies. Interested readers can refer to the works of Coto-Millan et al. [15], Notteboom et al. [16], Barros [17], Díaz-Hernández et al. [18], Panayides et al. [19], Wanke [20], Chang and Tovar [21, 22], and Tovar and Wall [23]. Likewise, there are some analytical studies that have been devoted to other aspects pertaining to seaports, such as seaport characterization and classification (e.g., [24]), port operations and resilience (e.g.,

[25–27]), port selection (e.g., [28–31]), and port competitiveness (e.g., [32, 33]). In this research we propose a unique set of determinants (*i*) port facility, (*ii*) port availability, (*iii*) port economics (*iv*) port service, (*v*) port connectivity, and (*vi*) port environment that impact the inland port performance. These determinants were derived based on Minimum Link Set (MLS) perspective. A MLS is a minimum set of operational factor or component required for the system to actively perform (Johansen and Tien; 2017), which implies failure of any factor or component within a system triggers cascading impact and leads to failure of the MLS (Jianag et al., 2016). Table 1 provides a summary of the current themes related to the different aspects of port literature. These themes serve as a baseline in the development of the proposed model.

Although there are many theoretical and empirical studies focused on the analysis and characterization of seaports, there is scant research that has attempted to quantify the performance of inland port using unique set of determinants. To address this gap, the following are the contributions made by this research:

- (i) Propose a new metric "*port performance indicator (PPI)*" to assess the probability of an inland port performance.
- (ii) Propose a probabilistic graphical model, a Bayesian network (BN), to predict the probability of port performance based on six criteria.
- (iii) Conduct different types of analysis such as belief propagation and sensitivity analyses to provide better insights regarding the results of the proposed model.
- (iv) Use BN as an effective tool in solving transportation and logistics management problems.

To the best of our knowledge, this is the first attempt to assess the probability of inland port performance using a Bayesian approach (BN). This research also presents the efficacy of BN tool in the context of transportation and logistics management. BN has some advantages over other approaches. BN is a powerful analytical tool that can be used for decision-making under uncertainty. Another important feature of BN is the ability to model both qualitative and quantitative variables which is different from other approaches such as swing weight, Analytical Hierarchy Process (AHP), or Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS). BN can also be used to conduct probabilistic scenario analysis known as *belief propagation analysis*. BN accounts for all causal factors to produce a final model, to reduce the burden of parameter acquisition, and to overturn the previous assumption by taking new evidences into consideration such as *subjective belief* and *objective data* [34]. The Bayesian approach has been used in different domains and applications such as electrical infrastructure system [35], security management [36], customer service management [37], traffic accidents [38], manufacturing systems [39], natural resource management [40, 41], power system [42], and data classification [43], electric vehicle [44], and supply chain and logistics [45, 46].

An overview of BN is presented below, followed by the identification of the criteria and subcriteria that impact the

TABLE 1: Current themes of the port literature.

Authors	Measures for port	Approach
Wiegmans et al. (2015)	Inland port performance	Statistical (regression) analysis
Shetty and Dwarakish [7]	Inland port performance and productivity	Statistical analysis
Kutin et al. [11]	Relative efficiencies	Data Envelopment Analysis (DEA)
Alamouh [12]	Port operational performance	Conceptual framework with analytical model
Oliveira and Cariou [13]	Port efficiency	Data Envelopment Analysis (DEA)
Bichou and Gray [14]	Port performance	Conceptual logistic and supply chain approach
Coto-Millan et al. [15]	Port economic efficiency	Stochastic Cost Frontier (SCF)
Notteboom et al. [16]	Relative efficiency of container terminal	Stochastic Cost Frontier (SCF)
Baros [17]	Technical efficiency	Stochastic Cost Frontier (SCF)
Díaz-Hernández et al. [18]	Technical and allocative efficiency	Stochastic Cost Frontier (SCF)
Panayides et al. [19]	Economic efficiency	Data Envelopment Analysis (DEA)
Wanke [20]	Physical infrastructure efficiency, shipment consolidation efficiency	Data Envelopment Analysis (DEA)
Chang and Tovar [21, 22]	Technical efficiency	Stochastic Distance Function (SDF)
Tovar and Wall [23]	Port productive efficiency	Directional technology distance function approach
Bichou and Gray [24]	Terminology for classifying seaport	Conceptual framework
Hosseini and Barker [25, 26]	Resilience	Bayesian approach
Sierra et al. [27]	Harbour operability	Numerical Model
Ugboma et al. [28]	Port selection	Analytical Hierarchy Process (AHP)
Chang et al. [29]	Port selection factor	Exploratory factor and confirmatory factor analyses
Gohomene (2008)	Port selection	Analytical Hierarchy Process (AHP)
Nur et al. [31]	Port selection	Stochastic Analytical Hierarchy Process (SAHP)
Song and Yeo [32]	Competitiveness of container ports	Analytical Hierarchy Process (AHP)
Yeo et al. [33]	Competitiveness of container ports	Fuzzy methodology

overall inland port performance and translate these criteria into BN model to assess the probability of PPI. From this, belief propagation and sensitivity analyses are presented, respectively. The paper concludes with implications, recommendations, and future research.

## 2. Fundamentals of Bayesian Network

BN is a Directed Acyclic Graph (DAG) which consists of vertices (nodes) and edges (arcs) where vertices represent the variables and edges signify the relationship between the two variables in the existing network. BNs are structured based on Bayes' theorem, capable of making statistical inferences in a rational way by updating the prior beliefs of any event which means entering an evidence in an *child node* will lead to backward belief propagation and ultimately result in updating the probability distributions for the *parent node(s)* and vice versa [34]. BN requires fewer probability parameters compared to a full joint probability model. Equation (1) represents the generic rule of the Bayesian theorem [47].

$$P_R(A_1, A_2, A_3 \dots A_n) = \prod_{i=1}^n P_R(a_i | \mu_i) = \prod_{i=1}^n \phi_{A_i | \mu_i} \quad (1)$$

In (1), the underlying network is specified by a pair  $R = \{G, \Theta\}$  where  $G$  is the directed acyclic consisting of a set of random

variables  $A_1, A_2, A_3 \dots A_n$ , and  $\Theta$  representing the set of the probability functions. Each  $a_i$  in  $A_i$  is provisioned on  $\mu_i$  for the set of the parameters of  $A_i$  in  $G$ .

In the underlying structure of a BN, the initial probabilities (unconditional) or prior information of the root nodes can be obtained from a subjective judgment (e.g., expert knowledge /historical data) or through a frequentist approach (observed data). The conditional probabilities refer to the quantitative degree of belief to describe uncertainty among nodes. In some cases, it is challenging to define the conditional probability table (CPT) for a large set of data. Thus, we used *AgenaRisk* software to offset this challenge, having said that Bayesian equation is used to calculate CPT with known initial probabilities of each node as shown in (2) [48].

$$P(A_j | B) = \frac{P(B | A_j) \times P(A_j)}{P(B | A_i) \times P(A_i)} \quad (2)$$

where  $i = 1, 2, \dots, n; j = 1, 2$ .

To illustrate the operational principle of BN networks, let us consider a BN structure with a set of variables  $R = \{A_1, A_2, A_3, A_4, A_5, A_6\}$  and a set of edges to show the conditional interdependencies among the variables (see

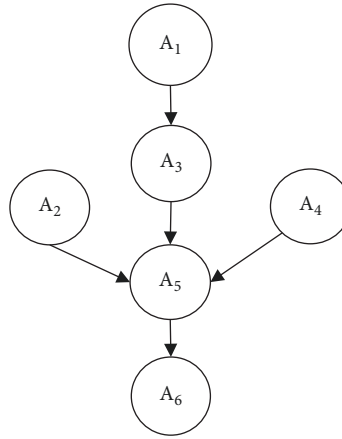


FIGURE 1: A sample of the Bayesian Network (BN) with six nodes.

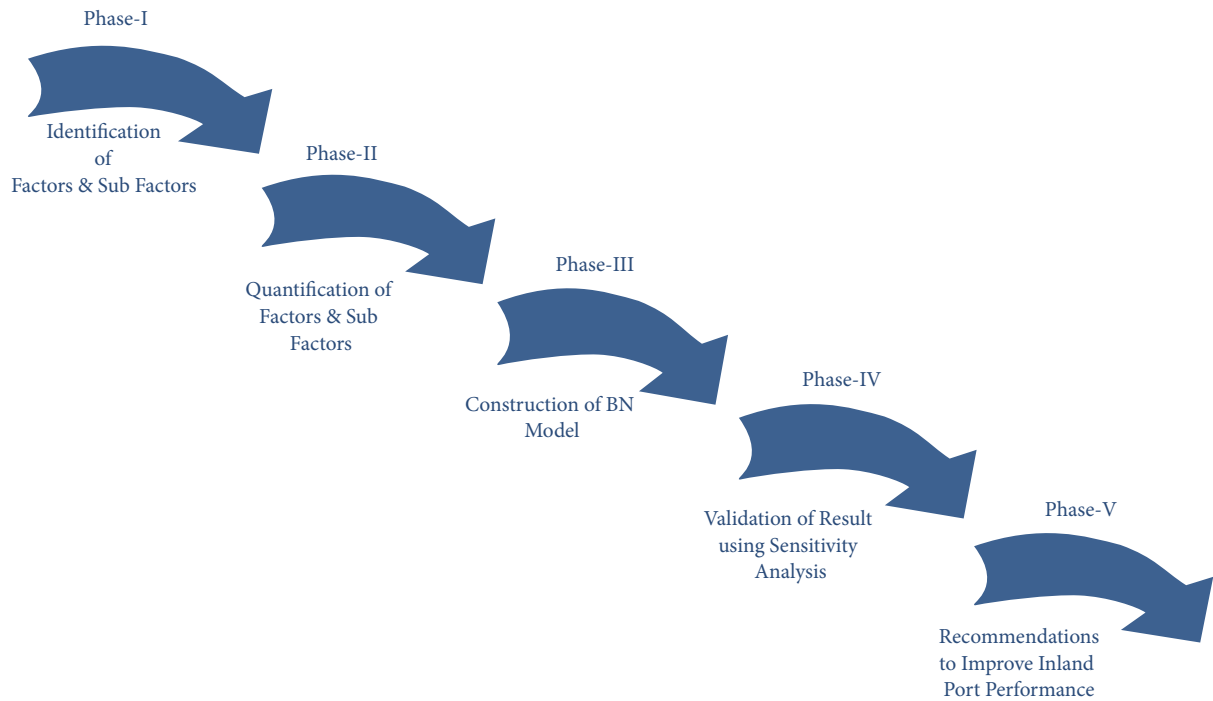


FIGURE 2: Proposed framework for inland port performance assessment.

Figure 1). The general expression of the full joint probability distribution can be represented as follows:

$$\begin{aligned}
 P(A_1, A_2, A_3, \dots, A_n) &= P(A_1 | A_2, A_3, \dots, A_n) \\
 &\cdot P(A_2 | A_3, \dots, A_n) \dots P(A_{n-1} | A_n) P(A_n) \\
 &= \prod_{i=1}^n P(A_i | \text{Parents}(A_i))
 \end{aligned} \quad (3)$$

The corresponding decomposition of the joint distribution of variables can be streamlined as follows:

$$\begin{aligned}
 P(A_1, A_2, A_3 \dots A_6) &= P(A_1) P(A_2) P(A_4) \\
 &\cdot P(A_3 | A_1) P(A_5 | A_2, A_3, A_4) P(A_6 | A_5) A_4
 \end{aligned} \quad (4)$$

### 3. Proposed Framework for Inland Port Performance Assessment

The proposed framework consists of five phases as illustrated in Figure 2.

- (i) *Phase I*. Identification of factors and subfactors: the first phase is to identify the factors and subfactors that could impact the performance of port infrastructure. First, the current research related to port performance is studied and analyzed, and initial subcriteria are constructed. Second, opinions from domain experts are incorporated into the scope of port performance management and the less important subcriteria are discarded, and finally all the subfactors are clustered

TABLE 2: Inland port performance index scale.

Intensity of PPI	Definition	Significance
PPI > 80%	High Standard	Maintains superior standard on all the performance criteria.
60% < PPI < 80%	Moderate Standard	Meets adequate performance standard, but the scope of improvement is available.
PPI < 60%	Low Standard	Performance level is lower than average and significant improvement is needed.

into six main criteria, namely, (i) *port facility*, (ii) *port availability*, (iii) *port economics*, (iv) *port service*, (v) *port connectivity*, and (vi) *port environment*.

- (ii) *Phase II*. Quantification and assessment of factors and subfactors: the second phase is to quantify the factors and subfactors. It also includes the determination of the likelihood of the related factors based on the subjective or frequentist approach.
- (iii) *Phase III*. Construction of BN model: a BN is used to quantify the probability of the port performance.
- (iv) *Phase IV*. Analysis of result: different techniques such as sensitivity analysis and belief propagation analysis were conducted to draw the insights from Phase III.
- (v) *Phase V*. Recommendation for port performance improvement: based on the analysis, different recommendations are provided to improve the overall port performance.

### 3.1. Performance Standard for Inland Waterway Port

3.1.1. *Proposed Inland Port Performance Index (PPI)*. Based on existing literature, a number of interrelated factors that influence the performance of inland waterway ports are identified. This research summarized all the possible factors and classified them into six criteria: (i) *port facility*, (ii) *port availability*, (iii) *port economics*, (iv) *port service*, (v) *port connectivity*, and (vi) *port environment*. The proposed Inland Port Performance Index (PPI) describes the probability of the performance standard that an inland waterway port can meet. For instance, the probability of PPI being 80% true means that there is 80% likelihood that the specific port will meet the performance standard based on the cited criteria.

The characterization of PPI incorporates the most significant parameters that impact the performance standard of the inland waterway port. In order to express PPI through a numerical scale, a value between 0 and 100 is assigned. It is important to note that the selected metric is based on expert knowledge within inland waterway port system and is used to highlight the overall performance of an inland waterway port. The subjective description of the metric values of PPI is explained in Table 2 and the base model of the BN for measuring PPI is illustrated in Figure 3.

3.1.2. *Port Facility (Criterion #1)*. An inland port is highly integrated with a maritime terminal to ensure smooth flow of logistical activities across the globe. Port terminal amenities and other key facilities such as warehouse area, outdoor ground storage, and dock-wall depth govern the overall facility of the inland port for freight handling and distribution.

(i) *Terminal facility*: it consists of three contributors port throughput, types of existing terminal, and number of diversified products.

(a) *Throughput*: it is volume of cargo or number of vessels that a port can handle over time. Throughput can be measured in terms of tons or transportation equivalent units (TEU). Different factors such as competition between ports, international and domestic cargo demand, and business arrangements can influence the terminal throughput [49].

(b) *Types of existing terminal*: from a transport facility viewpoint, a top-tier inland port possesses three kinds of terminals: *Satellite terminal* located near the port facility and used mainly for container trans-loading, *Freight distribution cluster or load center* dedicated to support warehousing and logistic functions, and *intermodal terminal* used to regulate freight circulation through intermodal facilities [50].

(c) *Number of diversified products*: based on United Nations Conference on Trade and Development (UNCTAD), five categories of seaborne trade that a port can handle are *containers*, *petroleum*, *crude oil*, *main bulk commodities*, and *other dry products*. However, for an inland port, diversity in the type of products handled is limited to two to three types.

(ii) *Key facilities*: warehouse area, outdoor ground storage, and dock-wall depth play a vital role in freight storage and port performance.

(a) *Warehouse area*: warehouse facility is tied up with freight storage and distribution operations. Sometimes warehouse areas are facilitated by staging areas to support loading and unloading operations.

(b) *Outdoor ground storage*: outdoor ground storage offers port expansion opportunity due to possible growth rate in port throughput. Some ports use outdoor ground storage as cargo staging/assembly zone, maintenance area, barge consolidation and deconsolidation facility, and container depot.

(c) *Dock-wall*: dock wall facilitates berthing area for vessel/cargo.

3.1.3. *Modelling of Port Facility*. In order to model port facility, three variables were used: (i) *Boolean variables* are



expressed in forms of a dichotomous response (true/false, yes/no) to present positive and negative outcomes respectively, (ii) *fixed variables* are modelled in constant values, and (iii) *continuous variables* are random variables with a known probability distribution.

A Boolean variable with two states of *true* and *false* are used to model facility, terminal facility, and key facilities nodes. The true state represents a positive outcome while the false state indicates a negative outcome. For instance, in Figure 3 the probability of facility being true or likelihood of meeting port facility is 87.41% while the probability of facility being false is 12.58%. Similar logic is also applicable to the other two Boolean nodes (terminal facility and key facilities).

Truncated normal distribution is used to model continuous variables such as throughput, warehouse facility, outdoor ground storage, and dock-wall area. Truncated normal distribution is a simple modification of a normal distribution that confines the mean values between lower and upper bounds. For example, the area of the warehouse facility cannot be negative and maximum warehouse area for the inland port does not generally exceed 1,50000 m<sup>2</sup>. Hence, the truncated normal distribution is found to be the most appropriate distribution to model the aforementioned continuous variables. The truncated distribution is defined in terms of four parameters:  $\mu$ , mean (i.e., central tendency);  $\sigma^2$ , variance (i.e., confidence in the results); lower bound and upper bound.

It is apparent from Figure 3 that port facility is conditioned upon terminal facility and key facilities. There might be other hidden factors contributing to port facility. This can be better described by the NoisyOR function. These hidden or missing parameters are known as “leak parameters” in NoisyOR function (see (5)).

$$\text{NoisyOR} (A_1, S_1, A_2, S_2, \dots, A_n, S_n, l) \quad (5)$$

Leak factor ( $l$ ) can be defined as the extent to which missing factors from the model can contribute to the consequence being true. It is the probability that  $B$  will be true when all of its causal factors are false. The conditional probability of  $B$  obtained with the NoisyOR function is presented below in (6).

$$\begin{aligned} P(B = \text{True} | A_1, A_2 \dots A_n) \\ = 1 \\ - \prod_{i=1}^n [(1 - P(B = \text{True} | A_i = \text{True})) (1 - P(l))] \end{aligned} \quad (6)$$

The modelling procedure for port facility and its contributors are summarized in Table 3.

In the proposed BN model, in order to calculate the posterior probability of the “port facility”, we used NoisyOR function, which is represented in (7). The equation means that, in order to meet port facility, both factor terminal facility and key facility are equally responsible (75%) and

other hidden factors are contributing rest of 25% to achieve desired port facility.

$$\begin{aligned} \text{Port Capacity} = \text{NoisyOR} (\text{Terminal Facility}, 0.75, \\ \text{Key Facility}, 0.75, 0.25) \end{aligned} \quad (7)$$

**3.1.4. Port Availability (Criterion #2).** Availability is the level to which the system (port) can self-organize itself to avoid any discontinuity of the system’s performance due to undesirable consequences. In terms of inland port, availability refers to the readiness of the adequate resources to perform the daily operation. Inland port availabilities can be measured through port resilience, the readiness of different kinds of equipment and labour support, dredging maintenance, and congestion rate.

- (i) *Port Equipment.* In order to perform daily operations such as handling cargo and stevedore operations, port authorities use different kinds of capital equipment, such as gang-tree/rubber-tree cranes, mooring instruments, forklifts, reach stackers, and towing vehicles. For the inland waterway port, gang-tree cranes and straddle carriers are most commonly used.
- (ii) *Port Resilience.* Ability of a port to bounce back to its normal operating condition after any type of disruption such as adverse weather conditions, human-made error, and/or cyberattack. Resilience capacities are the strategies to recover a region/entity from any shock or external perturbation due to disruption. Resilience capacities can expressed by means of *absorptive capacity, adaptive capacity, and restorative capacity* of the corresponding system [35, 39, 51–53]. It is generally designed based on metastructure under internal deterioration and external perturbation [54].

*Absorptive capacity* is an endogenous feature of the system and is also considered to be the first course of defense to minimize the impacts of the disruption [25, 26, 39, 52]. Maintenance, availability of additional capital equipment, and skillful response team are the mainstay of the port absorptive capacity. *Adaptive capacity*, which is considered to be the midline of defense, is described as the ability of a system to self-organize itself and provide immediate solutions to cope with the external shock without any recovery activity [25, 26, 35]. Alternate routing and relocating of resources are the key factors germane to the adaptive capacity within port infrastructure. *Restorative capacity* considered to be the last line of defense is the degree to which a system can efficiently repair or restore from the degraded state [35]. Within the restorative capacity of port infrastructure, two salient determinants are identified: restoration of resources and restoration of service.

- (iii) *Workforce.* Workforce is an asset to any port infrastructure. Operators and stevedores also ensure proper utilization of the available resources and reduce the delay during port operations such as loading and

TABLE 3: Modelling of variables contributed to port facility.

Variable Name	Modelling Technique	Modelling Description
Throughput	TNORM	Based on inland port statistics, a truncated normal distribution is used to approximate the annual throughput of an inland port with an average of 20 million ton/year.
Product Diversity	Trinagular	Product diversity is approximated with a triangular distribution with minimum, most likely, and maximum of 2, 3, and 5 respectively.
Types of Terminal	Arithmetic	Types of the terminal are fixed and equal to 3 to avail the proper terminal facility.
Warehouse Area	TNORM	Warehouse area is defined by truncated normal distribution with a mean of 7,5000 $m^2$ .
Outdoor Ground Storage	TNORM	The outdoor storage area is approximated using truncated normal distribution with an average of 40 acres.
Dock-wall	TNORM	A truncated normal distribution is used to approximate the dock wall parameter of an inland port with an average of 700 $m$ .
Terminal Facility	Comparative Expression	Threshold for throughput and product diversity are set as 20 million/year and 2 respectively while port should have exactly 3 types of terminal to avail the proper terminal facility.
Key Facilities	Comparative Expression	The key facilities will be suitable (true) for port operational activities if the parameters of dock-wall, warehouse area and outdoor ground Storage are higher than 600 $m$ , 50,000 $m^2$ , and 30 acres respectively.

TABLE 4: Modelling of equipment variable.

Variable Name	Modelling Technique	Modelling Description
Equipment	Comparative Expression	An IF logic is used for modelling an “equipment” node. The threshold for number of cranes and straddle is considered to be one. In order to perform the regular operation, the port equipment requirements will be met (true state) if the number of cranes and straddle is more than one and otherwise not (false state).

unloading, recouping, grading, fuel transfer from vessel to the pier, gate operations, and others.

- (iv) *Port Congestion*. A main parameter to measure port availability. Congestion occurs when vessels enter into a queue and wait for an extended period of time to access port facilities.
- (v) *Dredging Maintenance*. Dredging is important for inland ports to maintain the desired water depth at their approach channels, specifically for those ports where the waterway accumulates silt quickly. Enhanced dredging capability also positively impacts the availability of the port.

**3.1.5. Modelling of Port Availability.** Boolean variables were used to model the contributors of port availability. For instance, the prior distribution of the resilience variable with two states of True = 93.35% and False = 6.65%, which means that there is an 93.35% chance that a strong resilient port infrastructure would contribute to increase the availability of the port facility, while there is a 6.65% chance that it may fail. In other words, the resilience of the port system is successful 93.35% (True state) and fails 6.65% (False state) of the time. The same logic is applicable for other Boolean nodes under port availability variables. Table 4 provides detailed model description of the equipment variables.

Port resilience and availability criterion is designed using NoisyOR function and equation is presented below:

$$Port\ resilience = NoisyOR (absorptiev\ capacity, 0.70, adaptive\ capacity, 0.70, restorative\ caapcity, 0.80, 0.15) \quad (8)$$

$$Port\ Availabilily = NoisyOR (Dredging\ maintenance, 0.50, labor\ support, 0.50, port\ congestion, 0.20, port\ resilience, 0.50, equipment, 0.50, 0.15) \quad (9)$$

**3.1.6. Port Economics (Criterion #3).** *Port economics*: the solvency of the major stakeholders, the overall status of the global economy, and port pricing also influence the port economics.

Port associated cost consists of terminal-handling costs, port calling cost, and concession pricing.

- (i) *Terminal-handling cost (THC)*: it is related to the cost for loading or unloading, container service and clearance, storage, repacking, and forwarding. It includes all services essential for moving the freight



TABLE 5: Modelling of variables contributed to port pricing.

Variable Name	Modelling Technique	Description
Terminal Handling Cost (THC)	TNORM	Based on inland waterway port data, the average terminal cost is the \$6,500/barge with a variance of \$25 and cost varies from The \$5,000 to \$8,000 based on the size of the barge and other related factors.
Port Calling Cost (PCC)	TNORM	Port calling cost varies from \$1,800 to \$3,800 depending upon the size of the vessel with an average of \$2,500 dollar/ vessel.
Concession pricing (Upfront fee)	TNORM	Concession granting depends upon the area of the facility and follows a truncated normal distribution with an average of \$60 million upfront fee.

TABLE 6: Modelling of port pricing variable.

Variable Name	Modelling Technique	Description
Port pricing cost	Comparative Expression	IF the THC, PCC and concussion grant are lower than \$7000, \$2,800 and \$65 million, respectively then the port economy cost is within limit (true state), otherwise not (false state).

onwards through the port before being loaded onto a vessel. More precisely, beyond the sea freight, THC is the charge that is paid by shippers for handling the containers at the inland port.

- (ii) *Port calling cost*: it is the costs related to all types of services offered to handle a ship or vessel. More precisely, it is the summation of prices to be paid for various services including access to the terminal, pilotage, time costs, damage and delay, and bunkering.
- (iii) *Concession cost*: it is decided by the port governing body and it is the cost of acquiring a dedicated maritime facility such as a terminal, yard, or outdoor storage. It is mainly a leasehold agreement and used for a variety of reasons.

Tables 6 and 5 describe the modelling details of port pricing variables and its contributed factors, respectively.

NoisyOR function, which is discussed in the previous section, is applied to design the economic criterion as presented below.

$$\text{Port Economics} = \text{NoisyOR} (\text{Port pricing}, 0.50, \text{ solvency of major user}, 0.50, \text{ economic climate}, 0.50, 0.15) \quad (10)$$

**3.1.7. Port Service (Criterion #4).** An inland port's service level indicator is highly integrated with response rate, service availability, container dwell time, and vessel transit time at the port.

- (i) *Response rate*: a measure of port service including faster documentation, availability and quick updates of electronic information, early detection, and response to problems. The higher response rate reduces unnecessary cost pertaining to any port.
- (ii) *Service availability*: it is a measure of port performance that refers to port services at any time of the

day or the service is restricted for a fixed time. High service availability means the operational hours of a port is higher than normal and vice versa. Generally, for an inland waterway port the service hours vary from 8 to 24 hours per day.

- (iii) *Dwell time*: dwell time is measured by the amount of time a container waits to be picked up at a marine terminal after being offloaded from a ship or vessel [55]. This is considered as a key benchmark for port's service level indicator. Port authority always experiences a constant challenge to keep the dwell time down while accommodating inbound and outbound vessels.
- (iv) *Transit time*: transit time management is one of the main concerns of port authority. It is the amount of time that a vessel spends in different ports on the way to its destination port. This also includes waiting time dockside before loading/unloading.

**3.1.8. Modelling of Port Service.** As apparent from Figure 3, port service consists of four main contributors including response rate, service availability, dwell time, and transit time. Truncated normal distribution similar to what is explained in the previous section is applied to model the aforementioned four contributors. The modelling procedure of port service and its contributors are summarized in Tables 7 and 8. The modeling procedure of geographical location and port accessibility is summarized in Table 9.

**3.1.9. Port Connectivity (Criterion #5).** Connectivity refers to the level of ease that an inland port supports freight transportation through the supply chain network.

- (i) *Geographical Location*. Some geographic areas possess natural advantages for business flourishing. It is beneficial for an inland waterway port to have a logistic cluster, a major supplier, and an intermodal connection within its vicinity. Geographical locations

TABLE 7: NPTs of the variables describing response rate, service availability, dwell time, and transit time.

Variable Name	Modelling Technique	Modelling Description
Response rate	TNORM	Based on the inland port statistics, the response rate of inland port varies from 85% to 95% with an average of 90%.
Service availability	TNORM	At the worst possible scenario, the port operating hours are not lower than 16 hours and at the best possible the port provides 24 hours service a day.
Dwell time	TNORM	Dwell time is modelled with a truncated normal distribution with mean, LB, and UB of 3, 1, and 4 hours respectively.
Transit time	TNORM	A truncated normal distribution is used to approximate the transit time of a vessel with an average of 24 hr.

TABLE 8: Modelling of service variable.

Variable Name	Modelling Technique	Modelling Description
Port Service	Comparative Expression	If the values of response rate or service availability are greater than 90% or 12 hours, respectively AND dwell time or transit time is lower than 3 or 30 hours, respectively then the satisfactory service level is achieved (true state), otherwise not (false state)

TABLE 9: NPTs of the variables describing the contributor to geographical location and port accessibility.

Variable Name	Modelling Technique	Modelling Description
Distance to Logistic & Industrial Area	Arithmetic	Distance to logistic & industrial cluster is constant: 75 miles.
Distance to Major Supplier	Arithmetic	The distance between port and major supplier is constant: 50 miles.
Distance to Intermodal Connection	Triangular	The distance between the port and the intermodal connection is modelled with triangular distribution with mean, LB, and UB of 50, 35 and 65 miles respectively.
Geographical Location	Comparative Expression	Based on the historical data, if the logistics cluster, major supplier, and intermodal connection are within 75, 50, and 50 miles from the inland port location, then port has suitable geographical location for trade and commence (true state), otherwise not (false).
Port accessibilities	Boolean	We assume that 90% of the time, the port is accessible by all required modes of transportation. On very few occasions, port entire accessibility is halted by natural calamities, human error and/or cyber-attack.

associated with modal accessibility and availability of inland regional access influence shaping of the development of surrounding locality [56]. All these criteria boost productivity, save time, lower the logistic cost, and provide access to global markets.

- (a) *Proximity to the industrial area and logistics cluster*: logistics clusters provide integrated services in logistics. If the port location is close to the logistics cluster center, the port may perform better than port locations farther away from the cluster center. Port proximity to industrial areas has a great impact on port choice.
- (b) *Proximity to major supplier*: proximity to major suppliers will enhance national and international trade throughout the port. Traders can exploit economies of scale in shipping products and, in turn, will be benefitted from in time delivery and lower inventory holding costs.
- (c) *Proximity to intermodal connection*: the strength of inland intermodal transportation network includes the availability of railway, roadway, and rail spur in the port province. The ports that

are close to intermodal connection generally get better transportation facilities, such as highways, railroads, and airports.

- (ii) *Port Accessibility*. Port accessibility means the port location can be approachable by different modes of transportation. Port accessibility depends upon the location and the overall infrastructure of the port.

3.1.10. *Modelling of Port Connectivity*. NoisyOR function, as discussed in the previous section, is used here to calculate the conditional probability of connectivity criterion as defined in

$$\text{Port Connectivity} = \text{NoisyOR}(\text{port accessibility}, 0.75, \text{geographical location}, 0.75, 0.20) \quad (11)$$

The above equation means that port accessibility and geographical locations are equally responsible to obtain desired connectivity and there are other hidden factors directly or indirectly influencing to achieve preferred port connectivity.

3.1.11. *Port Environment (Criterion #6)*. Two main subcriteria, emission at port and probability of natural disaster,

TABLE 10: Modelling of variables related to emission at the port.

Variable Name	Modelling Technique	Description
CO <sub>2</sub> Level	TNORM	The level of CO <sub>2</sub> emission (million tonnes/year) follows a truncated normal distribution (TNORM) with an average of 15 and variance of 2.5. The level of CO <sub>2</sub> emission (million tonnes/year) never goes below 12 (lower bound) and 18.5 (upper bound), according to the historical data for the inland waterway port.
NO <sub>x</sub> Level	TNORM	The level of NO <sub>x</sub> (million tons/year) follows a truncated normal distribution with an average of 0.2, the variance of 0.005 with lower bound and upper bound of .1 million tons/year and 0.4 million tons/year, respectively.

TABLE 11: Modelling of emission at port variable.

Variable Name	Modelling Technique	Modelling Description
Emission at Port	Comparative Expression	An IF logic is used for modelling “emission at port” node. The threshold for CO <sub>2</sub> and NO <sub>x</sub> level are considered to be 18 and 0.2 respectively. The emission index will meet (true) if the levels of CO <sub>2</sub> and NO <sub>x</sub> at the port are less than 18 million tons/year and 0.2 million tons/year respectively.

TABLE 12: NPT for port environment.

<i>Emission</i>	True		True	
<i>Disruption</i>	False	True	False	True
False	0.2	0.15	0.15	0.05
True	0.8	0.85	0.85	0.95

are found as the main determinants to the environmental criterion for port performance.

- (i) *Emission at port*: shipping emission has a substantial impact on the overall environment of the port. Most shipping emissions in ports account for discharges of CO, SO<sub>x</sub>, and NO<sub>x</sub>. The quantity of total emissions depends on the type and size of vessel berth at the port. At the same time, emission due to regular port equipment also accounts for deterioration of the air quality of the port. In order to reduce these emissions, strong policy along with public awareness is required.
- (ii) *Probability of natural disaster*: the inland waterway port is often susceptible to different natural disasters such as hurricanes, cyclones, drought, or flood, combined with the prevailing port temperature and humidity.

**3.1.12. Modelling of Port Environment.** Figure 3 shows that port environment mainly conditioned upon two determinants: disruption of probability and emission at the port. The Boolean node is used to express the probability of disruption and emission at the port. For instance, disruption probability of 15% means that, according to the historical data, there is a 15% chance that the inland port might be impacted by adverse weather conditions. Tables 10 and 11 show the procedures of modelling for emission at the port and its contributed variables.

The NPT is the probability table that summarizes the occurrence probability between the causal relationship nodes. NPT can be developed manually or achieved by eliciting the distribution or related expression. For a node without

its parent node, the NPT would be simply the probability distribution of that specific node. NPT for port environment is shown in Table 12.

**3.1.13. Modelling of Port Performance Indicator.** The ultimate target node “port performance index” is conditioned on its contributed criteria (i) *port facility*, (ii) *port availability*, (iii) *port economics*, (iv) *port service*, (v) *port connectivity*, and (vi) *port environment*. The posterior probability of PPI is calculated as the weighted sum of its contributed criteria. Initially, it is assumed that the weight of each factor is equally distributed. The general equation associated with a weighted mean (WMEAN) is presented in (12), where (*i*) is the number of variables connected (six in this case) to the weighted average node of port performance index (see Figure 2) and  $W_i$  is the weight associated with the *i*th variable.

$$WMEAN = \sum W_i A_i, \quad i = 1, 2, \dots, n,$$

$$\forall i = 1 \quad 0 < W_i < 1, \quad \sum_i W_i = 1 \quad (12)$$

To compare the port performance index, based on above-mentioned criteria, the probability the probability of PPI being true is 87.82%, meaning that there is 87.82% likelihood (chance) that the specific port will meet the performance standard based on the cited criteria.

## 4. Validation of the Model

In order to validate the structure of the BN model, apart from traditional methods, *sensitivity analysis* (SA) is considered a powerful technique. It is a useful approach to examine the impact of the contributors on the target node within the same model, i.e., which node has more impact to its connected node. This is obtained by recalculating the outcomes of the targeted node under possible alternative assumptions. The object of the SA is to check that the outcomes generated from the propagation analysis are consistent with the expert’s expectation. To obtain more insights and better understanding of the simulation model, we used *AgenaRisk* software

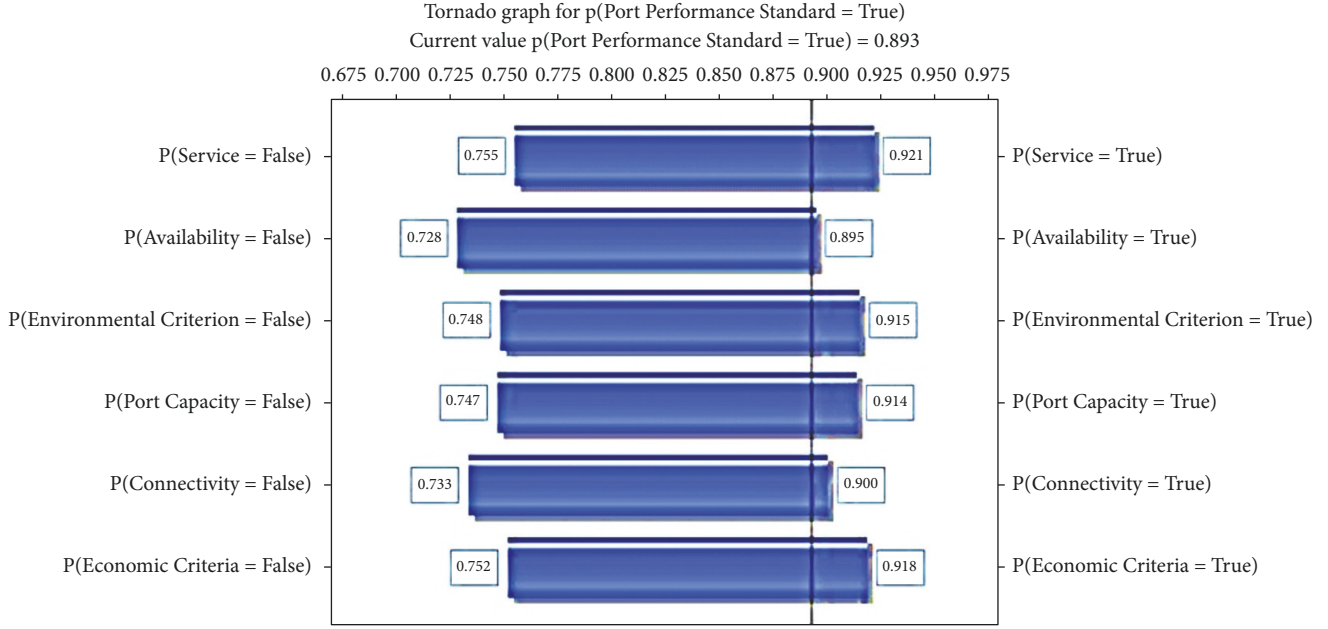


FIGURE 4: Sensitivity analysis of PPI.

to investigate the extent to which the six key performance contributors affect the port performance index. We performed SA on PPI as a target node with respect to its causal factors including (i) port facility, (ii) port availability, (iii) port economics, (iv) port service, (v) port connectivity, and (vi) port environment as subsequently shown in Figure 3. Tornado charts, generated during the SA analysis, identified the lowest and highest values of a posterior probability for each possible state of the target node if specific observations are inputted into the model. To be more specific, the length of the bars corresponding to each sensitive node in the tornado graph illustrates a measure of the impact of that corresponding node on PPI. Figure 4 shows the impact of those variables when the PPI is “true.” It is apparent from Figure 4 that the length of the bar chart for all the selected variables is almost same; however, port service has a slightly higher impact on PPI than other variables, whereas port economics has a lower impact on PPI among all the variables. To elaborate, from Figure 4 it is also apparent that the probability of PPI (“true”) for the first port given the result of port service goes from 0.755 (when port service is “false”) to 0.921 (when port service is “true”). In other words, the probability of PPI for first port is 0.921 when the port service is met. This range (0.755–0.921) is exactly the bar that is plotted in the tornado graph illustrated in Figure 4. This range varies from 0.752 to 0.918 for the port economics which implies port economics has the lowest impact on the PPI among all the variables. From Figure 4 it can also be interpreted that the probability of PPI for the port is more sensitive to the changes in the states of port service and least sensitive to changes in port economics. It can be concluded that although all the factors have almost same importance to the variability of PPI, port service ranked top in terms of contribution to the variability of PPI, and therefore the port

authorities and top management should emphasize more on port service than others determinants.

## 5. Propagation Analysis

The feature of the BN to disseminate the effect of evidence through the network is defined as “propagation analysis”. Special types of reasoning can be done through propagation analysis. During propagation analysis, different evidences (observations) can be entered anywhere in the underlying BN model to update the marginal probabilities of all unobserved variables. In this section, we have conducted forward propagation analysis to predict the probability distribution of PPI under the combination of the aforementioned six contributors. The related probability is represented in (13) (Zhou, 2018).

$$\begin{aligned}
 P(T = State_t) &= \sum_{n=1}^n (P(PPI = State_t | A_1 = a_j, A_2 \\
 &= a_j, \dots, A_n = a_j) \times P(A_1 = a_j, A_2 = a_j, \dots, A_n \\
 &= a_j))
 \end{aligned} \tag{13}$$

where  $n$  refers the number of parent nodes and  $a_j$  is the  $i$ th state of the parent node.  $(P(PPI = State_t | A_1 = a_j, A_2 = a_j, \dots, A_n = a_j))$  is conditioned probability distribution when  $T = State_t$ .

During the *forward propagation* analysis, we have designed two scenarios (1) pessimistic and (2) optimistic. Scenario 1 (pessimistic scenario) accounts for two assumptions: (i) the service hours of the port is set to 8 hours instead of truncated normal distribution with a mean of 16 hours and

TABLE 13: Summary of propagation analyses.

Scenario	Description of the Scenario	PPI	Standard of the Port	Significance of the propagation	Remarks
Base Case	Underlying BN Model	87.82%	Class B	-	-
Scenario 1 (Pessimistic Scenario)	Service hr = 8hr, Transit time=36hr (Other variables remain unchanged)	74.87%	Class C	Shows how probability of PPI changes with service level (service availability and transit time)	<i>Port Service</i> criterion has a significant impact on the probability of PPI.
Scenario 2 (Optimistic Scenario)	Throughput=30 million ton, Dredging Maintenance=100% True, Port environment =100 True% (all environmental criteria are met)	91.28%	Class A	Shows how probability of PPI varies with port facility (throughput), port availability (dredging maintenance) and port environment.	<i>Port facility, port availability</i> and <i>port environment</i> have less impact on the probability of PPI compared to port service

(ii) the transit hours of the port are set to 36 hours in lieu of truncated normal distribution with an average of 30 hours. Scenario 1 measures the changes in the probability of PPI of the first port if the service hours reduced to a constant value of 8 hours and transit time increased to a constant value of 36 hours. From Figure 5, the probability of PPI of the first port significantly reduced from 87.82% to 74.87% which indicates the importance of service hours and transit time on PPI. Scenario 2 (optimistic scenario) simulates the impact of throughput, dredging maintenance, and port environment on PPI for the inland port. We set the throughput to 30 million/year, dredging maintenance and environmental as 100% instead of their prior distribution parameters which increases the PPI from 87.82% to 91.28% (see Figure 6). This type of propagation analysis gives the capability to decision makers to make any number of observations especially on variables with inherent uncertainty and measures.

A summary of propagation analyses is given in Table 13. Scenario 1 and 2 are illustrated in Figures 5 and 6, respectively.

## 6. Conclusion

In this study, a novel dimensionless metric named port performance indicator (PPI) is introduced to assess the level of port performance based on six basic determinates named: *port facility, port availability, port economics, port service, port connectivity, and port environment*. In order to calculate the probability of PPI, we developed a Bayesian framework that captures the possible factors and subfactors pertaining to the level of port performance. The PPI indicates the level of performance that will be met by a specific port. It also provides a better understanding regarding the performance of a specific port under uncertainty. The PPI will aid port stakeholders in making better decisions in terms of the management of port supply chain and infrastructure. Such decisions include the number of port service hours, scaling port throughput, and others. In real-world practices, it is quite difficult to predict a port performance because of uncertainty and ambiguity (e.g., operational uncertainty, disruption uncertainty, etc.). In response, predicting the PPI through the Bayesian approach can help to substantially

reduce this uncertainty and will ensure better visibility for decision-making. Belief propagation feature of the Bayesian approach allows practitioners to run different future scenarios where assumptions and alterations in conditions or states can be tested and verified. Belief propagation analysis also demonstrates the weightage of interdependency among the different variables of the underlying BN structure. The BN structure is also validated through sensitivity analysis. The general interpretation of the sensitivity analysis indicates that all six criteria are important to predict PPI; however, port service has a slightly higher impact and port economics has a lower impact among all factors in predicting the probability of PPI. The novelty of this work is summarized.

- (i) The development of a model to assess port performance indicator (PPI).
- (ii) The underlying determinates pertaining to port performance were identified and classified with respect to six main factors named: *port facility, port availability, port service, port economics, port connectivity, and port environment*.
- (iii) The proposed model is then tested and validated through different types of analysis to draw better managerial insights to handle uncertainties. Results indicate that all the factors have almost same importance to the variability of PPI, port service ranked top in terms of contribution to the variability of PPI, and therefore the port management should stress more on service criterion than others factors.
- (iv) Demonstrate the efficacy of BN as an effective tool in solving transportation and logistics management problems.

This study can be extended in several research directions. In our study, NPT has been defined based on subjective judgment (expert opinion) and frequentist approach (historical data). Other methods such as swing weights, Delphi technique, and the classical method can be used to improve the accuracy of NPT. Furthermore, a deep investigation is required to identify the other related factors that might indirectly impact the PPI.

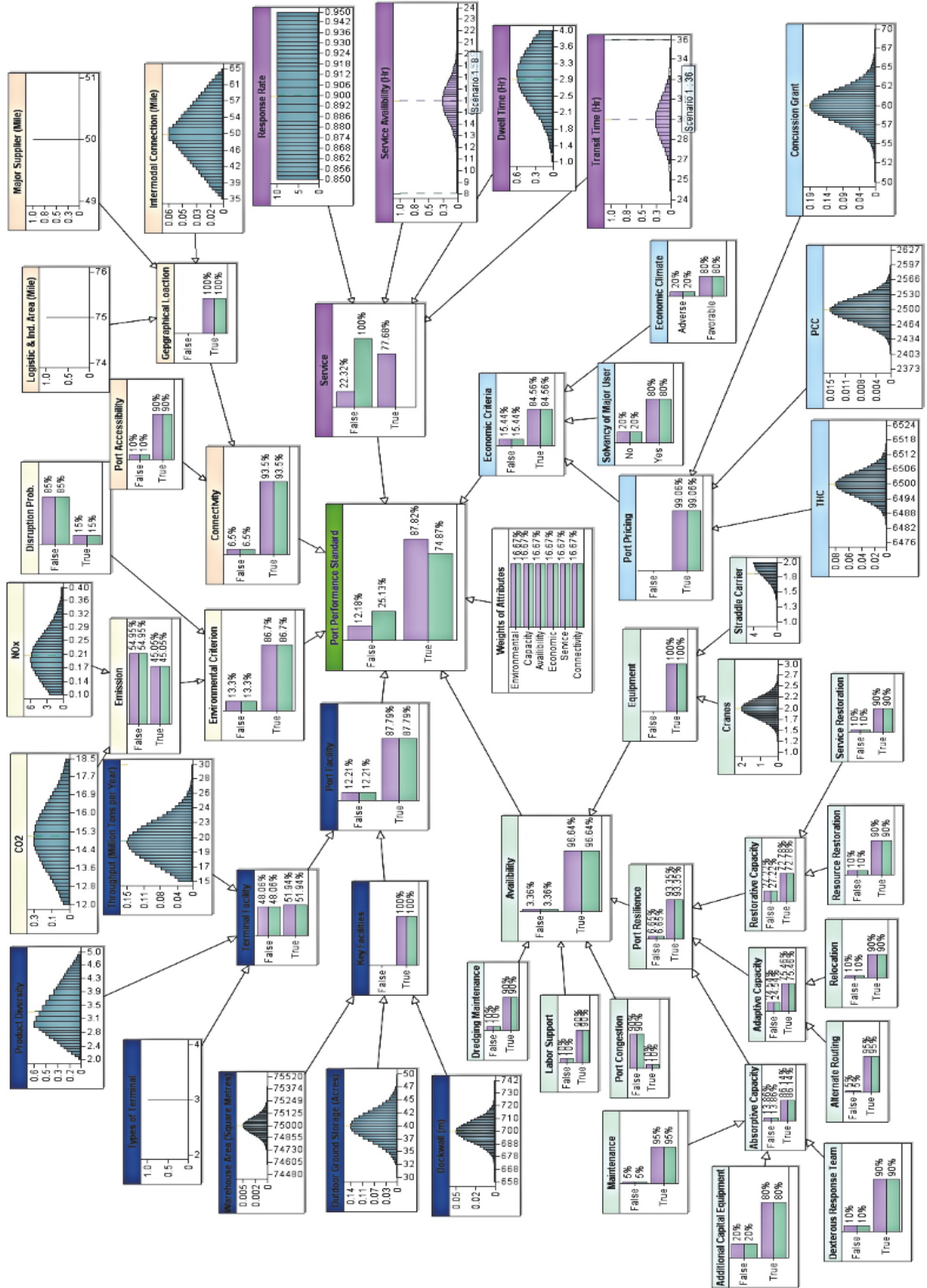


FIGURE 5: The developed BN model for scenario 1.

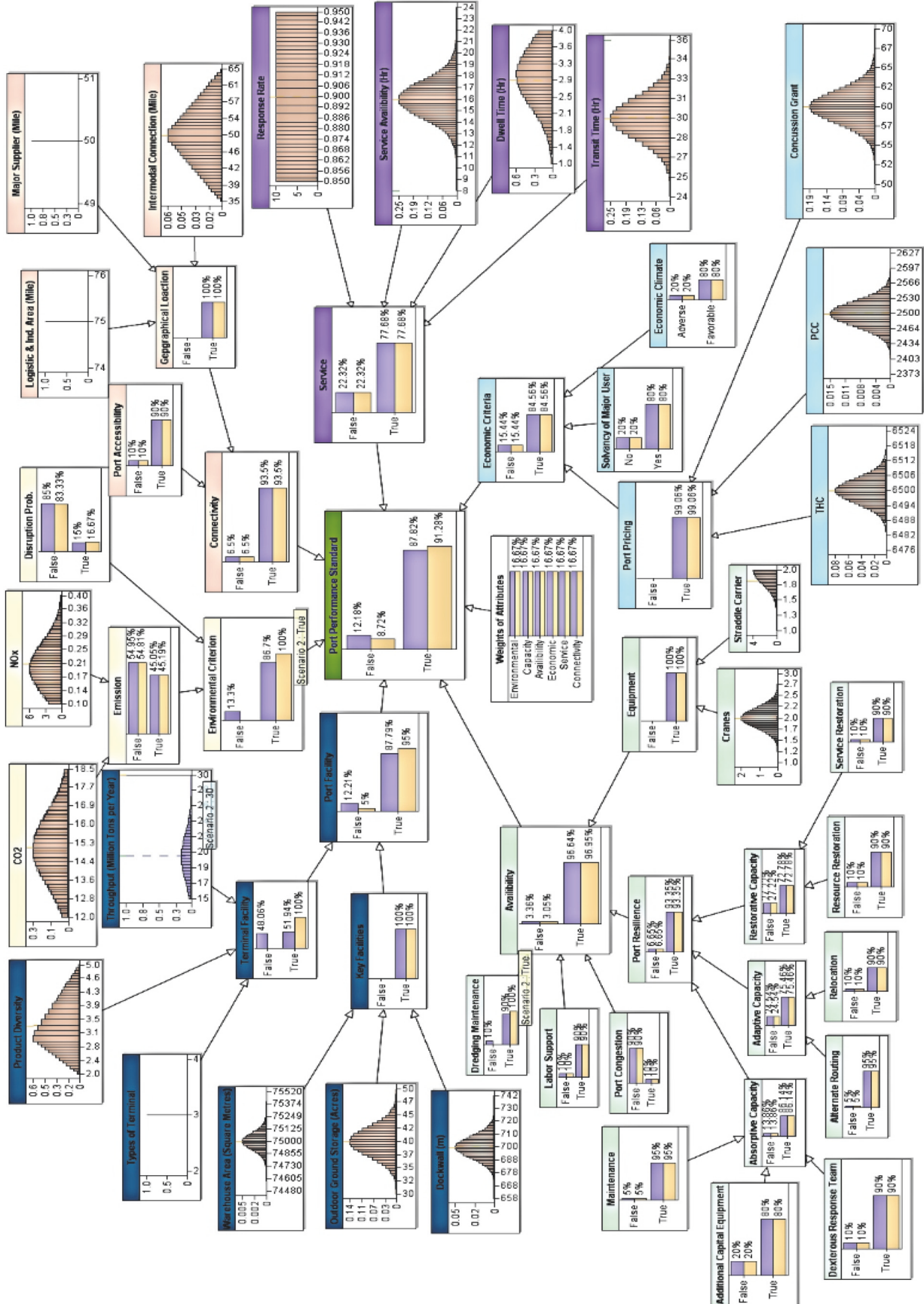


FIGURE 6: The developed BN model for scenario 2.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] M. Stopford, *Maritime Economics*, Routledge Press, 2013.
- [2] Transportation Research Board, *The Marine Transportation System and the Federal Role: Measuring Performance, Targeting Improvement*, vol. 279, Transportation Research Board, 2004.
- [3] C. R. Miller, "The evolving role of rural river ports as strategic economic development actors," *Water Resources and Rural Development*, vol. 9, pp. 28–38, 2017.
- [4] National Research Council, *Funding and Managing the US Inland Waterways System: What Policy Makers Need to Know*, Transportation Research Board Publishers, 2015.
- [5] X. H. Lim, P. Senter, J. Villars, and K. Watts-Fitzlerland, "Is the port authority underperforming? An operational efficiency analysis of port authority's divisions," 2014, [http://mpaenvironment.ei.columbia.edu/files/2014/06/Columbia\\_Port\\_Authority\\_Final\\_Report\\_05132014.pdf](http://mpaenvironment.ei.columbia.edu/files/2014/06/Columbia_Port_Authority_Final_Report_05132014.pdf).
- [6] P. de Langen, M. Nidjam, and M. van der Horst, "Indicators to measure port performance," *Journal of Maritime Research*, vol. 4, no. 1, pp. 23–36, 2007.
- [7] D. K. Shetty and G. S. Dwarakish, "Measuring port performance and productivity," *ISH Journal of Hydraulic Engineering*, pp. 1–7, 2018.
- [8] N. U. I. Hossain, F. Nur, and R. M. Jaradat, "An analytical study of hazards and risks in the shipbuilding industry," in *Proceedings of American Society for Engineering Management Annual Conference*, pp. 18–21, Charlotte, NC, USA, 2016.
- [9] N. U. I. Hossain and R. Jaradat, "A Synthesis of definitions for systems engineering," in *Proceedings of the International Annual Conference of the American Society for Engineering Management*, pp. 1–10, American Society for Engineering Management (ASEM), 2018.
- [10] A. Alfaqiri, N. U. Hossain, R. Jaradat et al., "A systemic approach for disruption risk assessment in oil and gas supply chains," *International Journal of Critical Infrastructures*, vol. 15, no. 3, 2019.
- [11] N. Kutin, T. T. Nguyen, and T. Vallée, "Relative efficiencies of ASEAN container ports based on data envelopment analysis," *Asian Journal of Shipping and Logistics*, vol. 33, no. 2, pp. 67–77, 2017.
- [12] A. S. M. Alamoush, *The impact of hinterland transport on port operational performance: a Jordanian case [Masters Dissertation]*, World Maritime University, Malmö, Sweden, 2016.
- [13] G. Figueiredo De Oliveira and P. Cariou, "The impact of competition on container port (in)efficiency," *Transportation Research Part A: Policy and Practice*, vol. 78, pp. 124–133, 2015.
- [14] K. Bichou and R. Gray, "A logistics and supply chain management approach to port performance measurement," *Maritime Policy & Management*, vol. 31, no. 1, pp. 47–67, 2004.
- [15] P. Coto-Millan, J. Banos-Pino, and A. Rodriguez-Alvarez, "Economic efficiency in Spanish ports: some empirical evidence," *Maritime Policy Management*, vol. 27, no. 2, pp. 169–174, 2000.
- [16] T. Notteboom, C. Coeck, and J. Van Den Broeck, "Measuring and explaining the relative efficiency of container terminals by means of Bayesian stochastic frontier models," *International Journal of Maritime Economics*, vol. 2, no. 2, pp. 83–106, 2000.
- [17] C. P. Barros, "Decomposing growth in Portuguese seaports: a frontier cost approach," *Maritime Economics & Logistics*, vol. 7, no. 4, pp. 297–315, 2005.
- [18] J. J. Díaz-Hernández, E. Martínez-Budría, and S. Jara-Díaz, "Parametric estimation of inefficiency in cargo handling in Spanish ports," *Journal of Productivity Analysis*, vol. 30, no. 3, pp. 223–232, 2008.
- [19] P. M. Panayides, C. N. Maxoulis, T.-F. Wang, and K. Y. A. Ng, "A critical analysis of DEA applications to seaport economic efficiency measurement," *Transport Reviews*, vol. 29, no. 2, pp. 183–206, 2009.
- [20] P. F. Wanke, "Physical infrastructure and shipment consolidation efficiency drivers in Brazilian ports: a two-stage network-DEA approach," *Transport Policy*, vol. 29, pp. 145–153, 2013.
- [21] V. Chang and B. Tovar, "Drivers explaining the inefficiency of Peruvian and Chilean ports terminals," *Transportation Research Part E: Logistics and Transportation Review*, vol. 67, pp. 190–203, 2014.
- [22] V. Chang and B. Tovar, "Heterogeneity unobserved and efficiency: A latent class model for west coast of south pacific port terminals," *Journal of Transport Economics and Policy*, vol. 51, no. 2, pp. 139–156, 2017.
- [23] B. Tovar and A. Wall, "Can ports increase traffic while reducing inputs? technical efficiency of spanish port authorities using a directional distance function approach," *Transportation Research Part A: Policy and Practice*, vol. 71, pp. 128–140, 2015.
- [24] K. Bichou and R. Gray, "A critical review of conventional terminology for classifying seaports," *Transportation Research Part A: Policy and Practice*, vol. 39, no. 1, pp. 75–92, 2005.
- [25] S. Hosseini and K. Barker, "Modeling infrastructure resilience using Bayesian networks: A case study of inland waterway ports," *Computers & Industrial Engineering*, vol. 93, pp. 252–266, 2016.
- [26] S. Hosseini and K. Barker, "A Bayesian network model for resilience-based supplier selection," *International Journal of Production Economics*, vol. 180, pp. 68–87, 2016.
- [27] J. P. Sierra, A. Genius, P. Lionello, M. Mestres, C. Mösso, and L. Marzo, "Modelling the impact of climate change on harbour operability: the Barcelona port case study," *Ocean Engineering*, vol. 141, pp. 64–78, 2017.
- [28] C. Ugboma, O. Ugboma, and I. C. Ogwude, "An analytic hierarchy process (AHP) approach to port selection decisions – empirical evidence from Nigerian ports," *Maritime Economics & Logistics*, vol. 8, no. 3, pp. 251–266, 2006.
- [29] Y.-T. Chang, S.-Y. Lee, and J. L. Tongzon, "Port selection factors by shipping lines: Different perspectives between trunk liners and feeder service providers," *Marine Policy*, vol. 32, no. 6, pp. 877–885, 2008.
- [30] D. A. Gohomene, Z. L. Yang, S. Bonsal, E. Maistralis, J. Wang, and K. X. Li, "The attractiveness of ports in West Africa: some lessons from shipping lines' port selection," *Growth and Change*, vol. 47, no. 3, pp. 416–426, 2016.
- [31] F. Nur, M. Marufuzzaman, R. Burch, S. M. Puryear, and E. S. Wall, "Analyzing the competitiveness of inland waterway ports: An application of stochastic analytical hierarchy process," in *Proceedings of the Industrial and Systems Engineering Conference*, Orlando, FL, USA, 2018.



- [32] D.-W. Song and K.-T. Yeo, "A competitive analysis of Chinese container ports using the analytic hierarchy process," *Maritime Economics & Logistics*, vol. 6, no. 1, pp. 34–52, 2004.
- [33] G.-T. Yeo, M. Roe, and J. Dinwoodie, "Measuring the competitiveness of container ports: Logisticians' perspectives," *European Journal of Marketing*, vol. 45, no. 3, pp. 455–470, 2011.
- [34] N. Fenton and M. Neil, *Risk Assessment and Decision Analysis with Bayesian Networks*, CRC Press, 2012.
- [35] N. U. I. Hossain, R. Jaradat, S. Hosseini, M. Marufuzzaman, and R. K. Buchanan, "A framework for modeling and assessing system resilience using a Bayesian network: A case study of an interdependent electrical infrastructure system," *International Journal of Critical Infrastructure Protection*, vol. 25, pp. 62–83, 2019.
- [36] M. Zhang, L. Wang, S. Jajodia, A. Singhal, and M. Albanese, "Network diversity: a security metric for evaluating the resilience of networks against zero-day attacks," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 5, pp. 1071–1086, 2016.
- [37] B. Song, C. Lee, and Y. Park, "Assessing the risks of service failures based on ripple effects: A Bayesian network approach," *International Journal of Production Economics*, vol. 141, no. 2, pp. 493–504, 2013.
- [38] M. Hänninen, "Bayesian networks for maritime traffic accident prevention: benefits and challenges," *Accident Analysis & Prevention*, vol. 73, pp. 305–312, 2014.
- [39] S. Hosseini, A. Al Khaled, and M. D. Sarder, "A general framework for assessing system resilience using Bayesian networks: A case study of sulfuric acid manufacturer," *Journal of Manufacturing Systems*, vol. 41, pp. 211–227, 2016.
- [40] A. Grêt-Regamey, S. H. Brunner, J. Altwegg, and P. Bebi, "Facing uncertainty in ecosystem services-based resource management," *Journal of Environmental Management*, vol. 127, pp. S145–S154, 2013.
- [41] E. Pérez-Miñana, P. J. Krause, and J. Thornton, "Bayesian Networks for the management of greenhouse gas emissions in the British agricultural sector," *Environmental Modeling and Software*, vol. 35, pp. 132–148, 2012.
- [42] L. M. Saini, "Peak load forecasting using Bayesian regularization, Resilient and adaptive backpropagation learning based artificial neural networks," *Electric Power Systems Research*, vol. 78, no. 7, pp. 1302–1310, 2008.
- [43] C. Arizmendi, D. A. Sierra, A. Vellido, and E. Romero, "Automated classification of brain tumours from short echo time in vivo MRS data using Gaussian Decomposition and Bayesian Neural Networks," *Expert Systems with Applications*, vol. 41, no. 11, pp. 5296–5307, 2014.
- [44] S. Hosseini and M. D. Sarder, "Development of a Bayesian network model for optimal site selection of electric vehicle charging station," *International Journal of Electrical Power & Energy Systems*, vol. 105, pp. 110–122, 2019.
- [45] S. Hosseini, D. Ivanov, and A. Dolgui, "Review of quantitative methods for supply chain resilience analysis," *Transportation Research Part E: Logistics and Transportation Review*, vol. 125, Part E, pp. 285–307, 2019.
- [46] S. Hosseini, N. Morshedlou, D. Ivanov, M. D. Sarder, K. Barker, and A. Al Khaled, "Resilient supplier selection and optimal order allocation under disruption risks," *International Journal of Production Economics*, vol. 213, pp. 124–137, 2019.
- [47] M. Perkusich, G. Soares, H. Almeida, and A. Perkusich, "A procedure to detect problems of processes in software development projects using Bayesian networks," *Expert Systems with Applications*, vol. 42, no. 1, pp. 437–450, 2015.
- [48] Y. Zhou, C. Li, C. Zhou, and H. Luo, "Using Bayesian network for safety risk analysis of diaphragm wall deflection based on field data," *Reliability Engineering & System Safety*, vol. 180, pp. 152–167, 2018.
- [49] U.S. Department of Transportation, "Port performance freight statistics program," 2017, <https://www.bts.gov/sites/bts.dot.gov/files/docs/browse-statistical-products-and-data/port-performance/216906/port-performance-2017-revised-2-12-18.pdf>.
- [50] J. P. Rodrigue, C. Comtois, and B. Slack, *The Geography of Transport Systems*, Routledge, 2009.
- [51] E. D. Vugrin, D. E. Warren, and M. A. Ehlen, "A resilience assessment framework for infrastructure and economic systems: quantitative and qualitative resilience analysis of petrochemical supply chains to a hurricane," *Process Safety Progress*, vol. 30, no. 3, pp. 280–290, 2011.
- [52] S. Hosseini, K. Barker, and J. E. Ramirez-Marquez, "A review of definitions and measures of system resilience," *Reliability Engineering & System Safety*, vol. 145, pp. 47–61, 2016.
- [53] B. Cai, M. Xie, Y. Liu, Y. Liu, and Q. Feng, "Availability-based engineering resilience metric and its corresponding evaluation methodology," *Reliability Engineering & System Safety*, vol. 172, pp. 216–224, 2018.
- [54] Q. Feng and Y. Ren, "Resilience design method based on meta-structure: a case study of offshore wind farm," *Reliability Engineering System Safety*, 2019.
- [55] Pacific merchant shipping association, "Container dwell time – Metrics for port efficiency & velocity," 2016, <http://www.pmsaship.com/pdfs/PMSA%20Press%20Release%20Container%20Dwell%20Time%202011-17-16.pdf>.
- [56] J.-P. Rodrigue and T. Notteboom, "The terminalization of supply chains: Reassessing the role of terminals in port/hinterland logistical relationships," *Maritime Policy & Management*, vol. 36, no. 2, pp. 165–183, 2009.

## Research Article

# Stability and Complexity of a Novel Three-Dimensional Environmental Quality Dynamic Evolution System

LiuWei Zhao <sup>1,2</sup> and Charles Oduro Acheampong Otoo<sup>2</sup>

<sup>1</sup>School of Business, Jiangsu University of Technology, Changzhou, Jiangsu 213000, China

<sup>2</sup>Computational Experiment Center for Social Science, School of Management, Jiangsu University, Zhenjiang, Jiangsu 212013, China

Correspondence should be addressed to LiuWei Zhao; 136901672@qq.com

Received 29 December 2018; Revised 24 February 2019; Accepted 10 March 2019; Published 17 April 2019

Guest Editor: Seyedmohsen Hosseini

Copyright © 2019 LiuWei Zhao and Charles Oduro Acheampong Otoo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, a novel three-dimensional environmental quality dynamic system is introduced. Bayesian estimation was used to calibrate environmental quality variables, and Genetic algorithm (GA) optimized Levenberg-Marquardt Back Propagation (LM-BP) neural network method was used to effectively identify the system parameters for calibration of various variables and official data. The studies found that the effect of increasing investment in environmental protection on energy intensity and environmental quality is not obvious, and it also aggravates the economic instability. Adjustment of peak parameters of pollution emissions can accelerate the evolution of energy intensity and environmental quality to a stable speed and eventually stabilize with a certain value. But if the peak value of pollution emissions reaches too early, it will pose a certain threat to the environment. Although the speed of ecological environment self-repair is increased, it cannot effectively reduce energy intensity, improve environmental quality, and maintain economic growth; it can control the stability of the control system or effectively control pollution. Therefore, in order to improve the environmental quality, we need to take more measures in parallel, use more means and resources for environmental governance, and ultimately achieve “win-win” between environmental quality and economy.

## 1. Introduction

With the rapid development of economy and society, a series of problems such as shortage of resources, environmental degradation, and ecological destruction have become increasingly prominent. Since the evolution of the environment complex giant system follows the nonlinear mechanism, the new quality will be continuously innovated, and the world will become more diversified and complicated. It can be said that nonlinearity is the basic guarantee for the enrichment and complexity of the environmental system. Just because more and more new substances are produced, evolution continues, substances become more complex, and the amount of information contained increases and accumulates by nonlinear action. Only in this way can development be sustainable. The effect of complex nonlinear is common between the environmental system and the external and internal elements of the system, so that the system forms a relatively stable structure, organizational model, and control

mechanism, which can stimulate or limit the evolution of the whole system. In addition, the exchange of matter, energy, and information in the cyclic motion of environmental systems and subsystems can lead to chaotic or disordered states, which requires humans to fully understand and guide the chaotic or disordered of environmental systems in a coordinated and orderly manner of orbit to evolve.

At present, the study of the complexity of the nonlinearity of ring systems has been more extensive. For example, Gudo Buenstor pointed out that the ecosystem is a self-organizing system, and its sustainable development is related to the self-organizing characteristics of the system. If various environmental problems will affect the process of economic development, then the thermodynamic changes of ecosystem (energy and entropy transformation) will have a guiding significance for the corresponding countermeasures of economic system [1]. Luis A et al. pointed out that there are relatively few studies on how human beings intervene in environmental systems and how natural phenomena

lead to changes in ecosystems [2]. Ranjit Kumar Upadhyay and others believe that structural complexity is one of the most fundamental problems of dynamic complexity of ecosystems. The development and application of dynamic system theory are helpful to understand the complexity of ecosystem. It is pointed out that structural complexity and dynamic complexity are interrelated. Simple dynamic models of ecosystem show that simple structural systems can produce very complex and unpredictable dynamic behavior in some cases. However, the ecosystems with complex structures may not necessarily produce complex dynamic behaviors [3]. Sazykina and Milan demonstrated that the nutrient structure of ecosystem is a self-organizing process through dynamic analysis of ecosystem models [4, 5]. Bianciardi proposed an experimental model for quantifying the complexity and self-organizing characteristics of complex ecosystems. When the input variables of the system change, the complexity of the system changes, and the system needs to reorganize and continue to change to another equilibrium state [6]. Wu et al. expounded that environmental system is a typical complex system, which has the characteristics of nonlinearity, irreversibility, multilevel, openness, self-organization, and criticality [7]. For example, there are many species in the environmental system. There are complex relationships among species, such as parasitism, symbiosis, and natural enemies. There are direct or indirect connections among all species, which constitute a complex ecological network. Because the environmental system has the characteristics of nonlinearity, self-organization, nonnegotiability, dynamics, openness, multilevel, self-similarity and so on, it is a typical complex system, and it is a large system with many elements, levels, and complex relationships. If we want to solve the current environmental problems facing mankind, we must apply complexity theory to environmental problems making the structure of environmental system more reasonable and the balance of environment, economy, and society rebuilt.

Chaos theory is one of the key methods to study complex dynamics. It was originally developed under the background of physics. However, many scholars have found that social, ecological, and economic systems are all nonlinear, while nonlinear relationships evolve dynamically over time. Based on the hypothesis of participant's bounded rationality, a novel Cournot Duopoly game model of carbon emission reduction is established, and the dynamic adjustment mechanism of emission reduction on enterprises participated in is also analyzed [8]. Fang et al. obtained the attractors of energy-saving and emission-reduction by a series of energy saving and emission reduction system models. Using the energy intensity formula varying with time, the effectiveness of energy-saving and emission-reduction was evaluated [9–12]. Yuan et al. analyzed the relationship between energy intensity and technological progress based on Douglas production function [13]. Long used game theory to analyze the optimal strategies between the government and enterprises in the process of implementing energy-saving and emission-reduction actions and obtained Nash equilibrium solution of mixed strategies [14].

At present, China's environmental protection is still lagging economic and social development. The problem of

multistage, multifield, and multitype pollution has accumulated on a long time. The environmental carrying capacity has reached or approached the upper limit, and the trend of ecological environment deterioration has not been fundamentally reversed. In the process of dealing with these problems, people gradually realized that how to deal with resource utilization problems and environmental problems and control pollutant emissions is the key to improving environmental quality. It is related to the sustained and healthy development of the economy and society. An important breakthrough in adjusting the economic structure and realizing the mode of economic growth is to establish a new economic competition and introduce a good opportunity for green development and environmental protection industry development.

In the process of further promoting environmental governance measures in China, the environmental problems left behind have become more complex and difficult to solve, and the work of further promoting environmental governance and quality improvement has become more arduous and complex. At present, with the acceleration of industrialization, urbanization, and modernization in China, the demand for energy is in a rapid growth stage. The sustained growth of fossil energy consumption and the prominent "high pollution" characteristic of "high emission" have become a major constraint on China's sustainable development. Although the quality of China's ecological environment has improved with active efforts, the complexity, urgency, and long-term nature of environmental issues have not changed. We must clearly understand and grasp the grim current situation of China's ecological environment, and we must rationally, objectively, and persistently promote the improvement of environmental quality. We should fully understand the complexity of the environment and the lag of environmental protection, reasonably reduce environmental pollution, and effectively improve economic growth to meet the people's pursuit of a beautiful ecological environment and higher quality of life. It plays an important role in alleviating the contradiction between human and nature and enhancing the effect of environmental protection. At the same time, the existing research extends from the focus on improving the effect of environmental governance to the idea of environmental governance system engineering, but the research on the mechanism of energy saving and economic growth to improve environmental quality needs to be improved. It is necessary to construct an environmental quality control system suitable for China's national conditions in this situation. Reasonably coordinate the relationship between various variables in the environmental quality management system, explore the vector sensitive to the environmental quality index in the environmental quality system, and find practical and feasible ways to reduce environmental pollution and have no significant impact on the economy. Provide theoretical basis and technical support for China to formulate policies and regulations to improve environmental quality and effectively promote the development of environmental management. This is the key to study the mechanism of improving environmental quality with strong theoretical expansion and practical application value.

This article is organized as follows. In Section 2, we have model building and interpretation, dynamic analysis, and showing the dynamic features of this system with numerical simulations, including bifurcation diagram, phase portrait, and sensitive dependence on initial conditions. In Section 3, we have data acquisition and processing of the model and LM-BP neural network is optimized based on genetic algorithm to identify system parameters. In Section 4, the depth of the key influential parameters is analyzed and studied in the identified model. Section 5 is the main research conclusion of this study.

## 2. Model Establishment and Analysis

**2.1. Modeling and Interpretation.** In this paper, we consider a new three-dimensional energy and economic growth system with environmental quality constraints, which can better simulate the actual situation and meet the requirements of practical development. The dynamic system is in the form of the following differential equations:

$$\begin{aligned}\dot{x} &= a_1 x \left(1 - \frac{x}{F}\right) + a_2 y \left(1 - \frac{y}{E}\right) - a_3 z \\ \dot{y} &= -b_1 x - b_2 y - b_3 z \\ \dot{z} &= -c_1 x + c_2 y \left(1 - \frac{y}{H}\right) + c_3 z \left(\frac{x}{P} - 1\right)\end{aligned}\quad (1)$$

where  $x(t)$ ,  $y(t)$ ,  $z(t)$ , respectively, are the level of pollution emission, the level of economic growth (GDP), and the level of environmental quality in period  $t$ ,  $a_i$ ,  $b_i$ ,  $c_i$  ( $i = 1, 2, 3, 4$ ),  $F$ ,  $E$ ,  $H$ ,  $P$  are normal numbers,  $t \in I$ , and  $I$  is an economic cycle.  $a_1$  represents the development rate elasticity coefficient of pollution emission  $x(t)$ ,  $a_2$  represents the influence coefficient of economic growth  $y(t)$  on pollution emission  $x(t)$ ,  $a_3$  denotes the inhibition coefficient of environmental quality  $z(t)$  on pollution emission  $x(t)$ ,  $F$  represents the peak value of pollution emission  $x(t)$  in an economic cycle,  $E$  represents the peak value of economic growth  $y(t)$  in an economic cycle;  $b_1$  represents the influence coefficient of pollution emission  $x(t)$  on economic growth  $y(t)$ ,  $b_2$  represents the restraint factor of investment in reduction pollution emission  $x(t)$  on economic growth  $y(t)$ ,  $b_3$  represents the restraint factor of investment in improving environmental quality  $z(t)$  on economic growth  $y(t)$ ;  $c_1$  represents the inhibition coefficient of pollution emission  $x(t)$  on environmental quality  $z(t)$ ,  $c_2$  represents the influence coefficient of economic growth  $y(t)$ 's investment on improving environmental quality  $z(t)$ ,  $c_3$  represents the speed coefficient of ecological environment self-repair without external intervention,  $H$  represents the peak value of the impact of economic growth  $y(t)$  on environmental quality  $z(t)$  in an economic cycle, and  $P$  represents the maximum amount of pollution that can be contained by the ecological environment.

In system (1),  $\dot{x}$  represents the level of change in pollution emissions in  $t$  period;  $\dot{y}$  represents the growth level of economic growth in the  $t$  period;  $\dot{z}$  represents the change level of environmental quality in  $t$  period. When the impact of pollution exceeds the maximum capacity of the ecological

environment (that  $\dot{z} > 0$ ), it means that the ecological environment is deteriorating. In this case, the ecosystem cannot rely on self-regulation to repair. Therefore, when this situation continues, the ecological environment will eventually be sold out; when  $\dot{z} \leq 0$ , it means that the environmental quality has reached a dynamic balance or gradually improved, and only under this condition can the ecological environment be the basis of human social development [15–17].

Energy mildness is the ratio of energy utilization to economic or material output. The energy intensity of a country or region is usually expressed in terms of energy consumption per unit of gross domestic product as follows:

Energy intensity

$$= \frac{\text{Energy consumption over an economic period}}{\text{GDP over an economic period}} \quad (2)$$

In other words,  $x^*(t) = \varphi_1(t, kx, y, z)$  and  $y(t) = \varphi_2(t, x, y, z)$  can be deduced from system (1), where  $x^*(t)$  represents energy consumption over time in an economic period;  $k = 1/k_0$ ,  $k_0$  represents the pollutant discharge coefficient of standard coal. Therefore, the energy intensity in an economic period can be expressed as follows:

$$U(t) = \frac{\varphi_1(t, kx, y, z)}{\varphi_2(t, x, y, z)}, \quad t \in I. \quad (3)$$

**Lemma 1** (see [18]). *This is the  $n$  dimensional discrete dynamical system and all the eigenvalues of Jacobian matrix  $J(x^*)$  of the right function in this system must satisfy the condition that  $|\lambda_n| < 1$  in order to stabilize the equilibrium point  $x^*$ .*

$$\begin{aligned}\dot{x}_1 &= f_1(x_1(t), x_2(t), \dots, x_n(t)) \\ \dot{x}_2 &= f_2(x_1(t), x_2(t), \dots, x_n(t)) \\ &\dots \\ \dot{x}_n &= f_n(x_1(t), x_2(t), \dots, x_n(t)).\end{aligned}\quad (A1)$$

**2.2. Dynamic Analysis of the Model.** Then we will analyze the dynamics of environmental quality system. System (1) is determined by the eigenvalues of the Jacobian matrix as follows:

$$J = \begin{bmatrix} \frac{(F-2x)a_1}{F} & \frac{(E-2y)a_2}{E} & -a_3 \\ -b_1 & -b_2 & -b_3 \\ -c_1 + \frac{zc_3}{P} & \frac{(H-2y)c_2}{H} & \left(\frac{x}{P} - 1\right)c_3 \end{bmatrix}. \quad (4)$$

System (1) is a very complex dynamic system, when  $a_i$ ,  $b_i$ ,  $c_i$ ,  $F$ ,  $E$ ,  $H$ ,  $P$  values are different, system (1) also shows different dynamic behavior. Therefore, in order to facilitate the study of system (1), the coefficients in system (1) will be set as follows:  $a_1 = 0.055$ ,  $a_2 = 0.896$ ,  $a_3 = 0.114$ ,  $b_1 = 0.759$ ,  $b_2 = 0.075$ ,  $b_3 = 0.003$ ,  $c_1 = 0.355$ ,  $c_2 = 0.674$ ,  $c_3 = 0.8127$ ,  $F = 1.6$ ,  $E = 2.55$ ,  $H = 0.25$ ,  $P = 2.4$ .

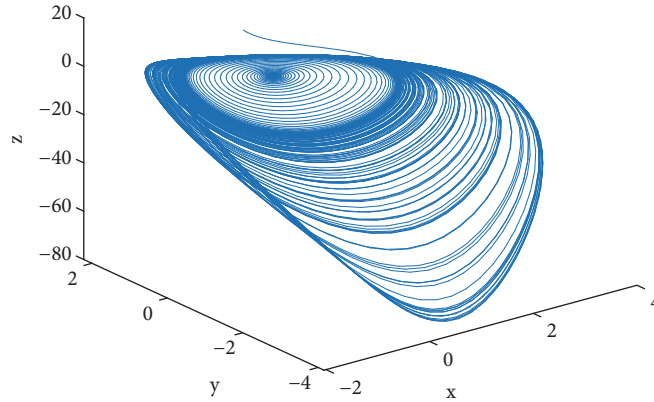


FIGURE 1: Three-dimensional attractor diagram of system (1): 3D  $(x - y - z)$ .

Setting parameters are brought into system (1). By calculation, we get the two real equilibrium points of system (1) respectively:  $O(0, 0, 0)$  and  $S(0.7118, -3.9918, -80.2967)$ .

For equilibrium  $O(0, 0, 0)$ , system (1) matrix of linear approximation system is as follows:

$$J_0 = \begin{bmatrix} a_1 & a_2 & -a_3 \\ -b_1 & -b_2 & -b_3 \\ -c_1 & c_2 & -c_3 \end{bmatrix}, \quad (5)$$

If  $f(\lambda)$  denotes the characteristic polynomial of the Jacobian matrix  $J(O)$ , then

$$f(\lambda) = \lambda^3 + A\lambda^2 + B\lambda + C = 0, \quad (6)$$

where  $A = b_2 + c_3 - a_1$ ,  $B = a_2b_1 + a_3c_1 + b_3c_2 + b_2c_3 - a_1(b_2 + c_3)$ ,  $C = a_3b_2c_1 - a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 + a_2b_1c_3 - a_1b_2c_3$ .

By simple calculation, we get three eigenvalues of the matrix  $J(O)$ :  $\lambda_1 = -0.7876$ ,  $\lambda_{2,3} = -0.0225 \pm 0.78595i$ . According to Lemma 1,  $\lambda_i < 1 (i = 1, 2, 3)$ , that equilibrium point  $O(0, 0, 0)$  is unstable.

Let  $b_1$  be any value, the corresponding characteristic equation is as follows:

$$f(\lambda) = \lambda^3 + 0.8327\lambda^2 + b_1(0.6513 + 0.896\lambda) - 0.0263\lambda - 0.0075 = 0. \quad (7)$$

Let  $p_1 = 0.8327$ ,  $p_2 = 0.896b_1 - 0.026319$ ,  $p_3 = -0.0075 + 0.6513b_1$ ; according to Routh-Hurwitz criterion, the necessary and sufficient conditions for the real part of all characteristic roots of the equation to be negative are as follows:  $p_1 = 0.8327 > 0$ ,  $p_1p_2 - p_3 > 0$ ,  $p_1p_2p_3 - p_3^2 > 0$ . By calculation, we know when  $0 < b_1 < 0.0114$  that equilibrium point  $O(0, 0, 0)$  is unstable. By calculating, the Jacobian matrix of system (1) at the equilibrium point  $S$ , we get the eigenvalue  $\lambda_1 = 0.8405$ ,  $\lambda_{2,3} = -0.7405 \pm 0.6863i$ .

$$\begin{aligned} \nabla V &= \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} = \frac{(F - 2x)a_1}{F} - b_2 + \left(\frac{x}{P} - 1\right)c_3 \\ &= a_1 - b_2 - c_3 + x\left(\frac{c_3}{P} - \frac{a_1}{F}\right) \end{aligned} \quad (8)$$

When  $c_3/P - a_1/F = 0$ ,  $a_1 - b_2 - c_3 < 0$ , system (1) is dissipative.

Next, the dynamic simulation of system (1) is carried out, setting the parameter values of the system is substituted with initial values  $[0.758, 1.83, 1.5]$ . Three-dimensional attractor diagrams of system (1) are given in Figure 1; two-dimensional attractor diagrams of system (1) are given in Figure 2; Figure 3 shows the time series diagram of system (1)  $(x(t), y(t), z(t))$ ; Figure 4 shows bifurcation diagram and corresponding Lyapunov exponential diagram of the variable  $b_1 \in [0.7, 1.3]$  change of system (1); sensitive dependence of the dynamical system (1) on initial conditions is given in Figure 5.

When time tends to be infinite, all orbits of unsteady flow on any bounded set tend to a bounded set called attractor, and such a set also has very complex geometric structure. Figure 1 shows that the trajectory of system (1) moves irregularly into the set of shapes with the increase of time. Figure 2 shows the different prints of the phase of system (1) on the two-dimensional plane, which also fully reflects that the trajectory of system (1) is irregular. Because the attractors are closely related to chaos, it is necessary to explore the properties of attractor sets to better reveal the law and structure of chaos in system (1). Figure 4 shows the bifurcation diagram and the corresponding Lyapunov exponential diagram of system (1) with the change of parameter  $b_1 \in [0.7, 1.3]$ ; we found the very complex dynamic behavior of system (1) from Figure 4. The Lyapunov exponential diagram is a quantitative quantity of system stability. The maximum Lyapunov exponent is greater than 0, which indicates that the system has chaotic behavior. For a four-dimensional system, if the maximum Lyapunov exponent at one point is equal to 0 and the rest is less than 0, it can indicate that the system has bifurcation. Impact of initial state on the system (1), Figure 5 gives the sensitive dependence on initial values. The sensitivity of system (when losing stability) with  $(x, y, z) = (0.758, 1.83, 1.5)$  and  $(x_1, y, z) = (0.757, 1.83, 1.5)$  and small changes of the initial conditions may cause the observed large changes of the system, which is sensitive to initial states. The butterfly effect is also an important symbol of chaotic motions. Figure 5 shows the difference among the different orbits with slightly deviated initial values which builds up rapidly after many iterations, although their initial states are indistinguishable.

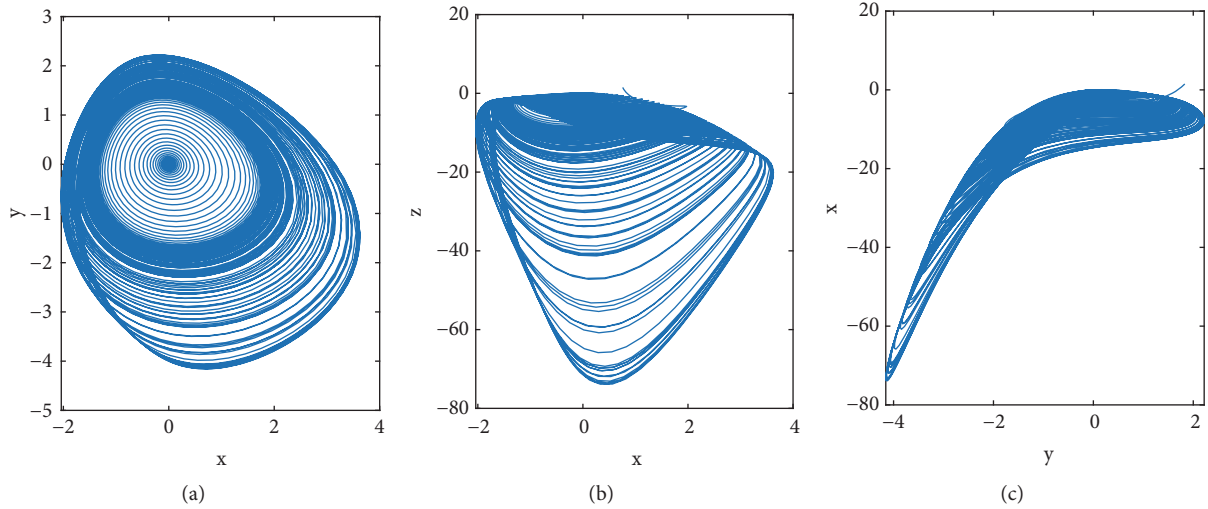


FIGURE 2: Two-dimensional attractor diagram of system (1): (a) 2D  $(x - y)$ , (b) 2D  $(x - z)$ , (c) 2D  $(y - z)$ .

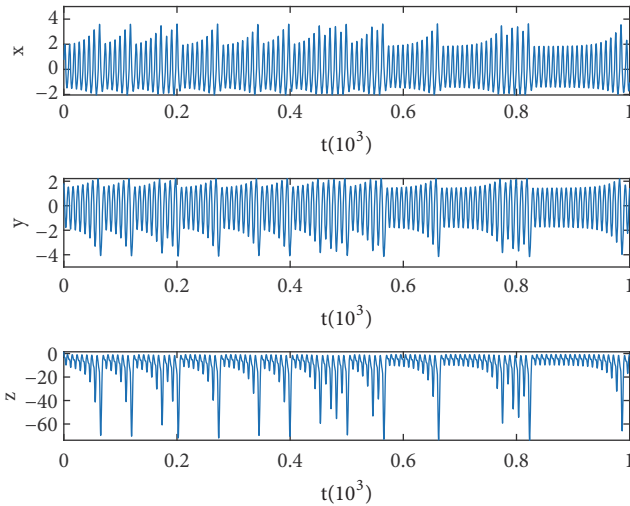


FIGURE 3: Time series diagram of system (1)  $(x(t), y(t), z(t))$ .

### 3. Parameter Identification of the System Based on GA Optimization LM-BP Neural Network Method

**3.1. Data Acquisition.** The system is based on the complex relationship between pollution emission reduction, economic growth, and environmental quality. The determination of parameters in the system is of great significance to practical research. Based on the statistical data of China, the parameters of the actual system are obtained by GA optimizing LM-BP neural network method. The evolutionary relationship between actual pollution emission, economic growth, and environmental quality is analyzed. The environmental quality, energy intensity, and economic growth are affected with those variables.

In this paper, according to China Statistical Yearbook 2018, the main indicators of pollution emissions are energy and carbon emissions caused by energy consumption in a

cycle, and GDP is the main choice for economic growth. Environmental quality is the key variable in this paper, but how to measure it accurately is one of the difficulties faced by relevant research. This paper mainly refers to the research on the quality of economic growth of Chao Xiaojing and Ren Baoping and calculates the environmental quality index using the principal component analysis method (PCA) based on covariance matrix [19]. In order to ensure the consistency of statistical caliber (in 2017, the Ministry of Ecology and Environment revised the standard system, survey methods, and related technical regulations of the statistical system. For this reason, the data selected in this paper are up to 2017.), we selected the per capita urban green space area (hectare/urban population), forest coverage (%), and the area occupied by nature reserves (%) from 2000 to 2017 as the positive indicators of environmental status. The per capita water resources (cubic meters/person), per capita industrial wastewater discharge (ton/person), and per capita industrial solid waste production (ton/person) were taken as the positive indicators of environmental status. Average industrial exhaust emissions (cubic meters per person) are regarded as the inverse indicators of potential environmental pressure, and all the inverse indicators are treated by reciprocal positive treatment, so that the effect of all indicators on environmental quality is similar. In view of the difference of dimension and magnitude, the results of main components will be biased to the indexes with larger variance or magnitude, so this paper treats all the indexes by means of dimensionless processing. By observing the cumulative contribution rate of the main components, the weight of each index is determined by the standard orthogonal eigenvector of the first main component. In addition, in order to make the environmental quality index positive, the logarithmic logistic model is adopted to standardize the data, and all the data are processed based on 1999. The final number is shown in Table 1.

**3.2. System Parameter Identification and Model Validation.** GA optimized LM-BP neural network method has better

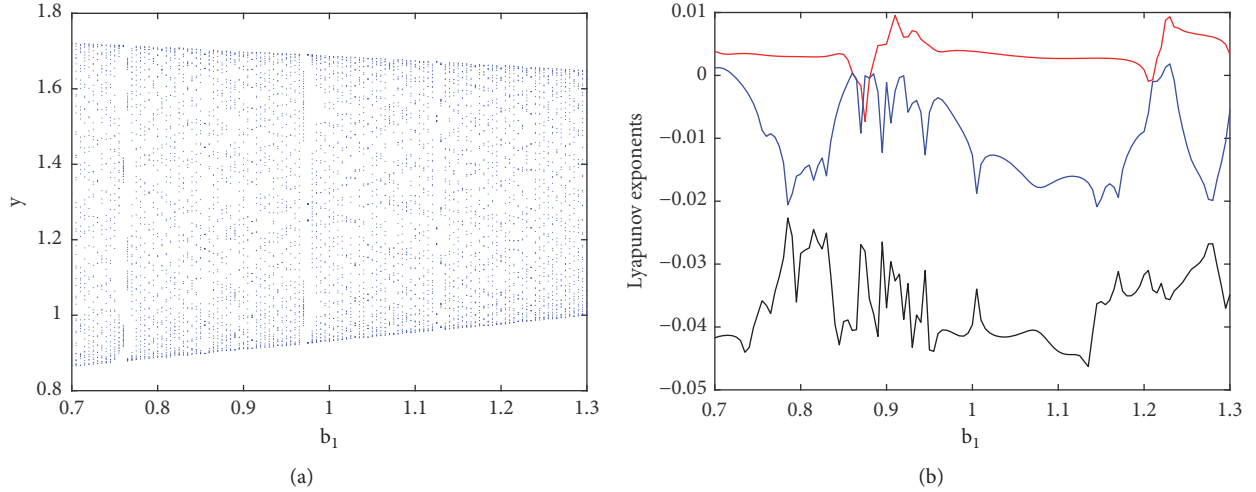


FIGURE 4: Bifurcation diagram and corresponding Lyapunov exponential diagram of the variable  $b_1 \in [0.7, 1.3]$  change of system (1): (a) bifurcation diagram and (b) Lyapunov exponential diagram of corresponding (a).

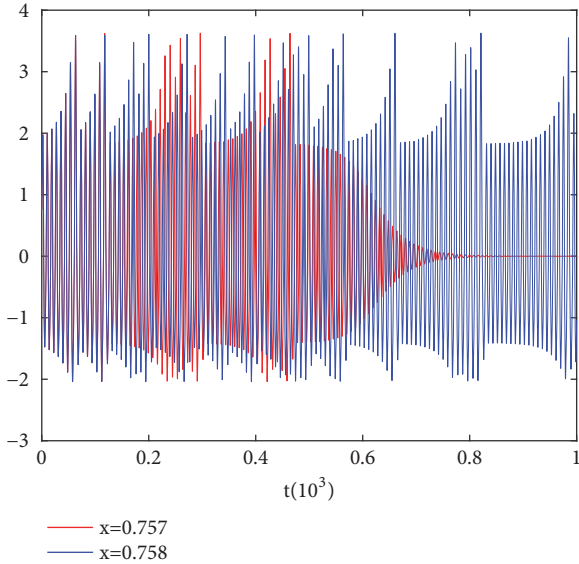


FIGURE 5: Sensitive dependence of the dynamical system (1) on initial conditions.

identification accuracy for the parameter identification of the nonlinear system. Firstly, system (1) is discretized and the following difference equations are obtained:

$$\begin{aligned}
 x(k+1) &= x(k) \\
 &\quad + T \left[ a_1 x \left( 1 - \frac{x}{F} \right) + a_2 y \left( 1 - \frac{y}{E} \right) - a_3 z \right] \\
 y(k+1) &= y(k) + T [-b_1 x - b_2 y - b_3 z] \\
 z(k+1) &= z(k) \\
 &\quad + T \left[ -c_1 x + c_2 y \left( 1 - \frac{y}{H} \right) + c_3 z \left( \frac{x}{P} - 1 \right) \right]
 \end{aligned} \tag{9}$$

TABLE 1: Statistics on energy consumption, economic growth, and environmental quality in China.

Year	x	y	z
2000	1.0455	1.1085	0.9837
2001	1.1066	1.2228	0.9595
2002	1.2064	1.3482	0.9347
2003	1.4020	1.5283	0.8747
2004	1.6382	1.8062	0.7741
2005	1.8594	2.0813	0.7242
2006	2.0380	2.4509	0.6403
2007	2.2156	3.0307	0.5455
2008	2.2808	3.5976	0.4752
2009	2.3912	3.8997	0.4061
2010	2.5656	4.6020	0.3693
2011	2.7534	5.4243	0.3565
2012	2.8608	6.0326	0.3210
2013	2.9659	6.6068	0.2817
2014	3.0292	7.2151	0.2575
2015	3.0583	7.6813	0.2322
2016	3.1004	8.2872	0.2159
2017	3.1942	9.2297	0.2147

The front  $n - 1$  sets of actual statistical data are used as input data of GA optimized LM-BP neural network after  $n - 1$  sets as output data of GA optimized LM-BP neural network, and the data are normalized in  $\bar{x}_i = (x_i - x_{\min}) / (x_{\max} - x_{\min})$  form. All other parameters are random numbers. The appropriate feedforward neural network is selected, and its output results are brought into the system (9). The results are compared with the target output. By comparing and controlling the error at  $10^{-6}$ , the identified system parameters are obtained as shown in Table 2.

In order to verify the feasibility of obtaining identification parameters, this paper chooses the data of 1980 as the initial

TABLE 2: Identifying the actual parameters of system (1) by identifying the actual data.

$a_1$	$a_2$	$a_3$	$b_1$	$b_2$	$b_3$	$c_1$
0.0945	0.0348	0.0161	0.0262	0.0834	0.0734	0.0321
$c_2$	$c_3$	$F$	$E$	$H$	$P$	
0.0151	0.1401	1.5616	2.7837	1.1943	2.3121	

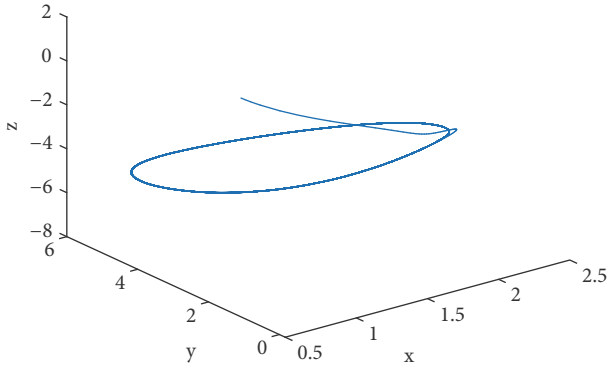


FIGURE 6: Three-dimensional phase diagram of system (1) with identified parameters: 3D ( $x - y - z$ ).

value of the system  $[0.658, 1.73, 1.1211]$  and obtains the actual system diagram as shown in Figures 6 and 7. From the phase diagram of the actual system, we can find that the system is developing steadily, which is consistent with the actual situation.

#### 4. Analysis of the Influencing Factors of the System

In order to better study the stability of system (1) and how to reduce energy intensity and improve environmental quality more effectively, it is necessary to do in-depth analysis and research on some key parameters of system (1).

Figure 8 shows under different values the influence of the system parameter  $c_2$  on the evolution path of energy intensity and environmental quality based on the parameters of system (1) as shown in Table 2 and the initial conditions are unchanged. When  $c_2 = 0.0151$ , system (1) evolves as shown in the red curve in Figure 8; when  $c_2 = 0.0159$ , system (1) evolves as shown in the blue curve in Figure 8. By comparing Figure 8, we can find that the effect of increasing investment in environmental protection on energy intensity and environmental quality is not obvious, and it also aggravates the economic instability; the fluctuation around a central value is not conducive to the control of environmental pollution with the course of time. Therefore, simply increasing investment in environmental protection with the stability of effectively reducing energy intensity and improving environmental quality is not obvious.

Figure 9 shows the effect of different values of the system parameter  $F$  on energy intensity and environmental quality based on the parameters of system (1) as shown in Table 2 and the initial conditions are unchanged. When  $F = 1.5616$ , system (1) evolves as shown in the red curve in Figure 9;

when  $F = 1.2616$ , system (1) evolves as shown in the blue curve in Figure 9. By comparing Figures 9(a) and 9(b), it can be found that when  $F = 1.5616$ , energy intensity and environmental quality fluctuate around a central value in the course of time. This phenomenon is not conducive to controlling pollution emissions (that is, it cannot effectively reflect the effect of energy saving and emission reduction) and improving environmental quality; when  $F = 1.2616$ , the evolution of energy intensity and environmental quality tends to be stable at a faster speed and eventually stabilizes with a certain value, which indicates that the development of a good control system or effective control of pollution can be achieved. In addition, it is found that the peak value of pollution emissions arrives too early, but it poses a certain threat to the environment from Figure 9(b). The reason for this phenomenon is that the peak value of pollution emissions exceeds the speed of natural environment itself and environmental remediation from the outside, and the pollutants in the environment rapidly accumulate to the maximum capacity of the environment. Therefore, it is very important to control the peak value of pollution emission reasonably for the stability of control system and the improvement of environmental quality.

Figure 10 shows the environmental quality under the influence of different system parameters  $c_3$ , based on the parameters of system (1) as shown in Table 2 and the initial conditions are unchanged. Figure 10 shows the parameter  $c_3 = 0.1401$ , and the evolution of system (1) is shown in the red curve; when parameter  $c_3 = 0.2401$ , the evolution form of system (1) is shown in the blue curve. By comparing Figure 10, when the speed of ecological environment self-repair is  $c_3 = 0.1401$ , energy intensity and environmental quality fluctuate around a central value in the course of time, which is not conducive to controlling pollution emissions and improving environmental quality; when  $c_3 = 0.2401$ , energy intensity, environmental quality, and economy eventually evolve into stable values. Although it cannot effectively reduce energy intensity, improve environmental quality, and maintain economic growth, it can control the stability of the control system or effectively control pollution.

Figure 11 shows the parameter  $P = 2.3121$ , and the evolution form of system (1) is shown in the red curve; when parameter  $P = 2.0121$ , the evolution form of system (1) is shown in the blue curve; when parameter  $P = 2.8121$ , the evolution form of system (1) is shown by green curve. By comparing Figure 11, it can be found that, in the short term, the environmental capacity has no obvious impact on the system, but when the environmental capacity decreases, the environmental quality fluctuates around a central value in the course of time and the vibration amplitude increases when the environmental capacity decreases to a certain extent (that is, when the pollution impact exceeds the ecological environment). When the amount of environmental pollution reaches the maximum capacity of the ecological environment system, the environmental system will collapse: that is, the ecological environment will deteriorate and the ecosystem cannot rely on self-regulation to repair in this situation. As a result, the situation continues, and the ecological environment is eventually sold out.



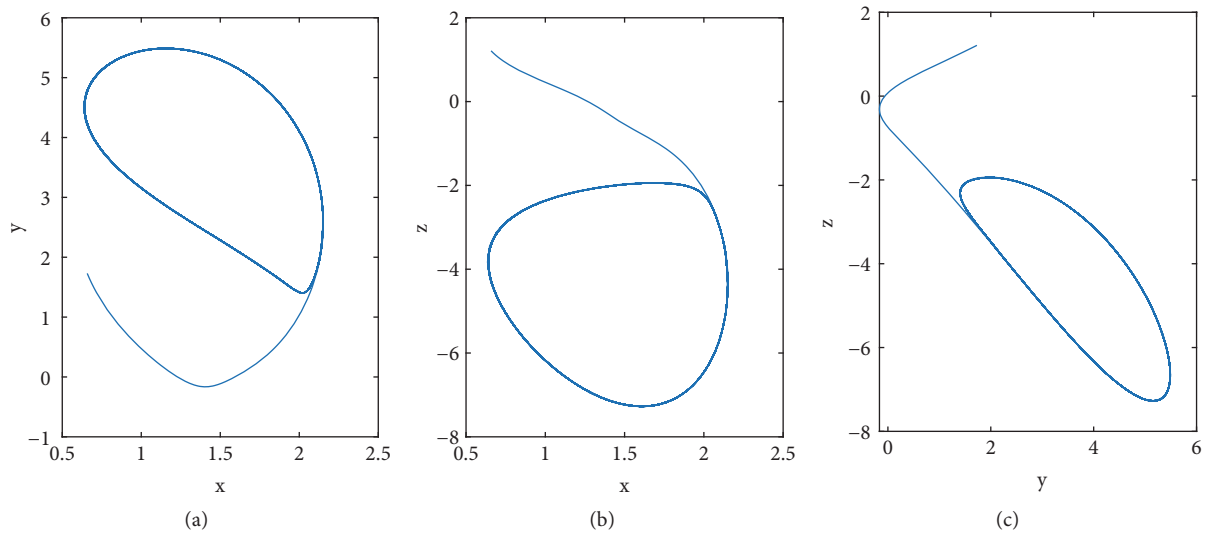


FIGURE 7: Two-dimensional phase diagram of system (1) with identified parameters: (a) 2D  $(x - y)$ , (b) 2D  $(x - z)$ , (c) 2D  $(y - z)$ .

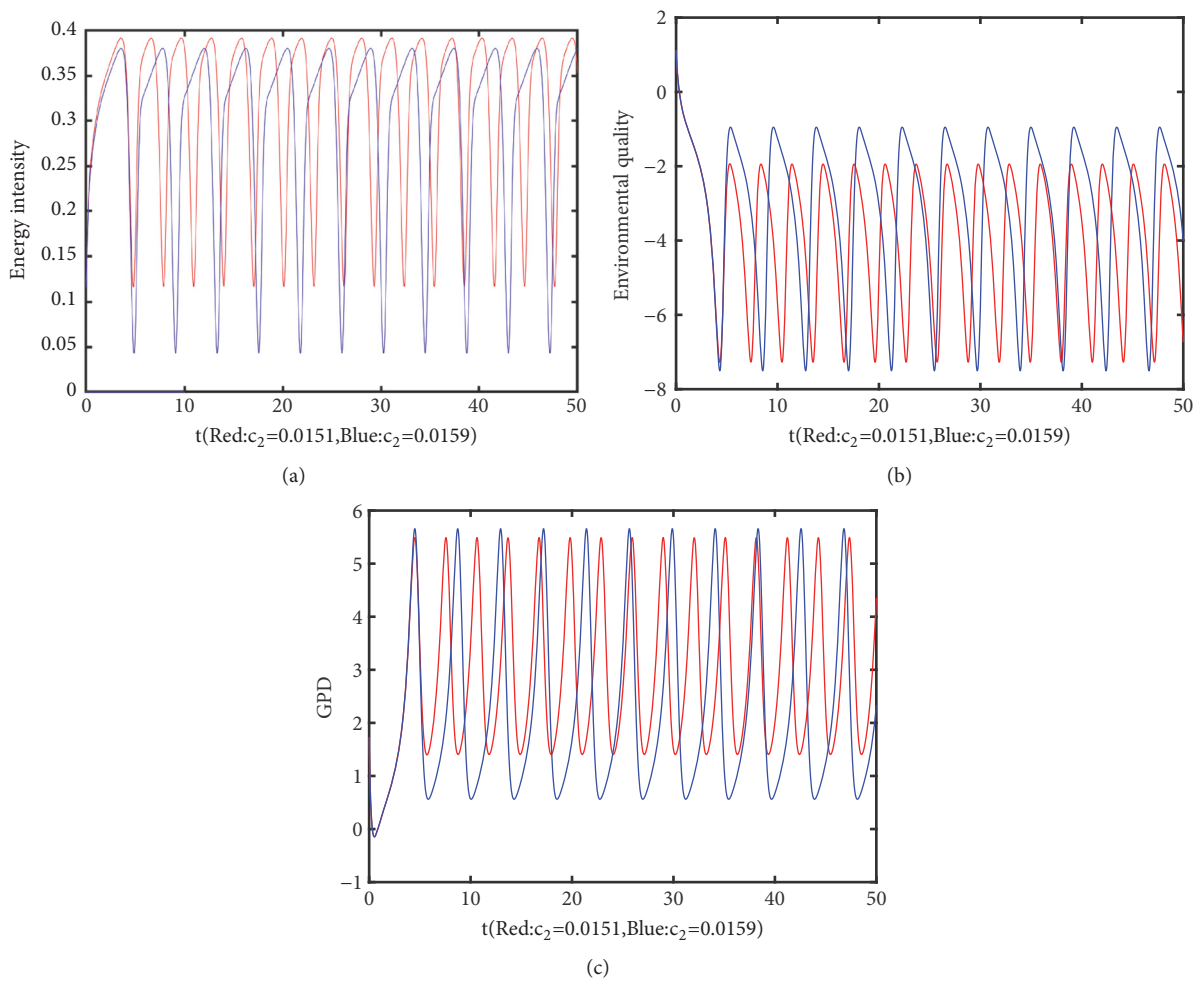


FIGURE 8: The influence of system (1) parameters  $c_2$  on energy intensity and environmental quality: (a) energy intensity, (b) environmental quality, and (c) GDP.

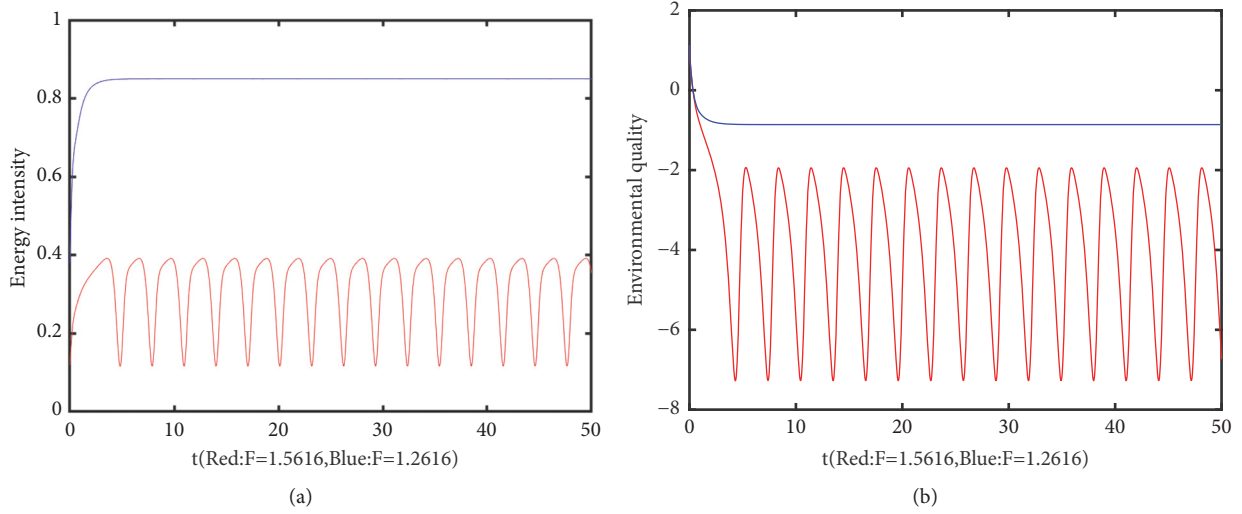


FIGURE 9: The influence of system (1) parameters  $F$  on energy intensity and environmental quality: (a) energy intensity and (b) environmental quality.

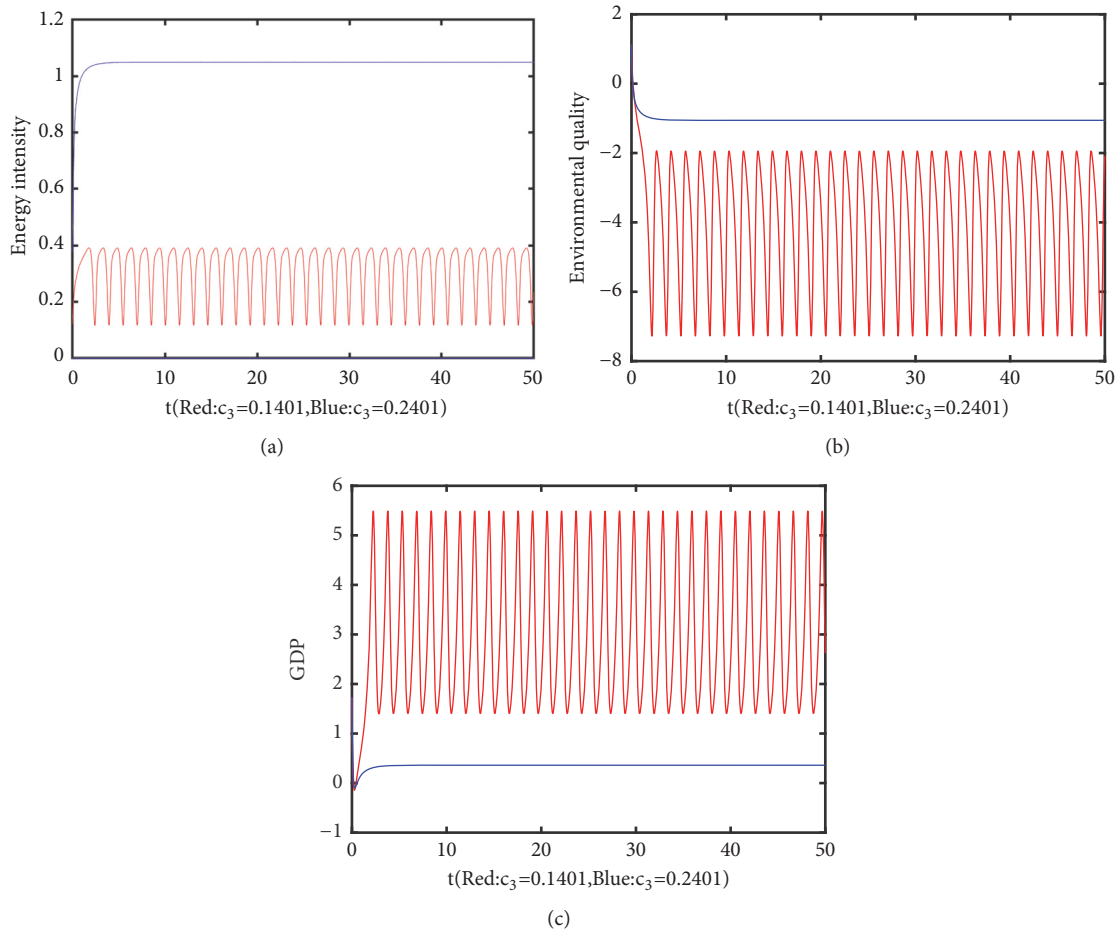


FIGURE 10: The influence of system (1) parameters  $c_3$  on energy intensity and environmental quality: (a) energy intensity, (b) environmental quality, and (c) GDP.

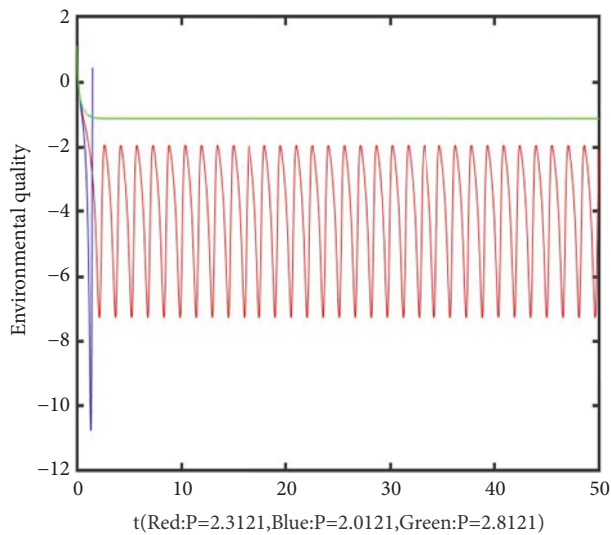


FIGURE 11: The identified system (1): the influence of environmental quality under the different parameters  $P$ .

## 5. Conclusion

In this paper, a model of energy-saving and emission-reduction with environmental constraints is constructed by using dynamic large-scale system modeling method. The stability and complexity of the model are analyzed in depth by applying system science analysis theory, game equilibrium theory, control optimization theory, complex system analysis, and decision-making theory model. Then the environmental quality variables are calibrated by Bayesian estimation, and the system parameters are effectively identified by GA optimized LM-BP neural network method for calibration of variables and official data. Finally, the influence of the change of key parameters on the stability, energy intensity, and environmental quality of the system is analyzed. Therefore, in order to improve the environmental quality, we need to take more measures in parallel, use more means and resources for environmental governance, and ultimately achieve “win-win” between environmental quality and economy.

## Data Availability

The numerical simulations data used to support the findings of this study were supplied by LiuWei Zhao under license and so cannot be made freely available. Requests for access to these data should be made to LiuWei Zhao, E-mail address: 136901672@qq.com.

## Conflicts of Interest

The authors of this paper have no conflicts of interest.

## Acknowledgments

This work is supported by the National Nature Science Foundation of China (No. 71471076) and Jiangsu University of Technology Talent Introduction Project (KYY18539).

## References

- [1] G. Buenstorf, “Self-organization and sustainability: energetics of evolution and implications for ecological economics,” *Ecological Economics*, vol. 33, no. 1, pp. 119–134, 2000.
- [2] A. Luis, B. Rquez-tapia, and L. Jus, “Integrating fuzzy logic, optimization, and GIS for ecological impact assessments,” *Environmental Management*, vol. 30, no. 3, pp. 418–433, 2002.
- [3] K. Ranjit Upadhyay and S. R. K. Iyengar, “Stability and complexity in ecological systems,” *Chaos, Solitons and Fractals*, vol. 11, pp. 533–542, 2000.
- [4] T. G. Sazykina, V. V. Alekseev, and A. I. Kryshev, “The self-organization of trophic structure in ecosystem models: the succession phenomena, trigger regimes and hysteresis,” *Ecological Modelling*, vol. 133, no. 1-2, pp. 83–94, 2000.
- [5] M. Straskraba, “Natural control mechanisms in models of aquatic ecosystems,” *Ecological Modelling*, vol. 140, pp. 195–205, 2001.
- [6] C. Bianciardi and S. Ulgiati, “Modelling entropy and exergy changes during a fluid self-organization process,” *Ecological Modelling*, vol. 110, no. 3, pp. 255–267, 1998.
- [7] W. Cifang and C. Meiqiu, “Complexity of land ecosystem,” *Chinese Journal of Applied Ecology*, vol. 13, no. 6, pp. 753–756, 2002.
- [8] L. Zhao, “Nonlinear complex dynamics of carbon emission reduction cournot game with bounded rationality,” *Complexity*, vol. 2017, Article ID 8301630, pp. 1–7, 2017.
- [9] G. Fang, L. Tian, M. Fu, and M. Sun, “The impacts of carbon tax on energy intensity and economic growth - A dynamic evolution analysis on the case of China,” *Applied Energy*, vol. 110, pp. 17–28, 2013.
- [10] G. Fang, L. Tian, M. Fu et al., “The effect of energy construction adjustment on the dynamical evolution of energy-saving and emission-reduction system in China,” *Applied Energy*, vol. 196, pp. 180–189, 2017.
- [11] G. Fang, L. Tian, M. Fu, and M. Sun, “Government control or low carbon lifestyle? - Analysis and application of a novel selective-constrained energy-saving and emission-reduction dynamic evolution system,” *Energy Policy*, vol. 68, pp. 498–507, 2014.
- [12] G. C. Fang, L. X. Tian, M. Sun, and M. Fu, “Analysis and application of a novel three-dimensional energy-saving and emission reduction dynamic evolution system,” *Energy*, vol. 40, pp. 291–299, 2012.
- [13] C. Q. Yuan and S. F. Liu, “Research on energy-saving effect of technological progress based on Cobb-Douglas production function,” *Energy Policy*, vol. 37, no. 8, pp. 2842–2846, 2009.
- [14] R. Y. Long and L. Yu, “Study on regulation design about energy-saving and emission-reduction based on game theory,” *Procedia Earth and Planetary Science*, vol. 1, no. 1, pp. 1641–1646, 2009.
- [15] I. Musu and M. Lines, “Endogenous growth and environmental preservation,” in *Environmental Economics: Proceeding of European Economic Association at Oxford*, G. Boero and A. Silberston, Eds., St Martins Press, London, 1995.
- [16] D. Fullerton and S. Kim, “Environmental investment and policy with distortional taxes and endogenous growth,” *NBER Working Paper*, 2006.
- [17] J.-H. Chen, C.-C. Lai, and J.-Y. Shieh, “Anticipated environmental policy and transitional dynamics in an endogenous growth model,” *Environmental and Resource Economics*, vol. 25, no. 2, pp. 233–254, 2003.

- [18] X. X. Liao, *The Stability of Power System Theory and Application*, National Defence Industry Press, Beijing, China, pp. 252–268, 2000.
- [19] C. Xiaojing and R. Baoping, “The fluctuation and regional difference of quality of economic growth in China,” *Economic Research Journal*, vol. 4, pp. 26–40, 2011 (Chinese).

## Research Article

# Cascading Failures Analysis Considering Extreme Virus Propagation of Cyber-Physical Systems in Smart Grids

Tao Wang <sup>1,2</sup>, Xiaoguang Wei <sup>3</sup>, Tao Huang <sup>1,4</sup>, Jun Wang,<sup>1,2</sup> Luis Valencia-Cabrera,<sup>5</sup> Zhennan Fan <sup>1,2</sup> and Mario J. Pérez-Jiménez<sup>5</sup>

<sup>1</sup>School of Electrical Engineering and Electronic Information, Xihua University, China

<sup>2</sup>Key Laboratory of Fluid and Power Machinery, Ministry of Education, Xihua University, Chengdu 610039, China

<sup>3</sup>School of Electrical Engineering, Southwest Jiaotong University, China

<sup>4</sup>Department of Energy, Politecnico di Torino, Italy

<sup>5</sup>Research Group on Natural Computing, Department of Computer Science and Artificial Intelligence, University of Seville, Spain

Correspondence should be addressed to Xiaoguang Wei; [wei\\_xiaoguang@126.com](mailto:wei_xiaoguang@126.com) and Tao Huang; [tao.huang@polito.it](mailto:tao.huang@polito.it)

Received 3 November 2018; Revised 14 January 2019; Accepted 18 February 2019; Published 13 March 2019

Guest Editor: Riccardo Patriarca

Copyright © 2019 Tao Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Communication networks as smart infrastructure systems play an important role in smart grids to monitor, control, and manage the operation of electrical networks. However, due to the interdependencies between communication networks and electrical networks, once communication networks fail (or are attacked), the faults can be easily propagated to electrical networks which even lead to cascading blackout; therefore it is crucial to investigate the impacts of failures of communication networks on the operation of electrical networks. This paper focuses on cascading failures in interdependent systems from the perspective of cyber-physical security. In the interdependent fault propagation model, the complex network-based virus propagation model is used to describe virus infection in the scale-free and small-world topologically structured communication networks. Meanwhile, in the electrical network, dynamic power flow is employed to reproduce the behaviors of the electrical networks after a fault. In addition, two time windows, i.e., the virus infection cycle and the tripping time of overloaded branches, are considered to analyze the fault characteristics of both electrical branches and communication nodes along time under virus propagation. The proposed model is applied to the IEEE 118-bus system and the French grid coupled with different communication network structures. The results show that the scale-free communication network is more vulnerable to virus propagation in smart cyber-physical grids.

## 1. Introduction

The smart grid, as a modern electrical network (EN) infrastructure, can enhance the efficiency, reliability, and security of traditional ENs based on the advancement of cyber-physical systems [1–3]. In a smart grid, the monitoring, control, and management of the EN depend closely on the smart information and communication (cyber) network [4–6], which works such that the EN ensures not only its own secure operation but also reliable operation of the entire communication network. Meanwhile, when the EN fails (especially, through cascading failure), fault cross-propagation between the electrical and communication networks (ECNs), called

interdependent network, occurs, which increases the complexity of fault propagation owing to interactions between the ECNs. For example, the Italian blackout of 2003 was triggered by effects of the ECN [7]. Therefore, exploring the propagation mechanism of interactive cascading failures [8] in an interdependent network has been receiving increasing attention.

To date, the connection between the different coupling modes between ECNs and the robustness/vulnerability of interdependent networks has been investigated widely [9–12]. Studies have demonstrated that the different types of links between ECNs greatly impact the robustness of the network. For example, [10] reveals the double-network link allocation

strategy is superior to single-network link allocation strategy. Therefore, reasonably allocating the interconnecting links between ECNs is vital for improving the robustness of interdependent networks. Accordingly, a few models (e.g., Petri nets) have been introduced to reveal the mechanism of interactions leading to catastrophic blackouts [13]. However, these works have been done merely from the perspective of the structure of the coupling of the ECNs.

Meanwhile, the physical and operational characteristics considering the interactions have been focused on as well. In [14], the impact of communication network vulnerability on power system operation was assessed considering both latency and communication interruptions. The data exchange model is introduced for modeling cascading failures in interdependent networks [15]. Reference [16] proposes a simulation platform to analyze the ECN vulnerability by considering the control strategy of power balance. Similarly, other simulations [17, 18] have been proposed to analyze the fault mechanism considering interactions between ECNs. Although these studies have included the interactions between the communication network and EN, their focus is only to study how to set up the simulation platform by considering both the communication network and the EN. Moreover, the operational characteristics are only studied from a steady state point of view.

In addition, as communication networks become increasingly smart and as smart grids are increasingly accessed using the Internet owing to advancement of the energy Internet [19, 20], cyber threats (e.g., virus propagation; hacking attacks) leading to interactive cascading failures should be focused on [21, 22]. For instance, the 2015 blackout in the Ukraine was a typical coordinated cyber-attack in which malicious code was employed to tamper with data and control the server of the monitoring system [23]. In the face of potential threats, measures to enhance information security of communication networks [24] and a few robust and efficient cyber infrastructures [25] have been proposed. For example, in Qinghai, China, to prevent viral infection of networks, an antivirus system was installed to manage the power dispatching data network. This system successfully detected and neutralized 3384 viruses in 2010-2012 [26]. It is manifest that there is a need to consider the virus propagation in the communication network although it is a low-probability event, as such an event can cause immense harm to the ECN owing to the subtlety of the virus and the high speed at which the network is infected by using advanced attack methods.

With this background, in our paper, we propose a framework to analyze the performance of the ECN by considering the interactions between two types of propagations: fault propagation in the EN and virus propagation in the communication network.

In the EN, cascading failures have been analyzed from the perspective of the overloaded mechanism [15, 27–29]. When a line fails, the power transmitted over the line will be redistributed in the network, and thus, a fault may cause increased flow in other branches and even overload them, leading to the fault propagation. However, those analyses are generally performed by through a steady state fashion. In

our study, we improved the dynamic power flow method to redistribute loads and adjust unbalanced power in the network during fault propagation by introducing the primary frequency regulation and the equations of rotors of generators. Notably, we only consider high-voltage transmission networks as the study objects.

For the virus propagation, virus spread models [30, 31] with time delay have been developed based on the complex network theory (CNT) from the perspective of the topological structure of the communication network. Generally, the communication network mainly has two topical topological structures: scale-free and small-world networks. This paper focuses on investigating the impacts of two types of networks on electrical networks during virus propagation. In other words, we analyze which type of communication network is more vulnerable from the network-wise perspective, i.e., once viruses are propagated in a communication network, which type of communication network can cause more damage to the electrical network. Meanwhile, we further analyze that in a communication network, the vulnerability of communication nodes with different degree is revealed by investigating the number of fault branches and blackout level of coupled electrical networks.

In addition, it should be noted that the most modern malware, surely most malware is used in known attacks to power grids and industrial controls, limits its own effectiveness by prematurely destroying/disabling nodes and is not self-replicating. Furthermore, currently in a real-world communication network vertexes will not be homogeneous; thus they will not support most of self-propagating malicious code. However, in order to consider a low-probability but high consequence scenario, in this paper, we assume a random constant spread malicious code (called “virus” in this paper) with the following features to investigate the impact of extreme case of self-propagating virus on the power system from the network-wise perspective:

- (1) the virus can block the communication between infected vertexes and the control center;
- (2) the virus can self-propagate among homogeneous vertexes;
- (3) a few infectious vertexes can be cured with the probability owing to the strengthening of security measures;
- (4) the differences of security level of each node are not considered.

The remainder of this paper is organized as follows. Section 2 describes the interactions between the ECNs in the coupling relationships and topological structures. In Section 3, the virus propagation models with time delay and information exchange model in the communication network are introduced. The dynamic power flow method and the overload mechanism are established in Section 4. The cascading failures model considering the interactions and the corresponding simulation analysis are described in Sections 5 and 6, respectively. In Section 7, we further discuss the contribution of this paper and the external validity of the modeling. Finally, conclusions are given with possible future work in Section 8.

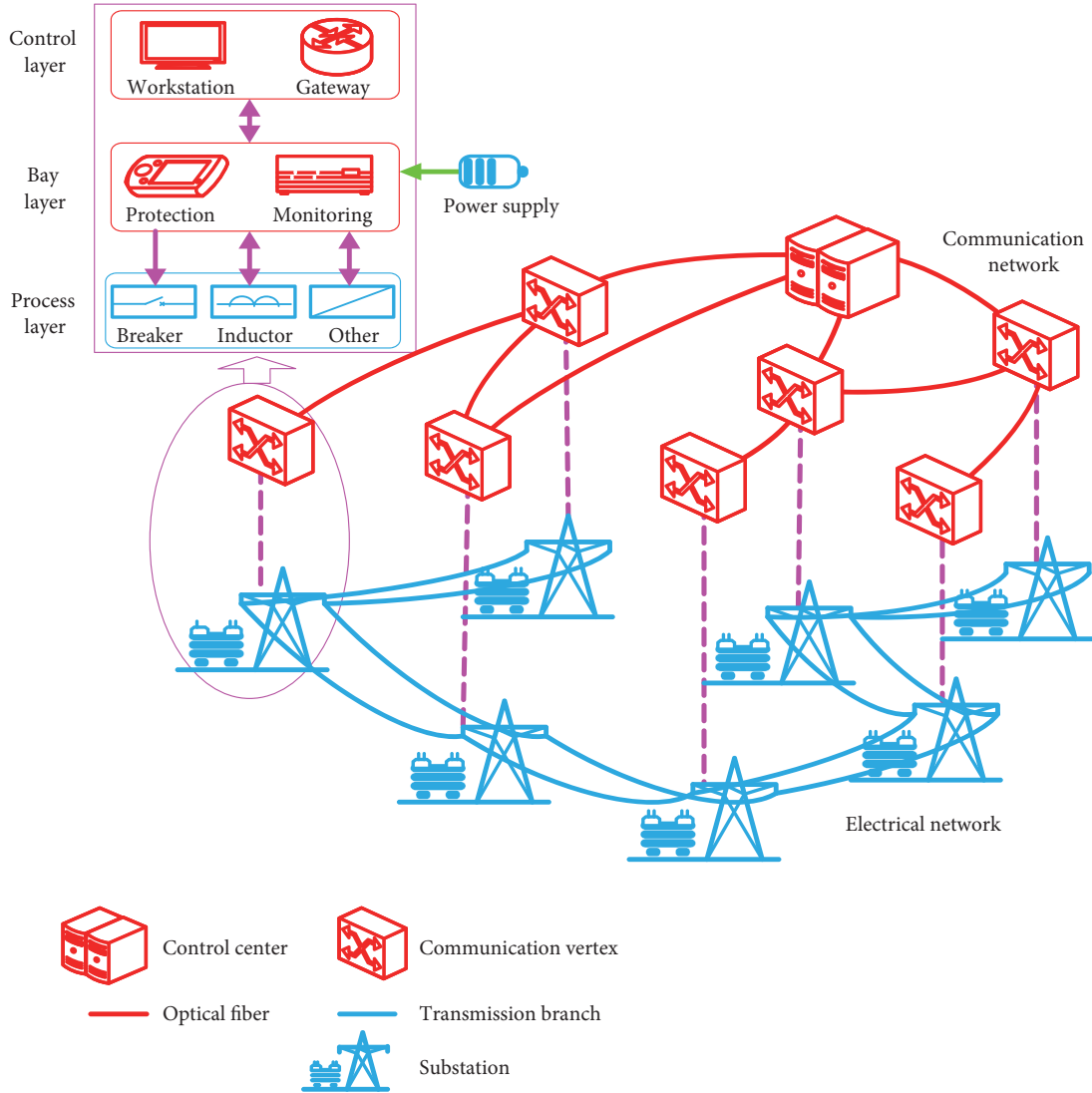


FIGURE 1: Diagram of ECN in smart grids.

## 2. Interaction between Electrical and Communication Networks in Smart Grid

### 2.1. Interdependent ECN in Smart Grids

*ECN Spatial Model.* The ECN in a smart grid is shown in Figure 1. The communication network, a hierarchical structure, is composed of optical fibers and synchronous digital hierarchies (SDHs), including control centers and communication vertexes [14, 32]. Generally, the substations (generators and loads) in the electrical network have the corresponding communication vertexes. The coupling between the substations and the communication vertexes is modeled by a smart communication module comprising three layers: process, bay, and control layers. Among the three layers, the bay layer is mainly responsible for accepting commands from the control layer to protect and monitor the electric network and realizing real-time interaction of information between the control and the process layers. The control

layer is responsible for sending real-time messages from the electrical network to the control center as well as for accepting commands from the control center through the communication network.

*ECN Operation Model.* We analyze the ECN operation model from the perspective of energy flow. In an ECN, there are two types of energy flows: power flow and communication flow, as shown in Figure 2. In an EN, power flow changes with time via buses or lines. Conversely, because the communication vertexes transmit and receive messages at regular intervals, the communication flow is transmitted based on discrete time.

*Communication Topology between Vertexes and Branches.* Because there is a consistent one-to-one match between each communication vertex and each bus node, each transmission branch  $B_i$  has two related communication vertexes  $V_{m1}$  and  $V_{m2}$  during normal operation. In this paper, only  $V_{m1}$

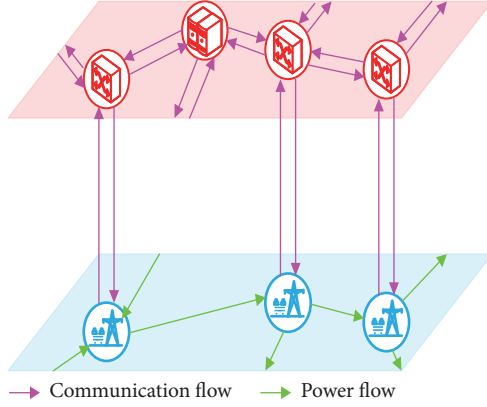


FIGURE 2: Information flows of electrical and communication networks in smart grids.

is considered for exchanging information packets with the control center [15].

## 2.2. Topological Structures of ECN

*Electrical and Communication Network as Graphs.* For simplifying analysis of the topological structures, we abstracted the ECN as graphs. The EN can be considered as a complex network with nodes and links. The buses, including generators, loads, and substations, can be viewed as nodes while transmission lines can be viewed as branches; therefore, the electrical network is represented as the graph  $\mathbb{G}_E = (\mathbb{N}, \mathbb{B})$ . The adjacent matrix  $\mathbf{G}_E = (a_{ir})_{M_N \times M_N}$  is employed to define  $\mathbb{G}_E$  as follows:

$$a_{ir} = \begin{cases} 1 & \exists N_i N_r = B_j \\ 0 & \neg \exists N_i N_r = B_j \end{cases} \quad (1)$$

where  $N_i N_r = B_j$  represents that there is a branch  $B_j$  between nodes  $N_i$  and  $N_r$ .

Similarly, the optical fibers and SDHs of the communication network can be considered as edges and vertexes, respectively; therefore, the communication network is represented as the graph  $\mathbb{G}_C = (\mathbb{V}, \mathbb{E})$ . The adjacent matrix  $\mathbf{G}_C = (a'_{mv})_{M_V \times M_V}$  is employed to define  $\mathbb{G}_C$  as follows:

$$a'_{mv} = \begin{cases} 1 & \exists V_m V_v = E_a \\ 0 & \neg \exists V_m V_v = E_a \end{cases} \quad (2)$$

where  $V_m V_v = E_a$  represents that there is an edge  $E_a$  between vertices  $V_m$  and  $V_v$ .

Because the buses are coupled with the corresponding communication vertexes by the communication module, which is represented as  $\mathbb{L} = \{L_b \mid L_b = N_i V_m, i = m\}$ , the ECN can be developed as an interdependent graph  $\mathbb{G} = \mathbb{G}_E \cup \mathbb{G}_C = (\mathbb{N} \cup \mathbb{V}, \mathbb{B} \cup \mathbb{E} \cup \mathbb{L})$ .

*Topological Structure Analysis.* We analyze the structural characteristics from the perspective of the CNT. Existing literature

indicate that ENs have small-world networks [33–35], which demonstrates that ENs have a relatively small average shortest path but a very large cluster coefficient. Thus, the small-world electrical networks reveal that if a node (or branch) in the network fails, the adjacent and even nonadjacent nodes (or branches) could fail, leading to cascading failures. Meanwhile, ENs has scale-free characteristics, as determined by analyzing changes in the network structure and function when one or more nodes (or branches) are removed from the network, which shows the networks are highly vulnerable under deliberate attacks but robust under random attacks [28]. However, fault propagation mechanism of ENs studied from the perspective of pure topological structure is not comprehensive and should more focus on the physical and operational features.

In communication networks, generally, there are two types of topological networks: scale-free networks and small-world networks [15, 36, 37]. Communication networks with scale-free structures contain a few nodes with high degree, and they can be considered center nodes. Compared to the small-world networks, the distributions of degree of which are more uniform, scale-free networks have higher communication efficiency but are more vulnerable to deliberate attacks.

From the perspective of pure topological structures, compared to ENs, fault (or virus) propagation in communication networks is largely determined by its topological structure. That is, a fault node (or branch) only causes neighboring nodes to fail. Therefore, we employ the CNT to develop the virus propagation models (VPMs) based on the topological structures.

In summary, the EN and communication network in smart grids have two essential differences in terms of the interactions of cascading failures.

*Features 1.* From the perspective of time scales, the power flow is transmitted based on continuous time, while the communication flow is transmitted based on discrete time.

*Features 2.* From the perspective of topological structures, fault propagation in the communication networks depends more on the network structures. Compared with the communication networks, fault propagation in ENs depends more on physical and operational modes because ENs comply with operational rules, for example, Kirchhoff's law.

## 3. Virus Propagation and Information Exchange Models in Communication Networks

Before analyzing VPMs, we introduce the following topological concepts:

The degree  $k_{V_m}$  of  $V_m$  is the number of neighboring vertexes connected to  $V_m$ , as expressed by

$$k_{V_m} = \sum_{v=1}^{M_V} a'_{mv} \quad (3)$$



Degree distribution  $p(k)$  is the distribution function of the degrees. That is, when a vertex is randomly selected from the network, the probability that its degree is equal to  $k$  is  $p(k)$ .

*3.1. Virus Propagation Model Based on CNT.* According to *Feature 1*, virus propagation in the communication network depends on the network structure; therefore, we employ the SI [38–40] and SIR [41, 42] models based on CNT to simulate virus propagation. In the SI model, the vertexes of the communication network are divided into two groups: susceptible set  $\mathbb{S}$  and infectious set  $\mathbb{I}$ . In  $\mathbb{S}$ , the probability that a susceptible vertex contracts the virus from the infectious vertexes is  $\beta$ . Meanwhile, because the virus spends some time in destroying the functions of susceptible vertexes (e.g., tampering with data or instructions), the susceptible vertexes take time to get infected. Therefore, we introduce time delay (virus infection cycle) to develop the SI model as follows:

$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta I(t - \tau_1) S(t - \tau_1) \\ \frac{dI(t)}{dt} &= \beta I(t - \tau_1) S(t - \tau_1)\end{aligned}\quad (4)$$

On the basis of the SI model, the SIR model considers that a few infectious vertexes can be cured with the probability owing to the strengthening of related antivirus measures (e.g., formatting operation). Thus, the infectious vertexes can obtain immunity in a certain virus removal cycle. Therefore, the vertexes add a group called removed set  $\mathbb{R}$ . The SIR model with time delay is given as

$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta I(t - \tau_1) S(t - \tau_1) \\ \frac{dI(t)}{dt} &= \beta I(t - \tau_1) S(t - \tau_1) - \alpha I(t - \tau_2) \\ \frac{dR(t)}{dt} &= \alpha I(t - \tau_2)\end{aligned}\quad (5)$$

Because an infectious vertex only transmits the virus to its neighboring vertexes, the topological structure of the network greatly influences virus propagation. When the communication network is a small-world network, which can be regarded as a uniform network owing to the relatively uniform distribution of degree [37], the degree  $k_{V_m}$  of  $V_m$  is approximately equal to  $\langle k \rangle$ , and the SIR model can be presented as follows:

$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta \langle k \rangle I(t - \tau_1) S(t - \tau_1) \\ \frac{dI(t)}{dt} &= \beta \langle k \rangle I(t - \tau_1) S(t - \tau_1) - \alpha I(t - \tau_2) \\ \frac{dR(t)}{dt} &= \alpha I(t - \tau_2)\end{aligned}\quad (6)$$

$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta k S_k(t - \tau_1) \frac{\sum_k k p(k) I_k(t - \tau_1)}{\langle k \rangle} \\ \frac{dI(t)}{dt} &= \beta k S_k(t - \tau_1) \frac{\sum_k k p(k) I_k(t - \tau_1)}{\langle k \rangle} \\ &\quad - \alpha I(t - \tau_2) \\ \frac{dR(t)}{dt} &= \alpha I(t - \tau_2)\end{aligned}\quad (7)$$

When the communication network is a scale-free network, the vertexes have different damage levels from the perspective of virus propagation because the distribution of degree follows a power law. That is, the greater the  $k_{V_m}$  of  $V_m$ , the more serious it is for the  $V_m$  to spread or contract the virus to more vertexes. The SIR model can be expressed as (7) in terms of the vertex degree [43, 44].

Generally, in the SI and SIR models, virus propagation is faster in scale-free networks owing to the power law distribution. In addition, by comparing SI and SIR models, once virus propagation occurs in the communication network, we can investigate whether the related antivirus measures with time delay can play an important role to prevent the fault from spreading across the EN.

*3.2. Information Exchange Model in Communication Network.* In the smart grid, the communication vertexes send operational data (parameters) associated with branches to the control center and receive commands from the control center step-by-step through the communication network in the form of information packets. At every step, the same information packets can be received and sent only by each communication vertex. Before constructing the information exchange model, three simplifications are made as follows.

(1) The communication blocks of vertexes (or edges) are not considered in process of the information transfer when the vertexes work orderly. That is, the capacity of the vertexes is adequate to exchange/handle the information packets.

(2) Because the vertexes send and receive information packets at intervals of 0.833 ms [45], the time required for information exchange between the vertexes and control center can be ignored because it is very small compared to the time required for fault propagation in the EN.

(3) Because we focus on the interactions between the ECNs, the methods of gathering and dealing with the information (such as the measuring units, transmission channels and protocols, encryption and decryption algorithms, etc.) are not considered.

Based on the above simplifications, the information exchange model is constructed based on the structure of the communication network.

*Communication Rules between Target Vertex and Control Center.* The vertexes abide by the rule of first-in-first-out to send out information packets to avoid exchange of the information packets to be in the same edge. At first, the target vertex produces information packets. Then, the information packets are sent to all its neighbor vertexes. If the control center is one of the neighbor vertexes, the information

transfer ends; otherwise, all neighbor vertexes, acting as target vertexes, continue to send the information packets to their corresponding neighbor vertexes until the information packets are sent to the control center. In the above process, the information packets are transmitted successfully between the target vertex and control center if a path exists between them in the communication network.

#### 4. Dynamic Power Flow and Overload Mechanism Models in Electrical Networks

*4.1. Dynamic Power Flow in Electrical Networks.* To redistribute power flow during disturbances, we employ primary frequency regulation [46, 47] and rotor equation [48] to model the dynamic power flow method.

*System Frequency Characteristics.* We employ primary frequency regulation to adjust the power flow. The characteristics of load and generation frequency are given by (8) and (9), respectively.

$$\Delta P_X = K_X \Delta f \quad (8)$$

$$\Delta P_W = K_W \Delta f \quad (9)$$

$K_X$  and  $K_W$  are calculated as follows:

$$K_X = \frac{\sum_{q=1}^{N_X} (K_{Rq} \cdot P_{Rrq})}{\sum_{q=1}^{N_X} P_{Rrq}} \quad (10)$$

$$K_W = \frac{\sum_{c=1}^{N_W} (K_{Wc} \cdot P_{Wrc})}{\sum_{c=1}^{N_W} P_{Wrc}} \quad (11)$$

*Unbalanced Power Redistribution.* To redistribute the unbalanced power  $P_{un}$  due to disturbances of the system, we first calculate the change in system frequency by using the primary frequency regulation and the rotor equation.

$$T_J \frac{d\Delta\omega}{dt} = P_{un} - K_E \cdot \Delta\omega(t) \quad (12)$$

In (12),  $P_{un}$  is calculated as follows:

$$P_{un} = \sum_{c=1}^{M_W} P_{Wc} - \sum_{c=1}^{M_X} P_{Xq} \quad (13)$$

When  $P_{un} < 0$ , we consider the characteristics of load and the generation frequency to adjust the system frequency:

$$K_E = K_X + K_W \quad (14)$$

When  $P_{un} > 0$ , we consider only the generation frequency characteristic:

$$K_E = K_W \quad (15)$$

By using (8)-(15), the changes in every generator and load can be expressed as follows:

$$\Delta P_{Xq} = K_{Xq} \cdot \Delta\omega(t) \quad (16)$$

$$\Delta P_{Wc} = K_{Wc} \cdot \Delta\omega(t) \quad (17)$$

Then, we employ the P-Q power flow to calculate the power flows of each bus ( $q = c = i$ ) as follows:

$$\begin{aligned} P_i &= -\Delta P_{Wq} + \Delta P_{Xc} \\ &+ \nu_i \sum_{u=1}^{M_N} \nu_u (G_{iu} \cos \theta_{iu} + B_{iu} \sin \theta_{iu}) \\ &= (-K_{Wq} + K_{Xc}) \Delta\omega \end{aligned} \quad (18)$$

$$\begin{aligned} &+ \nu_i \sum_{u=1}^{M_N} \nu_u (G_{iu} \cos \theta_{iu} + B_{iu} \sin \theta_{iu}) \\ Q_i &= \nu_i \sum_{u=1}^{M_N} \nu_u (G_{iu} \sin \theta_{iu} - B_{iu} \cos \theta_{iu}) \end{aligned} \quad (19)$$

*Dynamic Power Flow Method.* When branches fault during fault propagation in the network, the dynamic power flow can be calculated in Algorithm 1.

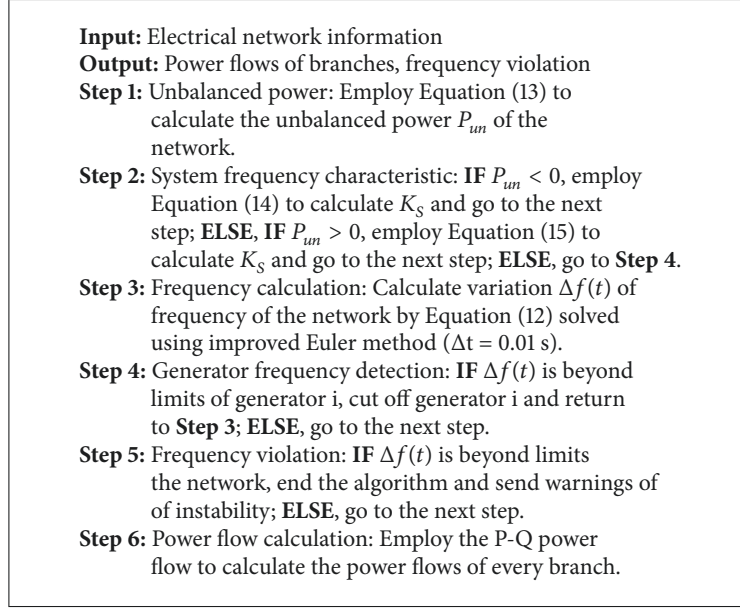
*4.2. Overload Mechanism of Electrical Networks.* In this paper, cascading failures in the EN are analyzed from the perspective of the overload mechanism. When one or more lines are cut off, the other lines are overloaded owing to the redistribution of power flow in the EN [27, 28]. When a branch is overloaded, the larger the power flow over the branch, the shorter is the operational time for which the branch is permitted to continue working [15]. As most of other studies [15, 49], in this paper, we assume that during the fault propagation, the control center tries to maintain the secure operation of the EN and lower the load shedding amount, thus the control strategy includes which branches to trip, how to adjust the generators output, as well as to shed which load of how many percentages, etc. Therefore, for some of the branches, the tripping command has to come from the control center. Of course, for some of the faulted lines, the tripping signal should be issued by a local protection unit. However, to simplify the process, we simply assume that the tripping command comes from the control center. Thus, under this simplification, when the corresponding communication vertexes send information about overloading to the control center via the communication network, the control center must quickly send trip commands to the target vertexes. We employ the inverse-time overcorrect protection scheme [15, 49] to calculate the overloaded operational time.

$$t_{B_j} = \frac{\kappa}{\left| \frac{I_{B_j}}{\bar{I}_{B_j}} \right|^\sigma - 1} \quad (20)$$

If the data exchange between the target vertex and control center to trip the branch  $B_j$  is completed within  $t_{B_j}$ , the control is successful; otherwise, the control is unsuccessful.

### 5. Interactive Cascading Failure Model

A diagram of cascading failures considering the interactions between the ECN is shown in Figure 3. During the cascading



ALGORITHM 1: Dynamic power flow method.

failures in the EN, if branch  $B_j$  is overloaded, according to the control strategy,  $B_j$  generates fault information packets, and the corresponding communication vertexes then send these packets to the control center via the communication network. Thereafter, the control center sends the commands back to  $B_j$  within  $t_{B_j}$ .

In addition, when the virus spreads through the communication network, the infectious vertex  $V_i$  will lose the function of information exchange and connectivity  $a'_{mi}$  with its neighbor vertexes, according to

$$a'_{mi} = 0 \quad (m = 1, 2, \dots, M_V) \quad (21)$$

Meanwhile, the virus will cause branches to trip or lead to an outage directly or indirectly because the infectious vertexes lose the function of information exchange. Accordingly, there are four types of fault branches.

*Type 1.* The branch is forced to trip because the corresponding communications get infected, also called forced outage branches  $\mathbb{B}_{FO}$ .

*Type 2.* The branch is tripped properly because the control center successfully sends commands to the corresponding vertexes based on the received overload information and control strategy with  $t_{B_j}$ , also called overload tripping branches  $\mathbb{B}_{OT}$ .

*Type 3.* The branch is damaged irreparably owing to control failures via the communication networks, leading to overload operational time exceeding  $t_{B_j}$ , also called irreparable fault branches  $\mathbb{B}_{IF}$ .

*Type 4.* The branch undergoes forced outage owing to network splitting, also called network splitting branches  $\mathbb{B}_{NS}$ .

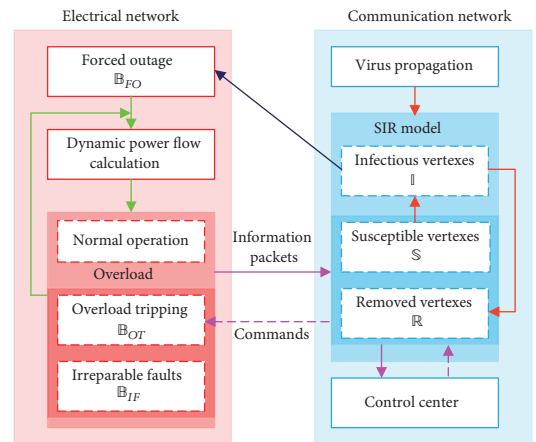


FIGURE 3: Cascading failures considering interactions between ECNs.

Based on the above analysis, cascading failure considering the interactions between the ECNs is modeled in Algorithm 2.

## 6. Case Study

The proposed model was applied to the IEEE 118-bus system and the French grid [50]. The small-world and scale-free networks were adopted to represent the respective communication networks. The computational work was performed in MATLAB running on a laptop. The laptop (Compaq, v3646TU) was equipped with an Intel® Core™ 2 Duo CPU T7250@2.00 GHz, 2.00 GB RAM, and 64-bit Windows 7 operating system.

**Input:** Electrical network information and parameters, communication network information and parameters,  $\eta, \mathbb{I}, \beta, \alpha, \tau_1$

**Output:**  $\mathcal{G}, \mathbb{B}_{FO}, \mathbb{B}_{OT}, \mathbb{B}_{IF}, \mathbb{B}_{NS}$

**Step 1:** Initialization:  $t = 0$  s,  $\mathcal{G} = \emptyset, \mathbb{B}_{FO} = \emptyset, \mathbb{B}_{OT} = \emptyset, \mathbb{B}_{IF} = \emptyset, \mathbb{B}_{NS} = \emptyset, \mathbb{B}_O = \emptyset, \mathbb{S} = \emptyset, \mathbb{R} = \emptyset, \mathbb{R}' = \emptyset$  and  $\mathbb{I}' = \emptyset$ .

**Step 2:** WHILE  $t < \eta$

**Electrical network:**

**Step 3:** Forced branch outage: **IF** the corresponding vertexes coupled with the branch  $B_j$  ( $j = 1, 2, \dots, M_B$ ) are in  $\mathbb{I}$ , add the branch  $B_j$  to  $\mathbb{B}_{FO}$  and trip it.

**Step 4:** Overloaded branch tripping: **IF** the overload operation time of  $B_x$  ( $B_x \in \mathbb{B}_O$ ) ( $x = 1, 2, \dots, M_O$ ) in  $t_{Bx} = t$ , and the fault packages between the corresponding vertexes and control center are exchanged successfully, add  $B_x$  to  $\mathbb{B}_{OT}$ , and delete  $B_x$  from  $\mathbb{B}_O$ ; **ELSE, IF**  $t_{Bx} > t$ , add  $B_x$  to  $\mathbb{B}_{IF}$ , and delete  $B_x$  from  $\mathbb{B}_O$

**Step 5:** Network splitting: Detect and split the electrical network. **IF** there exists the forced outage branch  $B_z$  ( $z = 1, 2, \dots, M_{NS}$ ) due to the splitting, add  $B_z$  to  $\mathbb{B}_{NS}$ .

**Step 6:** Network operational status: Calculate the power flow over  $B_j$  ( $j = 1, 2, \dots, M_B$ ). **IF**  $B_j$  is overloaded, calculate  $t_{Bj}$  of  $B_j$  by Equation (19), and add  $B_j$  to  $\mathbb{B}_O$ .

**Communication network:**

**Step 8:** Infectious vertexes detection: **IF** the infection time of the candidate vertex  $V_k$  ( $k = 1, 2, \dots, M_{I'}$ ) in  $\mathbb{I}'$  is equal to  $t$ , add  $V_k$  to  $\mathbb{I}$  and delete it from  $\mathbb{I}'$ .

**Step 10:** Removal of vertex detection: **IF** the removal time of the candidate vertex  $V_w$  ( $w = 1, 2, \dots, M_{R'}$ ) in  $\mathbb{R}'$  is equal to  $t$ , add  $V_w$  to  $\mathbb{R}$ , and delete it from  $\mathbb{R}'$ .

**Step 7:** Virus propagation: The susceptible vertex  $V_g$  ( $g = 1, 2, \dots, M_S$ ) contracts the virus with probability  $\beta$  according to Equation (6). **IF**  $V_g$  gets infected, delete  $V_g$  from  $\mathbb{S}$ , add  $V_g$  to  $\mathbb{I}'$ , and label its infectious time  $t + \tau_1$ .

**Step 9:** Vertex immunization: The infected vertex  $V_l$  ( $l = 1, 2, \dots, M_I$ ) is immunized with the probability  $\alpha$  according to Equation (6). **IF**  $V_l$  obtains immunity, delete  $V_l$  from  $\mathbb{I}$ , add  $V_l$  to  $\mathbb{R}'$ , and label its immunity time  $t + \tau_2$ .

**Step 11:**  $t = t + \Delta t$ ; **END WHILE.**

ALGORITHM 2: Cascading failure model considering interactions between ECNs (SIR model as an example).

**6.1. IEEE 118-Bus System.** We randomly chose the communication vertexes as the initial infectious vertexes and then performed 1000 cascading events to investigate load shedding and number of instances of the four types of fault branches ( $M_{IF}, M_{OT}, M_{FO}$ , and  $M_{NS}$ ) based on the different VPMs and topological structures of the communication network with the parameters  $\Delta t = 0.01$  s,  $\tau_1 = \tau_2 = 5$  s,  $\alpha = \beta = 0.3$ ,  $\kappa = 7$  and  $\sigma = 1.5$ . The averaged results are shown in Figure 4 and listed in Table 1.

Figure 4 shows that the load shedding changes with the passage of time based on the different topological structures of the communication networks and the SI model. Owing to space limits, the SIR-model-based load shedding is not given herein. In Figure 4, the propagation time of interactive cascading failures depends on the virus propagation time  $\tau_1$ , and the load shedding is the maximum when the propagation

time is approximately 10s. Compared with the small-world communication network, the propagation time of interactive cascading failures is longer in the scale-free communication network. Moreover, coupling with the scale-free communication network, the fault branches of the EN result in irreparable faults ( $M_{IF} = 3.415$ ) with higher probability than that ( $M_{IF} = 0.245$ ) in the case of coupling with the small-world communication network, as summarized in Table 1. Therefore, the coupling of EN with the scale-free communication network is affected more severely as the propagation time increases because the connectivity between vertexes often depends on a few hub vertexes (i.e., high-degree vertexes), and the exchange of information packets becomes difficult, leading to network paralysis once a few hub vertexes are infected.

A comparison of the SI and the SIR models shows that different VPMs have very small impacts on fault propagation

TABLE I: Average numbers of four types of fault branches.

VPMs	Structures	$M_{FO}$	$M_{IF}$	$M_{OT}$	$M_{NS}$
SI	SW	148.164	0.245	14.589	9.1554
	SF	142.977	3.415	14.643	10.1554
SIR	SW	148.222	0.139	14.658	9.336
	SF	143.104	1.371	16.588	10.064

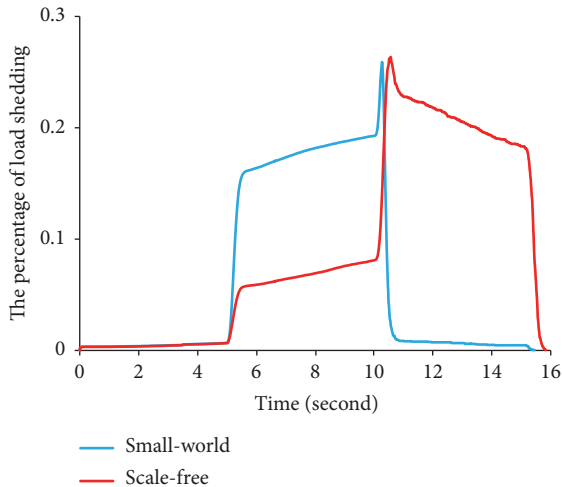


FIGURE 4: Load shedding of IEEE 118-bus system over time based on SI model.

in the EN, which indicates that once fault propagation occurs in the EN, the related antivirus measures with time delay can barely prevent the fault from spreading across the EN. However, the number of fault branches with irreparable faults can be reduced, especially in the case of coupling with the scale-free communication network. This is because the immune vertexes treated with the antivirus recover their function of data exchange, and a few overloaded branches can receive trip commands from the control center in a timely manner, thus avoiding irreparable faults.

Furthermore, we analyze the interactive cascading failures by selecting different initial infectious vertexes. We used the SI model as an example. Because the degree distributions of the small-world communication network are known, we take the scale-free network as the basis to select the high-degree (vertexes 115 and 116) and small-degree (vertexes 4 and 8) vertexes as the initial infectious vertexes. Figures 5 and 6 show the total and real-time load shedding changes with the passage of time for different virus propagation times  $\tau_1 = 2$  s, 5 s, and 8 s. The initial vertexes have small impacts on fault propagation in the EN owing to the known degree distribution. However, in case of the coupling of the EN with the scale-free communication network, the initial vertexes greatly impact fault propagation. Compared to the small-world communication network, when the initial vertexes are high-degree vertexes in the scale-free communication network, the propagation time is obviously shorter, which demonstrates the hub can rapidly spread the virus, leading to rapid collapse of the EN. By contrast, the propagation

time is longer when the initial vertexes are the low-degree vertexes, and when the load shedding peaks, as shown in Figure 6, the interactive cascading failures continue to spread, which indicates virus propagation times are longer than fault propagation times. That is, when fault propagation has stopped, virus propagation continues.

**6.2. French Grid.** A real French grid with 1951 nodes and 2956 branches was employed to simulate the interactive model. Owing to computational complexity, we only choose the high-degree vertexes as initial infectious vertexes considering the topological structures of the communication network with the parameters  $\Delta t = 0.01$  s,  $\tau_1 = \tau_2 = 2$  s,  $\alpha = \beta = 0.3$ ,  $\kappa = 7$ , and  $\sigma = 1.5$ . Figure 7 shows that the total and real-time load shedding of system changes with the passage of time. Compared to the small-world communication network, the propagation times are longer in the case of EN coupled with scale-free communication network, but the load shedding peaks at approximately 8.3 s under both topological structures.

Furthermore, we investigated the numbers of the three types of fault branches ( $M_{IF}$ ,  $M_{OT}$ , and  $M_{FO}$ ) at different moments, as shown in Figures 8(a) and 8(b). Between 6 s and 7 s, the fault propagation is at its height, which indicates that the numbers of infectious vertexes and forced outage branches  $\mathbb{B}_{FO}$  are the highest, leading to rapid collapse of the EN and surging load loss. In addition, in the fault propagation process in the EN, the fault branches with irreparable faults are not found when the EN is coupled with the small-world communication network. By contrast, there are many fault branches of this type at different moments in the case of EN coupled with the scale-free communication network. Therefore, when the communication network is scale-free, it is more vulnerable which cannot effectively resist the virus propagation leading to the more severe damage to the EN.

The conclusions obtained from these two cases are summed up in Table 2. In practice, when ENs are faced with hacker attacks, because the attackers find it relatively difficult to obtain complete information about the communication networks, such as topological structures, their attacks are random to some extent, which means scale-free communication networks are more appropriate for the ENs [10, 30, 31]. However, when ENs are faced with the threat of a cyber virus, the virus must be cleared promptly. Once the virus spreads, regardless of whether the initial infectious vertexes are selected randomly or deliberately, the infection will lead to a severe damage in the scale-free communication networks. Therefore, due to the propagation features of the scale-free communication networks, the simulation results show that software engineers should strengthen more the software

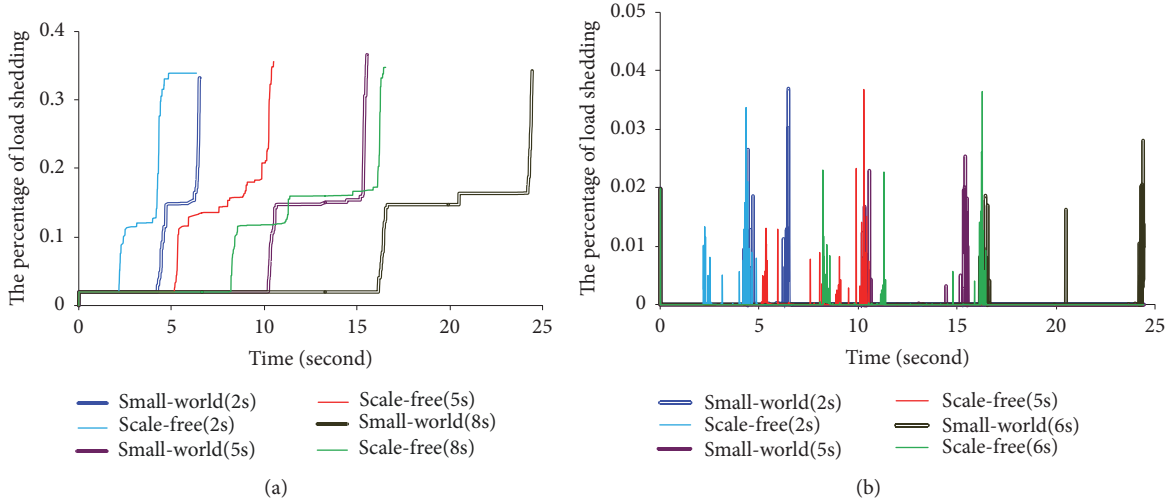


FIGURE 5: Load shedding of IEEE 118-bus system over time in the case of high-degree initial infectious vertices. (a) Total load shedding over time; (b) real-time load shedding.

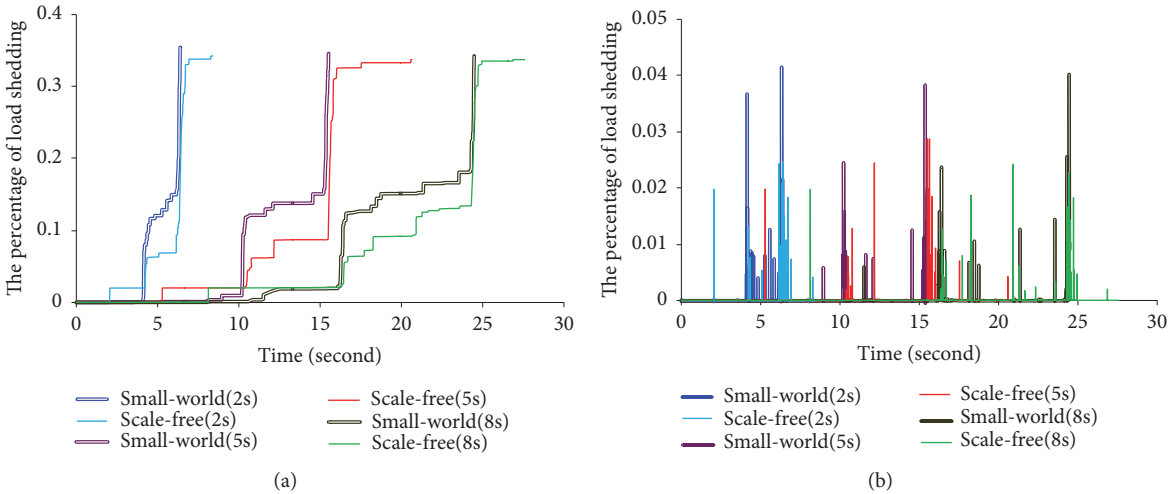


FIGURE 6: Load shedding of IEEE 118-bus system over time with low-degree initial infectious vertices. (a) Total load shedding over time; (b) real-time load shedding.

protections in the scale-free communication networks by means of more frequent update of the firewalls and antivirus software, strategies of automatic system restoration, etc.

## 7. Discussion

In this paper, we extend the state of the art for the study of the integrated communication network and electrical network. Compared with other literature, the main contributions of our paper are as follows:

We propose an interdependent fault propagation model which holistically considers the extreme virus propagation in the communication network to reveal the vulnerability of electrical network coupled with different communication network structures at the first time.

In the fault propagation model, to better reproduce the ex-post behavior of the electrical networks, we extended the

dynamic power flow by including the primary frequency regulation and the equations of rotors of generators.

To solve the issue of different time frames in the interdependent system, we adopt two time windows, i.e., the virus infection cycle of nodes and tripping time of overloaded branches during fault propagation to analyze the fault mechanism of both electrical branches and communication nodes along time.

It should be noted that even though the electrical network and communication network are both presented as graphs to conveniently describe their interdependent topological relationship in this paper, the modeling approach captured most of the relative features of the two networks.

For the electrical network, besides the commonly considered steady state physical and operational rules, we also adopt the rotor equation and system frequency to consider simple system dynamics in order to present the interactions between

TABLE 2: Comparison of propagation times between SW and SF networks.

Objects	Types of propagation	Initial factious vertexes	
		High-degree	Low-Degree
ECN	Interactive	SW>SF	SW<SF
Electrical network	Fault	SW=SF	SW=SF
Communication Network	Virus	SW>SF	SW<SF

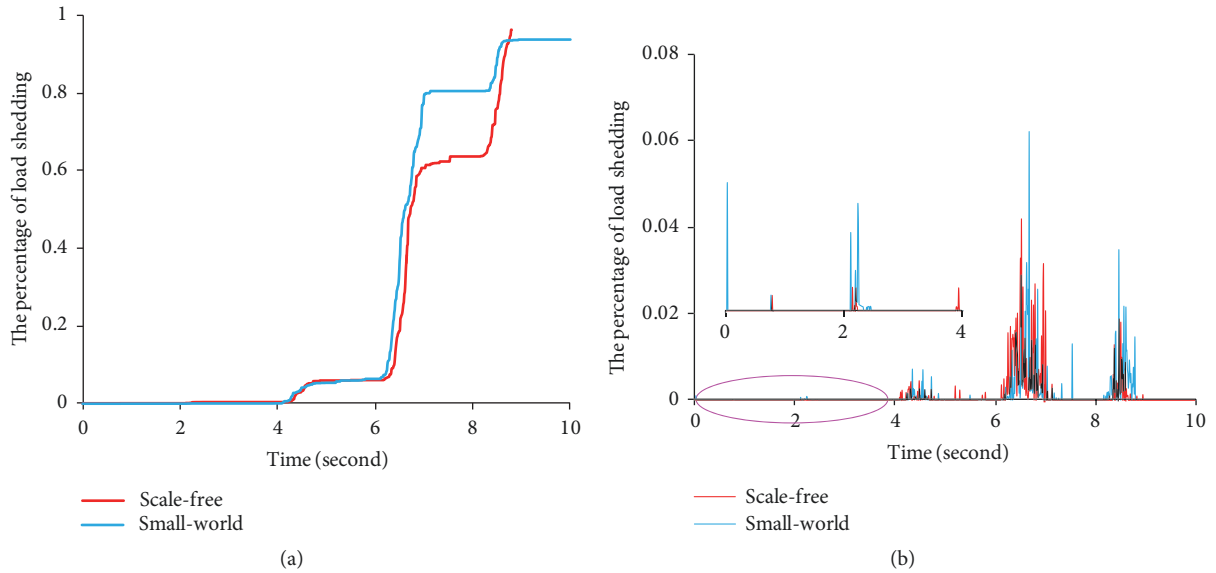


FIGURE 7: Load shedding of IEEE 118-bus system over time based on SIR model. (a) Total load shedding over time; (b) real-time load shedding.

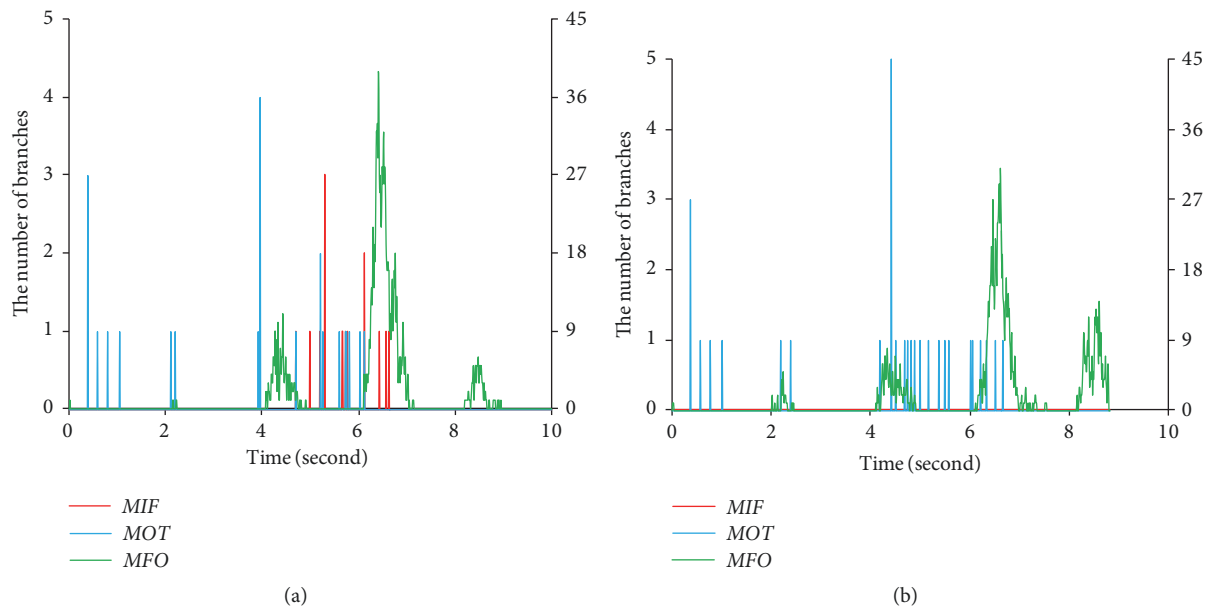


FIGURE 8: Numbers of different types of fault branches over time based on SIR model. (a) Scale-free network; (b) small-world network.

the electrical network and the communication network. As for the communication networks, we assume that the security level of each node is the same in terms of the infected rate. By contrast, in reality, the probability of the communication nodes got infected may vary for different nodes. However, the assumption made in our paper does not change the essence of the analysis and results in terms of evaluating which topology of communication network would have higher impacts on the electrical network during cyber-attacks.

## 8. Conclusions

The cyber-physical security of power systems is attracting increasing attention, especially after more and more evidences show that failures or attacks happening in the cyber system can greatly destroy the secure operation of power systems and bring tremendous consequences. To investigate the possible consequences, we propose an approximate interactive model to study cascading failures in ENs caused by virus in communication networks via two types of propagation. Our simulation on a standard study case, i.e., IEEE 118-bus system, and a realistic network, i.e., French grid, shows that the structure of the communication network has decisive impacts on the ECN in terms of the propagation time of cascading failures, loading shedding, number of faulted branches, etc. However, due to the simplification of the communication network and the virus propagation mechanism, the model can still be refined. In addition, the analysis is only focused on the overload of the system which may limit the results to part of behaviors of the EN.

Owing to the complexity of the propagation mechanism of interactive cascading failures, future work in this field will focus on considering more factors, such as data transmission delay, to simulate interactive cascading failures. Meanwhile, we also can investigate the impacts of differences of virus infection of nodes on interdependent fault propagation for electrical and communication networks. In addition, we can also analyze other aspects of the integrated CPS system, such as reliability, resilience, etc., under the virus propagation, to provide other dimensions for understanding the CPS.

## Nomenclature

ECN: Electrical and communication network  
 SDH: Synchronous digital hierarchy  
 CNT: Complex network theory  
 VPM: Virus propagation model  
 SM: Small-world  
 SF: Scale-free  
 EN: Electrical network.

Sets (Note That  $|\cdot|$  Represents the Dimension of a Set)

$\mathbb{N}$ : Set of nodes (i.e., buses) in an electrical network,  $\mathbb{N} = \{\dots, N_i, \dots\}$ ,  $|\mathbb{N}| = M_N$   
 $\mathbb{W}$ : Set of nodes with generators,  $\mathbb{W} = \{\dots, W_c, \dots\} \subseteq \mathbb{N}$ ,  $|\mathbb{W}| = M_W$   
 $\mathbb{X}$ : Set of nodes with loads,  $\mathbb{X} = \{\dots, X_q, \dots\} \subseteq \mathbb{N}$ ,  $|\mathbb{X}| = M_X$

$\mathbb{B}$ : Set of branches (i.e., lines) in an electrical network,  $\mathbb{B} = \{\dots, B_j, \dots\}$ ,  $N_i N_r = B_j$ ,  $|\mathbb{B}| = M_B$   
 $\mathbb{G}_E$ : Electrical network,  $\mathbb{G}_E = (\mathbb{N}, \mathbb{B})$   
 $\mathbb{V}$ : Set of vertexes (i.e., optical fibers) in a communication network,  $\mathbb{V} = \{\dots, V_m, \dots\}$ ,  $|\mathbb{V}| = M_V$   
 $\mathbb{E}$ : Set of edges (i.e., SDHs) in a communication network,  $\mathbb{E} = \{\dots, E_a, \dots\}$ ,  $V_m V_n = E_a$ ,  $|\mathbb{E}| = M_E$   
 $\mathbb{G}_C$ : Communication network,  $\mathbb{G}_C = (\mathbb{V}, \mathbb{E})$   
 $\mathbb{L}$ : Set of links which present the couples between electrical network and communication network,  $\mathbb{L} = \{\dots, L_b, \dots\}$ ,  $L_b = N_i V_m$ ,  $i = m$ ,  $|\mathbb{L}| = M_L$   
 $\mathbb{G}$ : Electrical and communication network,  $\mathbb{G} = \mathbb{G}_E \cup \mathbb{G}_C$   
 $\mathbb{S}$ : Set of susceptible vertexes,  $\mathbb{S} = \{\dots, V_g, \dots\} \subseteq \mathbb{V}$ ,  $|\mathbb{S}| = M_S$   
 $\mathbb{I}$ : Set of infectious vertexes,  $\mathbb{I} = \{\dots, V_l, \dots\} \subseteq \mathbb{V}$ ,  $|\mathbb{I}| = M_I$   
 $\mathbb{I}'$ : Set of candidate infectious vertexes,  $\mathbb{I}' = \{\dots, V_k, \dots\} \subseteq \mathbb{V}$ ,  $|\mathbb{I}'| = M_{I'}$   
 $\mathbb{R}$ : Set of removed vertexes,  $\mathbb{R} = \{\dots, V_h, \dots\} \subseteq \mathbb{V}$ ,  $|\mathbb{R}| = M_R$   
 $\mathbb{R}'$ : Set of candidate removed vertexes,  $\mathbb{R}' = \{\dots, V_w, \dots\} \subseteq \mathbb{V}$ ,  $|\mathbb{R}'| = M_{R'}$   
 $\mathbb{B}_{IF}$ : Set of branches with irreparable faults due to the control failures,  $\mathbb{B}_{IF} = \{\dots, B_e, \dots\} \subseteq \mathbb{B}$ ,  $|\mathbb{B}_{IF}| = M_{IF}$   
 $\mathbb{B}_{OT}$ : Set of fault branches with overload tripping due to the control successes,  $\mathbb{B}_{OT} = \{\dots, B_g, \dots\} \subseteq \mathbb{B}$ ,  $|\mathbb{B}_{OT}| = M_{OT}$   
 $\mathbb{B}_{FO}$ : Set of branches with forced outage due to the corresponding communication vertexes get infected,  $\mathbb{B}_{FO} = \{\dots, B_y, \dots\} \subseteq \mathbb{B}$ ,  $|\mathbb{B}_{FO}| = M_{FO}$   
 $\mathbb{B}_{NS}$ : Set of branches with forced outage due to the network splitting,  $\mathbb{B}_{NS} = \{\dots, B_z, \dots\} \subseteq \mathbb{B}$ ,  $|\mathbb{B}_{NS}| = M_{NS}$   
 $\mathbb{B}_O$ : Set of overloaded branches,  $\mathbb{B}_O = \{\dots, B_x, \dots\} \subseteq \mathbb{B}$ ,  $|\mathbb{B}_O| = M_O$ .

## Constants

$a_{ir}$ : The branch between  $N_i$  and  $N_r$   
 $a'_{mv}$ : The edge between  $V_m$  and  $V_n$   
 $P_{wrc}$ : Power rating of generator c  
 $P_{rrq}$ : Power rating of load q  
 $K_{wc}$ : cth generator unit power regulation  
 $K_w$ : Equivalent generator unit power regulation  
 $K_{xq}$ : qth load frequency regulation  
 $K_x$ : Equivalent load frequency regulation  
 $K_E$ : System unit power regulation  
 $T_j$ : Equivalent Inertia time constant  
 $T_{jc}$ : Inertia time constant of generator c



- $k_{vm}$ : Degree of vertex  $m$   
 $p(k)$ : Degree distribution  
 $\kappa$ : Proportional coefficient of inverse-time overcurrent protection  
 $\sigma$ : Power coefficient of inverse-time overcurrent protection  
 $\bar{I}_{B_j}$ : Current limit of branch  $B_j$ .

#### Variables

- $t$ : Time / clock  
 $t_{B_j}$ : Overloaded operational time of branch  $B_j$   
 $\eta$ : Simulation time  
 $\cdot(t)$ : Value / set of a variable at time  $t$   
 $\Delta t$ : Time step  
 $\cdot_k$ : Value / set of vertex(es) with  $k$  degrees  
 $P_{un}$ : System unbalanced power  
 $\Delta f$ : Frequency offset  
 $\Delta P_W$ : Changes of power of all generators  
 $\Delta P_{wc}$ : Changes of power of generator  $c$   
 $\Delta P_X$ : Changes of power of all loads  
 $\Delta P_{Xq}$ : Changes of power of load  $q$   
 $P_{Xq}$ : Power of load  $q$   
 $I_{B_j}$ : Current over branch  $B_j$   
 $P_i$ : Injection active power of node  $i$   
 $Q_i$ : Injection reactive power of node  $i$   
 $B_{iu}$ : Equivalent susceptance between nodes  $i$  and  $u$   
 $G_{iu}$ : Equivalent conductance between nodes  $i$  and  $u$   
 $\theta_{iu}$ : Voltage phase angle difference between nodes  $i$  and  $u$   
 $v_i$ : Voltage of node  $i$   
 $\Delta\omega$ : Changes of angular acceleration of equivalent generator  
 $\vartheta$ : The percentage of load shedding  
 $S$ : Percentage of susceptible vertexes,  $S = M_S/M_V$   
 $I$ : Percentage of infectious vertexes,  $I = M_I/M_V$   
 $R$ : Percentage of removed vertexes,  $R = M_R/M_V$   
 $\beta$ : Infection rate from susceptible vertex to infectious vertex  
 $\alpha$ : Recovery rate from infectious vertex to removed vertex  
 $\tau_1$ : Virus infection cycle  
 $\tau_2$ : Virus removal cycle.

#### Matrix

- $\mathbf{G}_E$ : Connectivity of the graph  $\mathbb{G}_E$ ,  $\mathbf{G}_E = (a_{ir})_{M_N \times M_N}$   
 $\mathbf{G}_C$ : Connectivity of the graph  $\mathbb{G}_C$ ,  $\mathbf{G}_C = (a'_{mv})_{M_V \times M_V}$ .

#### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

#### Conflicts of Interest

All authors declare that they have no conflicts of interest.

#### Acknowledgments

This research was partially funded by grants from the National Natural Science Foundation of China (51877181, 61703345, and 51607146), the Key Fund Project of the Sichuan Provincial Education Department (18ZA0459), the Key Scientific Research Fund Project of Xihua University (Z17108), and the Young Scholars Reserve Talents Support Project of Xihua University.

#### References

- [1] C. Vellaithurai, A. Srivastava, S. Zonouz, and R. Berthier, "CPIndex: Cyber-physical vulnerability assessment for power-grid infrastructures," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 566–575, 2015.
- [2] M. M. Rana, L. Li, S. W. Su, and W. Xiang, "Consensus based smart grid state estimation algorithm," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3368–3375, 2018.
- [3] M. M. Rana and L. Li, "An overview of distributed microgrid state estimation and control for smart grids," *Sensors*, vol. 15, no. 2, pp. 4302–4325, 2015.
- [4] D. T. Nguyen, Y. Shen, and M. T. Thai, "Detecting critical nodes in interdependent power networks for vulnerability assessment," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 151–159, 2013.
- [5] M. M. Rana, L. Li, and S. W. Su, "Cyber attack protection and control of microgrids," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 2, pp. 602–609, 2018.
- [6] J. Le, C. Wang, W. Zhou, Y. Liu, and W. Cai, "A novel PLC channel modeling method and channel characteristic analysis of a smart distribution grid," *Protection and Control of Modern Power Systems*, vol. 2, no. 2, pp. 146–158, 2017.
- [7] A. Berizzi, "The Italian 2003 blackout," in *Proceedings of the 2004 IEEE Power Engineering Society General Meeting*, pp. 1673–1679, Denver, CO, USA, June 2004.
- [8] Q. Sun, L. Shi, Y. Ni, D. Si, and J. Zhu, "An enhanced cascading failure model integrating data mining technique," *Protection and Control of Modern Power Systems*, vol. 2, no. 2, pp. 19–28, 2017.
- [9] Z. Chen, J. Wu, Y. Xia et al., "Robustness of interdependent power grids and communication networks: a complex network perspective," *IEEE Transactions on Circuits and Systems II*, vol. 65, no. 1, pp. 115–119, 2018.
- [10] X. Ji, B. Wang, D. Liu et al., "Improving interdependent networks robustness by adding connectivity links," *Physica A: Statistical Mechanics and its Applications*, vol. 444, pp. 9–19, 2016.
- [11] O. Yagan, D. Qian, J. Zhang et al., "Optimal allocation of interconnecting links in cyber-physical systems: interdependence cascading failures, and robustness," *IEEE Transactions on Parallel and Distribution Systems*, vol. 23, no. 9, pp. 1708–1720, 2012.
- [12] J. Wang, C. Jiang, and J. Qian, "Robustness of interdependent networks with different link patterns against cascading failures," *Physica A: Statistical Mechanics and its Applications*, vol. 393, pp. 535–541, 2014.
- [13] K. Schneider, C.-C. Liu, and J.-P. Paul, "Assessment of interactions between power and telecommunications infrastructures," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1123–1130, 2006.

- [14] Q. Wang, M. Pipattanasomporn, M. Kuzlu, Y. Tang, Y. Li, and S. Rahman, "Framework for vulnerability assessment of communication systems for electric power grids," *IET Generation, Transmission & Distribution*, vol. 10, no. 2, pp. 477–486, 2016.
- [15] Y. Cai, Y. Cao, Y. Li, T. Huang, and B. Zhou, "Cascading failure analysis considering interaction between power grids and communication networks," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 530–538, 2016.
- [16] P. Huang, Y. Wang, and G. Yan, "Vulnerability analysis of electrical cyber physical systems using a simulation platform," in *Proceedings of the IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, pp. 489–494, Beijing, China, October 2017.
- [17] Y. Deng, H. Lin, S. Shukla et al., "Co-simulating power systems and communication network for accurate modeling and simulation of PMU based wide area and measurement systems using a global event scheduling technique," in *Proceedings of the Workshop on MSCPES*, p. 1, Berkeley, CA, USA, May 2013.
- [18] S. Tan, W. Song, D. Huang, Q. Dong, and L. Tong, "Distributed software emulator for cyber-physical analysis in smart grid," *IEEE Transactions on Emerging Topics in Computing*, vol. 5, no. 4, pp. 506–517, 2017.
- [19] M. Rana, "Architecture of the internet of energy network: An application to smart grid communications," *IEEE Access*, vol. 5, pp. 4704–4710, 2017.
- [20] S. E. Collier, "The emerging enernet: convergence of the smart grid with the internet of things," *IEEE Industry Applications Magazine*, vol. 23, no. 2, pp. 12–18, 2017.
- [21] D. Dzung, M. Naedele, T. P. von Hoff, and M. Crevatin, "Security for industrial communication systems," *Proceedings of the IEEE*, vol. 93, no. 6, pp. 1152–1177, 2005.
- [22] E. Bou-Harb, C. Fachkha, M. Pourzandi, M. Debbabi, and C. Assi, "Communication security for smart grid distribution networks," *IEEE Communications Magazine*, vol. 51, no. 1, pp. 42–49, 2013.
- [23] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine blackout: implications for false data injection attacks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3317–3318, 2017.
- [24] G. N. Ericsson, "Cyber security and power system communication-essential parts of a smart grid infrastructure," *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1501–1507, 2010.
- [25] R. K. Pandey and M. Misra, *Cyber Security Threats-Smart Grid Infrastructure*, NPSC, Bhubaneswar, India, 2016.
- [26] D. Wang and H. Qin, "Design and application of antivirus system in Qinghai electric power dispatching data network," *Qinghai Electric Power*, vol. 31, no. 3, pp. 61–63, 2010.
- [27] J. Yan, Y. Zhu, H. He, and Y. Sun, "Multi-contingency cascading analysis of smart grid based on self-organizing map," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 4, pp. 646–656, 2013.
- [28] X. Wei, J. Zhao, T. Huang, and E. Bompard, "A novel cascading faults graph based transmission network vulnerability assessment method," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2995–3000, 2018.
- [29] X. Wei, S. Gao, T. Huang et al., "Identification of two vulnerability features: a new framework for electrical networks based redistribution mechanism of complex networks," *Complexity*, vol. 2019, Article ID 3531209, 14 pages, 2019.
- [30] B. Qu and H. Wang, "SIS epidemic with heterogeneous infection rates," *IEEE Transaction on Network Science and Engineering*, vol. 4, no. 3, pp. 177–186, 2018.
- [31] I. Tomovski and L. Kocarev, "Simple algorithm for virus spreading control on complex networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 4, pp. 763–771, 2012.
- [32] Y. Tang, X. Han, Y. Wu et al., "Electric power system vulnerability assessment considering the influence of communication system," in *Proceedings of the*, vol. 35, pp. 6066–6074, 2015.
- [33] Z. Lu, Z. Meng, and S. Zhou, "Cascading failure analysis of bulk power system using small-world network model," in *Proceedings of the International Conference on Probabilistic Methods Applied to Power Systems*, pp. 635–640, Ames, IA, USA, 2004.
- [34] P. Crucitti, V. Latora, and M. Marchiori, "A topological analysis of the Italian electric power grid," *Physica A: Statistical Mechanics and its Applications*, vol. 338, no. 1-2, pp. 92–97, 2004.
- [35] M. Rosas-Casals, S. Valverde, and R. V. Solé, "Topological vulnerability of the European power grid under errors and attacks," *International Journal of Bifurcation and Chaos*, vol. 17, no. 7, pp. 2465–2475, 2007.
- [36] G. Li, W. Ju, X. Duan et al., "Transmission characteristics analysis of the electric power dispatching data network," *Proceedings of the CSEE*, vol. 32, no. 22, pp. 141–148, 2010.
- [37] J. Hu, Z. Li, and X. Duan, "Structural feature analysis of the electric power dispatching data network," *Proceedings of the CSEE*, vol. 29, no. 4, pp. 53–59, 2009.
- [38] M. Romero-L and L. Gallego, "Analysis of voltage sags propagation in distribution grids using a SI epidemic model," in *Proceedings of the 3rd IEEE Workshop on Power Electronics and Power Quality Applications, PEPQA 2017*, Bogota, Colombia, June 2017.
- [39] B. Jia, S. Liu, T. Zhou, and Z. Xu, "Opportunistic transmission mechanism based on si in mobile crowd sensing networks," in *Proceedings of the 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, pp. 211–215, Prague, Czech Republic, July 2017.
- [40] P. Crépey, F. P. Alvarez, and M. Barthélemy, "Epidemic variability in complex networks," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 73, no. 4, Article ID 046131, 2006.
- [41] D. Xu and X. Xu, "Modeling and control of dynamic network SIR based on community structure," in *Proceedings of the 26th Chinese Control and Decision Conference, CCDC 2014*, pp. 4648–4653, China, June 2014.
- [42] Z. Sun, B. Wang, J. Sheng, Y. Hu, Y. Wang, and J. Shao, "Identifying influential nodes in complex networks based on weighted formal concept analysis," *IEEE Access*, vol. 5, pp. 3777–3789, 2017.
- [43] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics and endemic states in complex networks," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 63, no. 6, Article ID 066117, 2001.
- [44] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, "Epidemic outbreaks in complex heterogeneous networks," *The European Physical Journal B*, vol. 26, no. 4, pp. 521–529, 2002.
- [45] Y. Tang, X. Han, and Y. Wu, "Electric power system vulnerability assessment considering the influence of communication system," *Proceedings of the CSEE*, vol. 35, no. 23, pp. 6066–6074, 2015.

- [46] A. Delavari and I. Kamwa, "Improved optimal decentralized load modulation for power system primary frequency regulation," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1013–1025, 2018.
- [47] Y. Guo, D. Zhang, J. Wan, and D. Yu, "Influence of direct air-cooled units on primary frequency regulation in power systems," *IET Generation, Transmission & Distribution*, vol. 11, no. 17, pp. 4365–4372, 2017.
- [48] X. Yu and C. Singh, "A practical approach for integrated power system vulnerability analysis with protection failures," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1811–1820, 2004.
- [49] Y. Cao, Y. Zhang, and Z. Bao, "Analysis of cascading failures under interactions between power grid and communication network," *Dianli Zidonghua Shebei/Electric Power Automation Equipment*, vol. 33, no. 1, pp. 7–11, 2013.
- [50] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

## Research Article

# Identification of Two Vulnerability Features: A New Framework for Electrical Networks Based on the Load Redistribution Mechanism of Complex Networks

Xiaoguang Wei,<sup>1</sup> Shibin Gao ,<sup>1</sup> Tao Huang ,<sup>2</sup> Tao Wang ,<sup>3</sup> and Wenli Fan<sup>1</sup>

<sup>1</sup>School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China

<sup>2</sup>Department of Energy, Politecnico di Torino, Torino 10129, Italy

<sup>3</sup>School of Electrical Engineering and Electronic Information, Xihua University, Chengdu 610039, China

Correspondence should be addressed to Shibin Gao; 1514754029@qq.com and Tao Huang; tao.huang@polito.it

Received 24 September 2018; Accepted 25 October 2018; Published 16 January 2019

Guest Editor: Seyedmohsen Hosseini

Copyright © 2019 Xiaoguang Wei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a new framework to analyze two vulnerability features, impactability and susceptibility, in electrical networks under deliberate attacks based on complex network theory: these two features are overlooked but vital in vulnerability analyses. To analyze these features, metrics are proposed based on correlation graphs constructed via critical paths, which replace the original physical network. Moreover, we analyze the relationship between the proposed metrics according to degree from the perspective of load redistribution mechanisms by adjusting parameters associated with the metrics, which can change the load redistribution rules. Finally, IEEE 118- and 300-bus systems and a realistic large-scale French grid are used to validate the effectiveness of the proposed metrics.

## 1. Introduction

Critical component identification is an important part of security analyses for electrical networks [1–3]. The main idea is to rank the weakness of the equipment in an electrical network via a set of metrics.

As an artificial network, electrical grids have topological similarities to other general networks. They also exhibit several typical features of complex networks, such as small-world properties [4–6]. Therefore, complex network theory (CNT) is a popular method to assess the vulnerability of electrical networks [3, 4, 7–13]. The construction of structural metrics is an important branch of vulnerability evaluations [14] based on CNT. CNT uses the connectivity information abstracted from the network to create indices based on statistics and, sometimes, physical features of the network are added to improve the effectiveness of the indices [10, 11].

However, there are still several problems with this method. Compared to general networks (or systems), an electrical network has its own characteristics that limit the wide application of CNT. First, analyzing the topological

structures of electrical networks without considering their operational status does not disclose the real features of the systems [10, 11]. Secondly, in most general networks, when a vertex (or an edge) of a network fails, the direct neighbors are the first to be affected or have the largest impact based on CNT. However, this is not generally true for electrical networks [15]. Moreover, the structural metrics are static indices [12, 13, 16] that only consider the normal operational status of the network. To overcome the above problems, statistical graphs [17, 18] are employed to analyze the vulnerability or cascading failures of electrical networks. For example, [19, 20] proposed a sequential attack graph (SAG) to identify critical nodes while [21] proposed a correlation matrix. In addition, [22] proposed influence graphs to analyze cascading failures. Statistical graphs have also provided promising options for security, because they comprehensively consider the topological, physical, and operational characteristics of a system.

In addition, another problem in which features of vertices (edges) (For clarity, hereinafter the terms “network, branch and node” are used only for electric systems and “graph,

edge and vertex” only for complex networks.) in complex network vulnerability detection, especially in electrical networks, should be distinguished, is often overlooked. For example, some vertices can easily spread faults leading to a high probability of a network failure event. Conversely, some vertices are easily affected by propagated faults. Therefore, it is necessary to devise a method to identify these two features of vertices and to better reveal the vulnerabilities of networks.

In summary, our main contributions are as follows.

First, we propose a new framework that employs statistical graphs to represent the useful information for analyzing the network vulnerability from the original physical grid, using CNT, compared to a traditional framework that employs the original topological structure of the grid.

Secondly, inspired by [19–24], we propose a correlation graph (CG) generated via critical paths to analyze the electrical network vulnerability.

Thirdly, using benchmarks, we analyze the topological properties of the CGs based on CNT via the cumulative distributions of the vertex degrees. According to the analysis, the CGs are scale-free graphs, which verifies that electrical networks have scale-free properties under deliberate attacks, as opposed to traditional complex network methods, which verify the properties by correlating the drop in the network demand (or efficiency) with the attacked branches.

Finally, we define two vulnerability features from the perspective of CNT and then map the features onto electrical networks. Further, we employ the scale-free structures of the CGs to construct vulnerability metrics for the first time to differentiate the two features from the perspective of the load redistribution mechanism of CNT. The features of the metrics are explained in detail, including their relationship with the degree.

In addition, note that, even though dynamic models analyzed by real-time simulation platforms[25] are more comprehensive for security analyses in the real world, they require much longer simulation times and result in an immense computational burden, which makes it difficult to analyze a large-scale network. Meanwhile, as a media connecting equipment in the power system, the transmission network has notably fast dynamics/transients, compared to rotating devices. In other words, the transmission network *per se* can usually be considered to be a static component. Therefore, static models from the perspective of the load redistribution are widely employed to analyze the network vulnerability in existing literature [10–24]. Based on above, we focus on understanding the nature of the transmission network using static models by the load redistribution from the entire network.

## 2. Correlation Graph

We constructed a CG to incorporate both the structural features and the operational status of power systems, using critical paths from the point of view of *load redistribution mechanism* (LRM). The constructed graph considers both the topological structures and the operational features under fault operation of the system. For example, branches of an

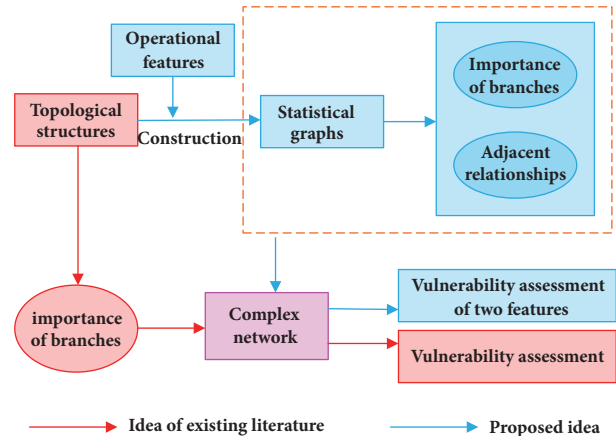


FIGURE 1: Comparison between proposed idea and existing literatures based on CNT.

electrical network can be transformed into vertices in a new graph while edges are formed to reflect the adjacent relationships between branches.

**2.1. Vulnerability Assessment: A New Framework.** To overcome the limitations of structural vulnerability identification methods by applying CNT to the electrical network vulnerability assessment, we need to consider the following two aspects: (1) the importance of a vertex and (2) the adjacent relationships between vertices. To assess the importance of a vertex, there are many indices (e.g., degree and betweenness) that can be used to qualify it from the perspective of LRM. Comparatively, there are few indices for quantifying the importance of branches because it is difficult to assess edges under LRM. In addition, in most general networks, when a vertex (or an edge) of a network fails, the adjacent relationship between vertices usually imply that the immediate neighbors are the first to be affected or suffer the largest impact based on CNT. However, this is not generally true for electrical networks; sometimes, nonadjacent branches are the first to be affected due to the physical laws of electric circuits and the physical and operational constraints [15]. Therefore only using the information of the structure of an electrical network cannot effectively identify the critical branches.

In summary, it is spatially insufficient to analyze the network vulnerability using only the topological structures of the grids. Therefore, we propose a new framework that employs statistical graphs [19–24] to represent information useful for analyzing the network vulnerability from the original physical grid, using CNT, as shown in Figure 1. In the existing methods, the topological structures are employed to assess the electrical network vulnerability based on the CNT on the original physical networks. Its main idea is to focus on the importance of branches by constructing statistical metrics without the involvement of the operational feature of the system. However, we construct statistical graphs comprehensively considering topological, physical and operational features of power systems, and further based on the constructed statistical graphs which can reveal not

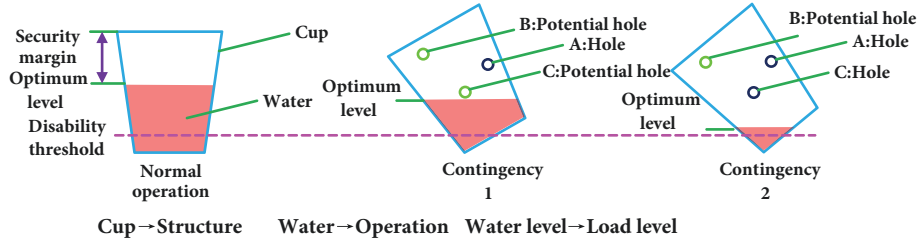


FIGURE 2: An example to explain the electrical vulnerability.

only importance of branches but also adjacent relationships among branches we assess the vulnerability with two features by replacing the original electrical networks.

**2.2. Correlation Graph.** Although many statistical graphs are proposed in references [19–21], there are still many limitations in the construction of statistical graphs to identify critical branches. In references [19, 20], a SAG was constructed by investigating different node combinations under sequential attacks. However, a SAG cannot be used to identify the critical branches of a large-scale electrical network because the different branch combinations will increase exponentially with increasing network scales. In reference [21], a correlation matrix was constructed under  $N-1$ , but an  $N-k$  contingency was not considered and the proposed method was not verified on a large-scale grid. Meanwhile, the above statistical graphs are only analyzed from the perspective of static statistical indices (e.g., degree).

Therefore, it is necessary to construct a new statistical graph to apply to the identification of vulnerable branches of large-scale networks under the  $N-k$  contingency and the corresponding properties should be analyzed in depth. To explain the rationale of the construction of a statistical graph, we use a cup of water as an example. The cup represents the topological structure of an electrical network and the water in the cup represents the operational status, as shown in Figure 2. Under normal operation, the system operator decides on an optimal/appropriate operational point, considering different constraints, including the necessary security margin. The optimal/appropriate point corresponds to a certain electric load level in the electrical network, represented by the water level in the cup. When a hole “A” in the cup is created, for example, by a contingency 1 due to structural damage, the optimal level for the water will change. Therefore “A” decides the optimal/appropriate water level, analogous to the electric load level, which can be viewed as the importance of the elements inside the contingency 1. Further, we assume that there are two potential holes, i.e., “B” and “C,” which can only be revealed after contingency 1, and that “C” is more decisive for the appropriate water level than “B.” This infers that the adjacent relationship “A”→“C” is more important than that of “A”→“B.” Therefore, the adjacent relationship “A”→“C” and the properties of “C” decide the appropriate water level. Accordingly, “A”→“C” can be viewed as the adjacent relationship between two branches during fault propagation.

Note that, in every step, we only need to pick the most decisive “hole” in the cup, which is nearest to the optimal level of the water.

To trace the adjacent relationships between branches, we need to consider different combinations. Now, the computational burden becomes an issue. For example, if we consider the  $N-k$  criterion, for an electrical network with  $N_L$  branches, we need to calculate  $N^k$  contingencies. For a French grid with 2596 branches, we need to calculate 17.5 billion contingencies for  $N-3$ .

Therefore, to simplify the calculation, we constructed critical paths [17, 18] to trace the adjacent relationships. We employed a *Branch loading assessment index* (BLAI) introduced in our precious work [17, 18, 26] to select an attacked branch, having the largest impact on the electrical network, as the next contingency. In addition, a commonly used termination condition, i.e. the blackout size, was used to mark the end of the critical paths [20].

To select a branch, the BLAI is employed to reflect the loading burden and its possibility of failure under the current contingency from the perspective of the load redistribution. The index can be calculated as

$$\alpha_j = \frac{f_j^x - f_j^0}{f_j^M} \exp\left(\frac{f_j^x - f_j^M}{f_j^M}\right) \quad (1)$$

The *blackout size* is adopted to mark the end of the critical paths and is viewed as a measure of the gravity of a critical path. The blackout size is defined as

$$\delta_z^x = 1 - \frac{\sum_{d \in B_z} P_d^x}{\sum_{d \in B_z} P_d^{(x-1)}} \quad (2)$$

$$\Lambda = \sum_{x=1}^{N_S-1} \sum_{z=1}^{Z^x} \delta_z^x \quad (3)$$

and when  $\Lambda \geq \Delta$ , we terminate the process.

Based on the above-mentioned considerations, we employ the structural features and the operational status of the electrical network to construct a CG to reveal the adjacent relationships between the branches. Using the CG, the spatial association network between branches was translated into a CG. However, before introducing the CG, we define the vulnerability relationship to describe the relationship between the two branches.

**Vulnerability Relationship.** We denote two adjacent links on a critical path as having a vulnerability relationship.

*Critical Path Generation Method.* To explore the vulnerability relationship, we use the critical paths of a network to construct a CG. For a network with  $N$  branches, we treat every branch as a triggering fault. We can obtain  $N$  paths of a critical path. Note that some paths may contain only one vertex. DC-OPF (optimal power flow) is employed to optimize the operation status in different topological structures. The critical path is generated as follows.

*Step 1.* Input the electrical network information. Initialize  $S = \phi$  and  $\Delta$ . Select a branch as a triggering fault.

*Step 2.* Remove the selected branch from the electrical network and add it to  $S$ .

*Step 3* (island detection and partition). Calculate the DC power flow [27–29] of every island. Employ (1) to calculate  $\alpha$  for every branch.

*Step 4.* Take (2) as the objective function. Calculate the minimum  $\delta$  of every island using the DC-OPF algorithm [1, 28]. Employ (3) to calculate  $\Lambda$ .

*Step 5.* If  $\Lambda \geq \Delta$ , end the critical path generation process; otherwise select the branch whose  $\alpha$  is the maximum of all the branches as the candidate branch under next contingency scenario and go to Step 2.

*Step 6.* Output  $S$ .

By above process, we can simply and quickly develop critical paths and efficiently reduce the computational burden.

*CG Generation Method.* To map a critical path  $S^i = \{L_1^i, L_2^i, \dots, L_{N_s}^i\}$  onto a graph  $G^i$ , let the branches in  $S^i$  be the vertices of  $G^i$ , i.e.,  $V^i = \{v_j \mid v_j = L_j^i, j = 1, 2, \dots, N_{N_s}^i\}$ . Then the edges can be defined as  $E^i = \{e_q^i \mid e_q^i = L_j^i L_{j+1}^i, j = 1, 2, \dots, N_{N_s}^i - 1\}$ . By merging corresponding  $N_L$  graphs, we can obtain the CG, i.e.,  $G = G^1 \cup G^2 \cup \dots \cup G^{N_L}$ . Finally the CG is represented as

$$G = \{(V, E) \mid V = V^1 \cup V^2 \cup \dots \cup V^{N_L}, E = E^1 \cup E^2 \cup \dots \cup E^{N_s}\}. \quad (4)$$

Obviously, the CG is an undirected and unweighted graph.

*CG Topological Features.* To analyze the topological features of the CG, we employ four benchmark systems (described in Table 1). We set the threshold  $\Delta = 20\%$  [20]. The CG of the IEEE 14-bus system, shown in Figure 3, manifests the adjacent vulnerability relationship between the branches. Using the CG, the spatial association network between the branches in the electrical network can be translated into a statistical graph.

The cumulative distributions of the vertex degree [29]  $P(K > k) = \sum_{K>k} P(k)$  in CGs are all power laws whose  $r$  and  $R^2$  are given in Table 2, except for the CG of the

TABLE 1: Description of test benchmarks.

Test benchmarks	$N_B$	$N_W$	$N_L$
IEEE 14-bus system	14	5	20
IEEE 118-bus system	118	54	186
IEEE 300-bus system	300	69	411
French Grid	1951	391	2596

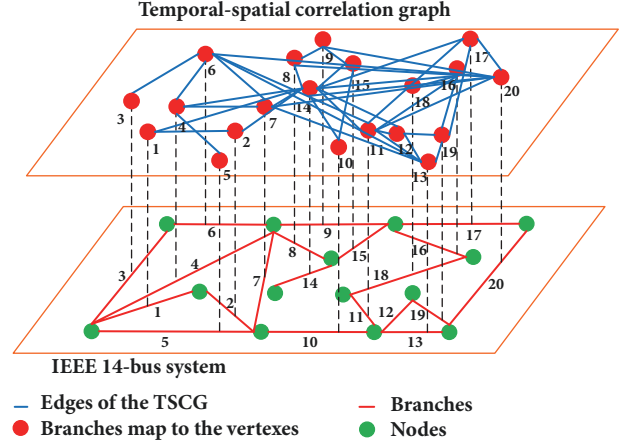


FIGURE 3: Mapping between IEEE 14-bus system and CG.

TABLE 2: Parameters  $r$  and  $R^2$  of cumulative distributions of the vertex degree in three CGs. Generally,  $R^2 \geq 80\%$  has a satisfactory fitting effect.

Test benchmarks	$r$	$R^2$
IEEE 14-bus system	0.8880	0.7254
IEEE 118-bus system	1.3023	0.9338
IEEE 300-bus system	1.3500	0.9166
French grid	1.2180	0.9497

IEEE 14-bus system because its vertices are too few to allow statistical conclusions to be drawn. Table 2 indicates that CGs are scale-free graphs (i.e.,  $P(K > k) \sim x^{-r}$ ), which have high robustness under random vertex attacks, but low robustness under intentional attacks. In addition, we can employ the CGs to verify the scale-free properties of electrical networks under deliberate attack, and compared them to traditional complex network methods which verify the properties by correlating the drop in the network demand (or efficiency) to the attacked branches. Due to its scale-free features, its statistics of under faults operation and its vulnerability relationships between branches, we can indirectly assess the electrical network vulnerability using the CG.

### 3. CG Based Vulnerable Indices with Two Vulnerability Features Using CNT

In this section, we define the two vulnerability features. Then we propose CG based metrics to differentiate the two features to assess the electrical network vulnerability from the perspective of LRM. Before defining the two vulnerability

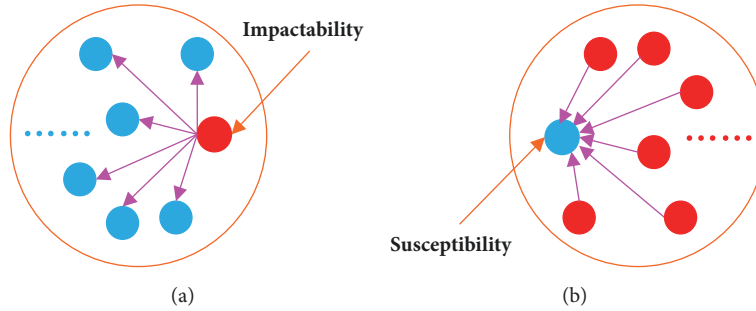


FIGURE 4: The diagram to explain impactability and susceptibility.

features, we briefly introduce the taxonomy from CNT used in this paper.

*Flow* is a tangible or intangible substance that exists in a specific network. For example, in physical networks, such as water networks, the water flow is the flow. In social networks, the flow is the communication between people.

*Load* is defined as the quantity of a flow that a vertex owns for a certain status of the network [30].

*Load capacity* is the maximum load that a vertex can withstand [31]. Once the load of a vertex is beyond the load capacity, the vertex is then in *overload*.

*Load redistribution* is the process where, when a vertex fails or is removed, its corresponding load needs to be reallocated or transferred to other vertices following certain rules. In particular, when the corresponding load is reallocated to the adjacent vertices that are directly connected to the fault vertex, the redistribution rule is called *neighbor distribution rule (NDR)* [32]; meanwhile when the load is reallocated to all the other vertices, the redistribution rule is called the *global distribution rule (GDR)* [31]. When the corresponding load is evenly reallocated to selected vertices, the redistribution rule is called a *uniformity distribution rule*.

In this paper, we analyze the vulnerability from the angle of LRM of CNT [33, 34]. That is, we assess the vulnerability of the network by the changes due to load redistributions when a vertex fails.

**3.1. Two Vulnerability Features.** At present, a popular vulnerability analysis is to identify the critical vertices (or edges) of an electrical network, which easily leads to a network failure event under deliberate attacks [1]. Such approaches only identify critical branches that can easily affect the network when they fail, as shown in Figure 4(a). However, they ignore the other feature of critical branches which are easily affected by faults of other branches, as shown in Figure 4(b). Therefore it is necessary to differentiate the two vulnerability features and we call *impactability* and *susceptibility*.

When a branch fails (is attacked), it causes obvious or serious changes in the original status of the network in one or more aspects, such as the topological structure and the function, and in such a case, the branch is called an *impactable vertex*.

When one or more network branches fail, if a branch is easily affected by the faults, leading to changes in the original status of this branch, such as the load increasing

(even overload), branch failure, then the branch is called a *susceptible vertex*.

The impactability of a branch in an electrical network describes how the failure of that branch can cause a considerable load increase in other branches, leading to serious system changes. The susceptibility of a branch in an electrical network describes, when other branches fail, the branch, and how easily the branch is affected by a fault, leading to a severe load increase or failure.

The differentiation of these two features of branches has practical implications. First, it provides useful lists of critical branches corresponding to different operation states for power dispatchers to monitor. For example, under normal operation, dispatchers need to give priority to impactable branches because they can easily spread faults when they are attacked. Conversely, under fault operation, dispatchers should also pay attention to susceptible branches because they can easily be affected by other faults. Secondly, analyzing the impactability of branches provides suggestions to offenders about how to cause significantly large disturbances to a system and to defenders about how to protect the safe operation of a system at the lowest cost. By analyzing the susceptibility of branches, it can reduce or avoid further deterioration of the system under a deliberate attack. In summary, the differentiation of the two features can improve the management of system security.

**3.2. Proposed Vulnerability Metrics with the Two Features.** Previous studies of the load in LRM in complex networks have primarily focused on the load model (e.g., the initial load) and the load redistribution strategy. For example, the betweenness or degree is generally employed to define the initial load of a vertex [32, 33, 35–37]. Similarly, we employ the degree to define the load of a vertex because the degree can reflect its importance in the propagation process. For the load redistribution strategy, [33, 36] investigated NDR while [37] proposed stochastic probability redistribution models. However, the redistribution rule they proposed is not adjustable. References [32, 34] first considered the adjustable redistribution rules for the load. To define the vulnerability metrics by means of LRM, we adopted the adjustable load redistribution strategy to study LRM. Before defining the metrics, we introduce some new concepts based on CNT as follows.



The *vulnerability flow* is a virtual substance that reflects the vulnerability relationship between the branches. For example, we can define the flow that exists in the CG as the vulnerability flow. The vertices of the CG carry a certain proportion of the vulnerability flow called the *load*. The edges of the CG reveal the paths of transmitted load. When a branch of an electrical network fails, it causes changes in other branches in terms of their loads. If we map a contingency onto the CG, it describes the corresponding vertex failures, causing the reallocation of the load onto other vertices (i.e., load redistribution) via the edges.

*Initial Load.* In a CG, a higher degree vertex plays a more important role in the fault propagation process; therefore we employ the degree of the vertices to quantify the amount of the initial load [34]. The initial load of vertex  $v_j$  is expressed as

$$\rho_j = \frac{D_j^\tau}{\sum_{j=1}^{N_L} D_j^\tau} \quad (5)$$

*Distance.* In a CG, there may be more than one path between any two vertices. Therefore, to quantitatively depict the distance between any two vertices, we use the minimum path between them [32, 34] to quantify their vulnerability relationship.

In an electrical network, the failure of a branch will cause the redistribution of the initial power flow in the fault branch to other branches, leading to an increase in the transmitted power over other branches. Correspondingly, when we map a contingency onto the CG, the relevant fault vertex will lead to the redistribution of its initial load to other vertices. Therefore we use the increase in the vulnerability flow at other vertices to quantify the impact of the corresponding fault branch on the electrical network.

The load redistribution  $\Delta\rho_{j \rightarrow k}$  from vertex  $v_j$  to vertex  $v_k$  in the set of affected vertices  $Q_j$  ( $v_k \in Q_j$ ) is defined as

$$\Delta\rho_{j \rightarrow k} = \begin{cases} \frac{l_{jk}^{-\lambda}}{\sum_{\gamma \in Q_j} l_{j\gamma}^{-\lambda}} \rho_j, & k \in Q_j \\ 0, & k \notin Q_j \end{cases} \quad (6)$$

$$Q_j = \left\{ k \mid \frac{l_{jk}^{-\lambda}}{\sum_{\gamma=1}^{N_L} l_{j\gamma}^{-\lambda}} > \eta, 0 < \eta \leq \frac{1}{N_L - 1} \right\} \quad (7)$$

Obviously, the load redistribution is proportion to the distance between  $v_j$  and  $v_k$ .  $\lambda$  controls the portion of the load that  $v_j$  reallocates to  $v_k$ , and  $\eta$  is used to define the set  $Q_j$ . Equation (6) reflects not only the importance of a vertex but also the vulnerability relationships between that vertex and others. We employ the parameters  $\tau$  and  $\lambda$  to adjust the proportion between the *importance of a vertex* and its *vulnerability relationships*.

*Impactability Metric (IM).* To describe the impactability of a fault vertex  $v_j$  in a CG, we introduce the entropy

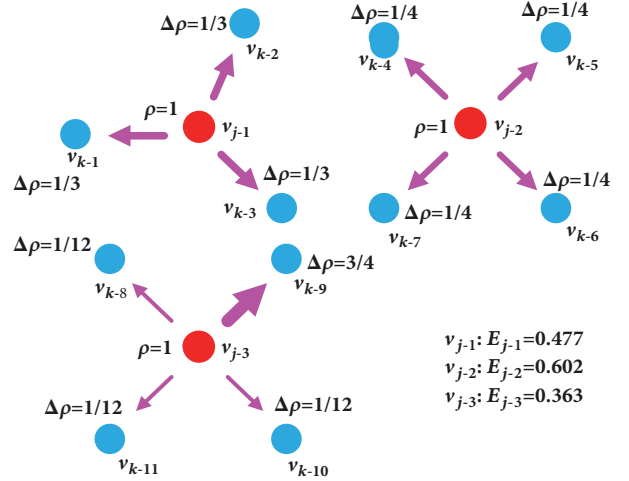


FIGURE 5: An example to understand (8).

expressed in the following to measure the load change in the graph.

$$E_j = - \sum_{k \in Q_j} \Delta\rho_{j \rightarrow k} \ln \Delta\rho_{j \rightarrow k} \quad (8)$$

Eq. (8) considers the severity of the fault in terms of the affected number of vertices and measures the uniformity of the load redistribution. Further,  $E_j$  is larger for more severe faults and/or more even load redistributions between affected vertices. To illustrate the two aspects we mentioned for Equation (8), we use 3 vertices depicted in Figure 5. Assuming under even load redistribution strategy, the  $E_{j-2}$  of  $v_{j-2}$  is greater than  $E_{j-1}$  of  $v_{j-1}$  solely because  $v_{j-2}$  affects more vertices. In contrast, with the same number of affected vertices,  $E_{j-2}$  of  $v_{j-2}$  is greater than  $E_{j-3}$  of  $v_{j-3}$  due to a more uniform load redistribution.

However, note that some vertices may exist whose entropy is greater than that of others, due to more even load redistributions between larger affected number of vertices, yet cause small load increases for other vertices, resulting in a small overall impact. To exclude these vertices, we further refine

$$IV_j = \frac{1}{N_Q} E_j = - \frac{1}{N_Q} \sum_{k \in Q_j} \Delta\rho_{j \rightarrow k} \ln \Delta\rho_{j \rightarrow k} = - \frac{1}{N_Q} \cdot \sum_{k \in Q_j} \frac{l_{jk}^{-\lambda}}{\sum_{r \in Q_j} l_{jr}^{-\lambda}} \left( \frac{D_j^\tau}{\sum_{j=1}^{N_L} D_j^\tau} \right) \cdot \ln \frac{l_{jk}^{-\lambda}}{\sum_{\gamma \in Q_j} l_{j\gamma}^{-\lambda}} \left( \frac{D_j^\tau}{\sum_{j=1}^{N_L} D_j^\tau} \right) \quad (9)$$

We define  $0 \ln 0 = 0$ .

*Susceptibility Metric (SM).* To describe the susceptibility of an affected vertex  $v_k$  in a CG, we use the average incremental

load redistribution from all other fault vertices into that vertex:

$$SV_k = \frac{1}{\sum_{j=1}^{N_L} \sigma(\Delta\rho_{j \rightarrow k})} \sum_{j=1}^{N_L} \Delta\rho_{j \rightarrow k} \quad (10)$$

$$\forall v_j \in G, j \neq k$$

$$\sigma(\Delta\rho_{j \rightarrow k}) = \begin{cases} 1, & \Delta\rho_{j \rightarrow k} > 0 \\ 0, & \Delta\rho_{j \rightarrow k} \leq 0 \end{cases} \quad (11)$$

**3.3. Load Redistribution Rules for IM and SM.** In this subsection, we discuss two rules for the load redistribution: NDR and GDR which explore the relationship between the two metrics and the degree.

**Neighborhood Distribution Rule.** When  $\lambda = +\infty$ , the load redistribution rule is NDR.

(1) For IM, (9) is simplified as

$$IV_j = -\frac{D_j^{\tau-1}}{\sum_{j=1}^{N_L} D_j^{\tau}} \ln \frac{D_j^{\tau-1}}{\sum_{j=1}^{N_L} D_j^{\tau}} \quad (12)$$

When  $0 < \tau < 1$ ,  $IV_j$  is inversely proportional to  $D_j$ . When  $\tau = 1$ ,  $IV_j$  is constant, i.e.,  $IV_1 = IV_2 = \dots = IV_{N_L}$ . Clearly,  $IV_j$  is invalid in terms of identifying the impactability for IM under NDR. When  $\tau > 1$ ,  $IV_j$  is proportional to  $D_j$ .

(2) For SM, (10) is simplified as

$$SV_k = \frac{1}{D_k} \sum_{j=1}^{D_k} \frac{D_j^{\tau-1}}{\sum_{j=1}^{N_L} D_j^{\tau}} \quad (13)$$

When  $\tau = 1$ ,  $SV_k$  is a constant and  $SV_1 = SV_2 = \dots = SV_{N_L}$ . When  $\tau \neq 1$ ,  $SV_k$  is determined by  $D_j$  and  $D_k$ . If for two existing vertices  $D_{k1} = D_{k2}$  and  $\sum_{j=1}^{D_{k1}} D_j^{\tau-1} > \sum_{j=1}^{D_{k2}} D_j^{\tau-1}$ , then  $SV_{k1} > SV_{k2}$ . This indicates that a vertex is more susceptible to its high vertex degree neighbors.

**Global Distribution Rule.** When  $\lambda = 0$ , then the load redistribution rule is the GDR.

(1) For IM, (9) is simplified as

$$IV_j = -\frac{1}{N_L - 1} \left( \frac{D_j^{\tau}}{\sum_{j=1}^{N_L} D_j^{\tau}} \right) \ln \frac{1}{N_L - 1} \left( \frac{D_j^{\tau}}{\sum_{j=1}^{N_L} D_j^{\tau}} \right) \quad (14)$$

(2) For SM, (10) is simplified as

$$SV_k = \frac{1}{(N_L - 1)^2} \left( 1 - \frac{D_k^{\tau}}{\sum_{j=1}^{N_L} D_j^{\tau}} \right) \quad (15)$$

Equations (14) and (15) show that  $IV_j$  is proportional to the degree  $D_j$ , while  $SV_k$  is inversely proportional to the degree  $D_k$ .

The analysis shows that when  $\tau$  has different values; the impactability and susceptibility of the branches are different under the same redistribution rule. In addition, one can infer that compared with all other topologies, the central vertex in a star graph has the largest impactability and susceptibility (refer to the proof in the appendix).

## 4. Simulation and Analysis

The simulations were performed for the IEEE 118-bus system, the IEEE 300-bus system, and a French grid and were implemented in MATLAB to verify the validity of the proposed method by sequentially attacking branches to calculate the total amount of affected loads and the relevant load decreasing speeds. For all the simulations conducted below, we set the parameter  $\eta = 5.40 \times 10^{-3}$ ,  $2.43 \times 10^{-3}$ , and  $3.85 \times 10^{-4}$  in the IEEE 118- and 300- bus systems and the French grid, respectively.

**4.1. Relationship between the Two Metrics and the Degree under Different Load Redistribution Rules.** To visualize the relationships under different rules, we employ the TSCG of French grid and change the values of  $\lambda$  ( $\lambda = 0, +\infty, 0.5$ ) and  $\tau$  ( $\tau = 0.1, 1, 1.5$ ).

(1)  $\lambda = +\infty$  (NDR). It can be seen from Figure 6 that when  $\tau = 1$ , both IM and SM are constant and do not change with the degree of vertexes. When  $\tau = 0.1$  and  $\tau = 1.5$ , the values of the IM of a vertex are inverse and direct power law functions of the vertex's own degree, respectively. In contrast, the values of the SM of a vertex are not closely related to the vertex's own degree. In other words, the degree is invalid to identify the vulnerability of the vertexes. However, in this case, our proposed SM is still valid to identify the vulnerable vertexes.

(2)  $\lambda = 0$  (GDR). Figure 7 shows that the IM and SM of a vertex are inverse and direct proportional to the vertex's own degree, respectively. It demonstrates that, under the GDR, the metrics of a vertex are only dependent on its own degree.

(3)  $\lambda = 0.5$  (an example of some value in  $(0, +\infty)$ ). When  $\lambda \in (0, +\infty)$ , the corresponding redistribution rule can be regarded as somewhere in between the GDR and the NDR. For example, in Figure 8(a), when  $\tau = 0.1$ , the IM of a vertex is not obviously in proportional relationship with its own degree. It indicates that the distance plays a more important role in (12) in this case. On the contrary, when  $\tau = 1$  or  $\tau = 1.5$ , the values are approximately of power law with respect to a vertex's own degree. By simulations performed on different systems for many times, the authors find that when  $\lambda$  takes any fixed value in  $(0, +\infty)$ , with the increase of  $\tau$ , the degree have more impact on the IM than the distance. In Figure 8(b), the SM is decided jointly by both the distance and degree, and when  $\lambda$  takes any fixed value in  $(0, +\infty)$ , the SM increases with the growth of  $\tau$ .

Further, we analyze the changes of metrics with  $\lambda$  when  $\tau$  is fixed. Figures 9(a) and 9(b) display the changes of two randomly chosen branches: the IM of branch 513 and the SM of branch 2372, respectively. When  $\lambda$  changes from 0.1

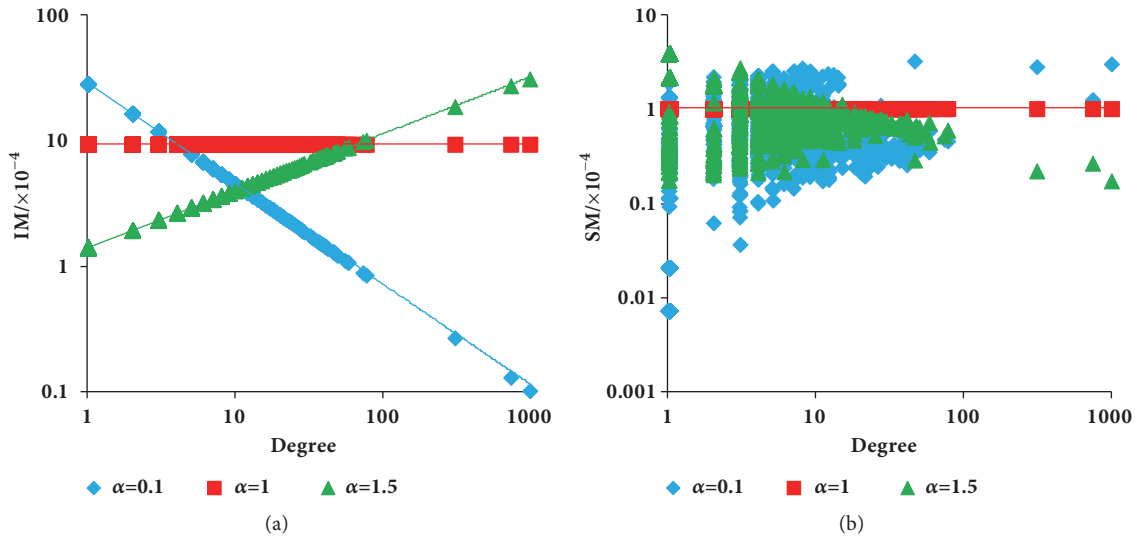


FIGURE 6: The relationship between metrics and degree in NDR. (a)IM and (b)SM.

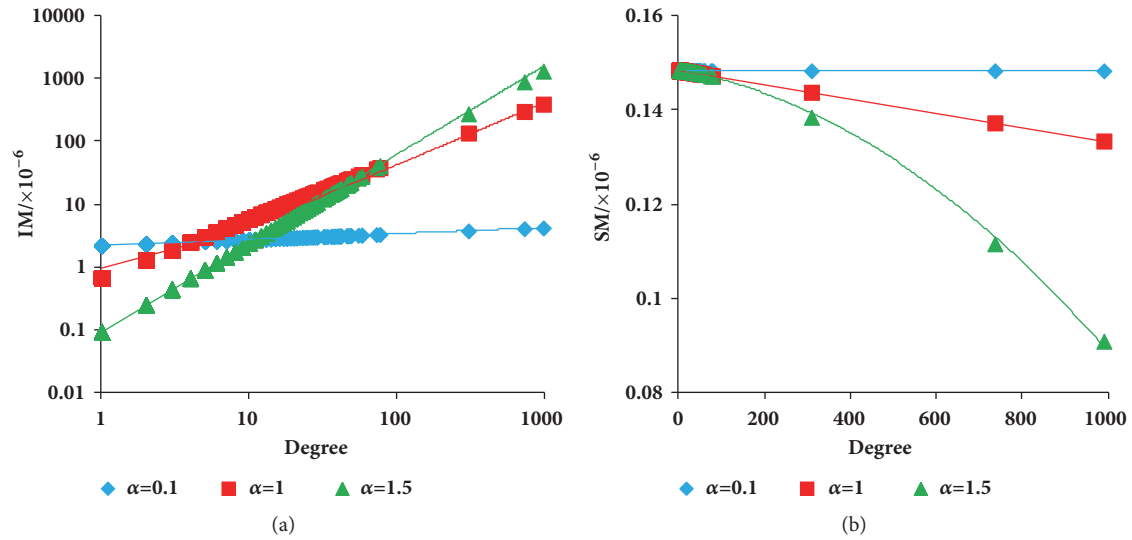


FIGURE 7: The relationship between metrics and degree in GDR. (a)IM and (b)SM.

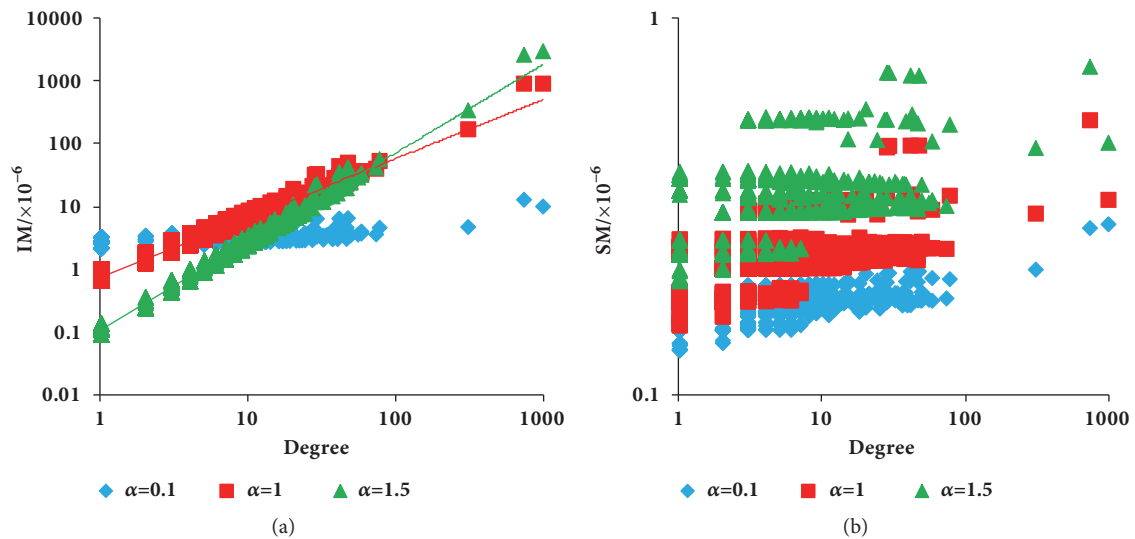


FIGURE 8: The relationship between metrics and degree in between. (a)IM and (b)SM.

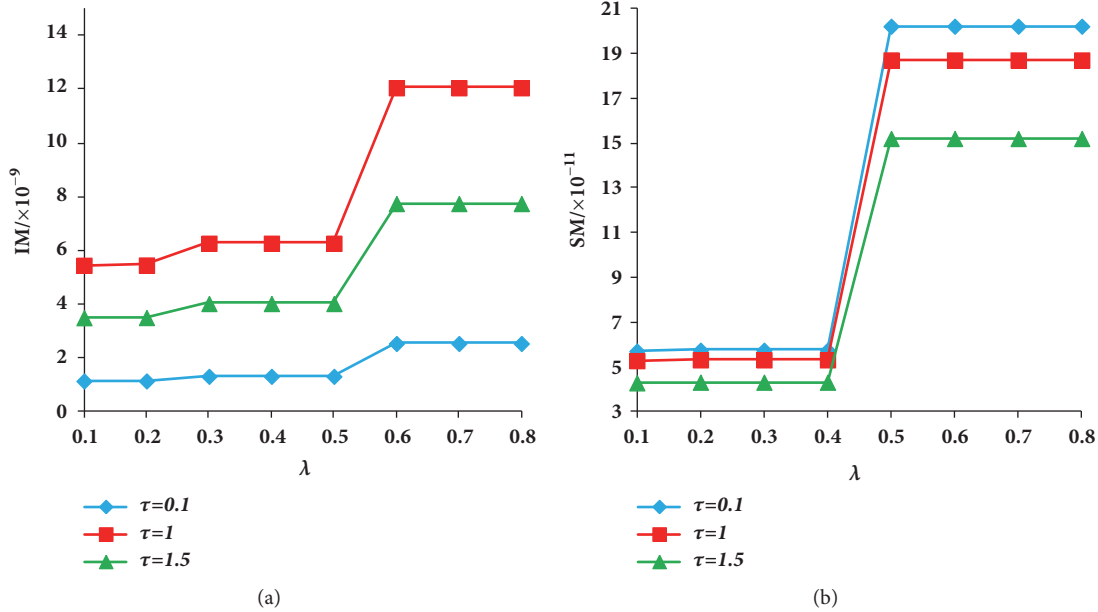


FIGURE 9: The trend of the change of metrics with  $\lambda$  increasing. Here, (a) IM of branch 513 and (b) SM of branch 2372.

to 0.8 with an interval of 0.1, both the IM and the SM exhibit stepwise features, which indicates that there exists points that divide  $\lambda$  into different segments, and the values of the IM and SM are insensitive to  $\lambda$  in each segment. In practice, in order to enlarge the discriminative ability of the proposed metrics, a larger  $\lambda$  is recommended.

In summary, our proposed metrics can identify the vulnerability when degree fails. The  $\lambda$  is vital in the two metrics as it defines the elements in  $Q_j$ , which further affects the relationship between the metrics and degree. The proposed metrics can be applied to both neighborhood and global redistribution rules by simply adjust  $\lambda$  and  $\tau$ .

**4.2. Vulnerability Analysis of Electric Networks.** To verify the validity of the proposed method, we sequentially attack the 20, 40 and 140 critical branches in the IEEE 118- and 300-bus systems and the French grid, respectively, as identified by IM and SM, and evaluated the remaining load and rate of load decrease. For all the attacks, after removing an identified branch, DC-OPF was used to redispatch the network, with the objective of minimizing the load shedding.

The *remaining load* was used to evaluate the gravity of the attack, and is obtained at the end of the simulation.

The *rate of load decrease* was adopted to reflect the speed at which the load decrease reached important vertices. We calculated the slope between the adjacent samples and then the average slope was used to represent the rate of load decrease, as shown in

$$\kappa = \frac{1}{Y} \sum_{y=2}^Y (\psi_y - \psi_{y-1}) \quad (16)$$

In our simulations,  $\lambda$  varied from 0 to 2 with an interval of 0.2 and  $\tau$  varied from 0 to 2 with an interval of 0.2. In addition, to consider NDR, we also set  $\lambda = +\infty$ . With all these

TABLE 3: Performances of the metrics by remaining load.

Test benchmarks	$\psi$	IM	SM
IEEE 118-bus system	>80%	<b>52.81%</b>	80.32%
	>90%	<b>23.87%</b>	58.74%
IEEE 300-bus system	>80%	<b>35.09%</b>	74.67%
	>90%	<b>16.02%</b>	43.32%
French grid	>80%	100%	100%
	>90%	<b>52.46%</b>	100%

TABLE 4: Comparisons between the IM & SM using remaining load.

Test benchmarks	$\psi_{IM} < \psi_{SM}$	$\psi_{IM} > \psi_{SM}$
IEEE 118-bus system	96.28%	3.72%
IEEE 300-bus system	94.29%	5.71%
French grid	94.47%	5.53%

TABLE 5: Comparisons of the metrics by the rate of load decrease.

Test benchmarks	$\kappa_{IM} < \kappa_{SM}$	$\kappa_{IM} > \kappa_{SM}$
IEEE 118-bus system	76.37%	23.63%
IEEE 300-bus system	79.21%	20.79%
French grid	67.19%	32.81%

combinations, for the IM and SM, we can obtain  $12 \times 11 = 132$  sets of critical branches.

The simulation results are given in Tables 3–5. Table 3 gives the statistical results of the 132 attacks. IM is better at identifying critical branches than SM. For example, removing branches according to the IM can cause the grid to lose more than 10% of its load in approximately half of the simulations of the French grid, compared to SM.

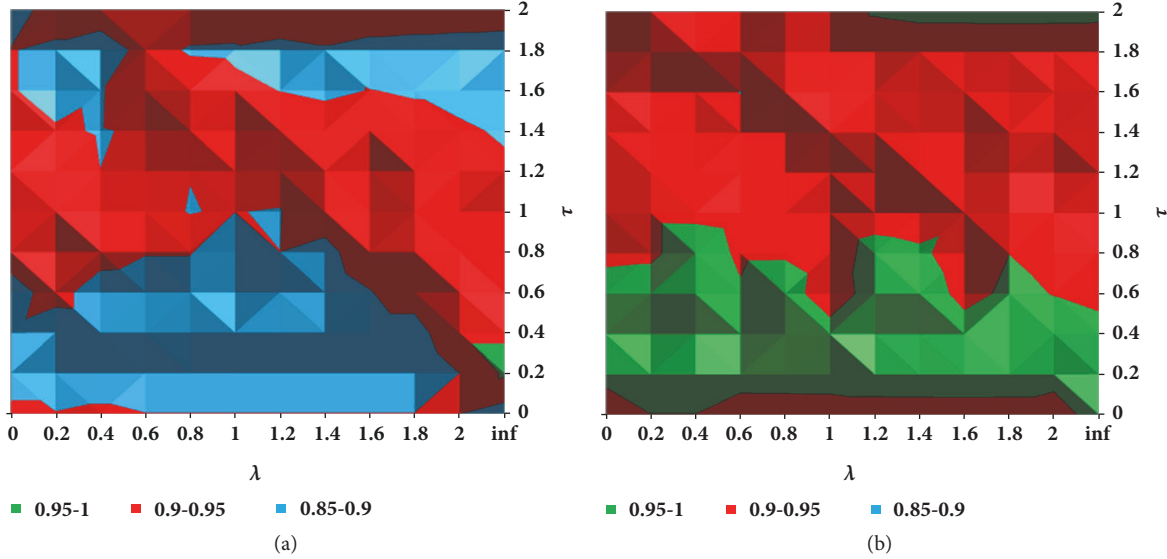


FIGURE 10: Distribution of removing load under different parametric combinations on the French grid. (a)IM and (b)SM.

To be more quantitative, Table 4 shows the performance difference of the two metrics according to the residual load of the systems under the same  $\lambda$  and  $\tau$ . The second to the third columns list the percentage of the two metrics in which (1) IM is better than SM and (2) SM is better than IM. It is clear that a load associated with a fault can easily be redistributed to many other branches; therefore, the branches can distinctively exhibit one of the two features, i.e., impactability or susceptibility.

In addition to the percentage of the remaining load, we compared the performance of the metrics according to the speed of the load decrease, as shown in Table 5. Similarly, in the majority of the simulations, the speed of the load loss after an attack following IM is faster than that after SM in the three benchmarks. The speed signifies the intensity of the attack. The faster the load decreases the more difficult it is for the system operator to apply control strategies to stop cascades.

Combining the results from Tables 3–5, we can infer that the identification of the two features becomes increasingly important, because they reveal different features in terms of the fault propagation. In general, both from the perspective of a power grid and a general graph, the branches with high impactability more easily spread faults; therefore, they decide the speed of the consequences (in the power system cases used in this paper, this is the loss of the load) and the affected area. Conversely, when branches with high susceptibility are attacked, the fault propagation is slow and load loss is small. This is because the susceptible vertices do not propagate faults as easy as impactable vertices do; they usually work as propagation sinks. So they define the consequences of a cascade.

Accordingly, in practice, with the distinction of the two features of the branches, a system operator can deploy better defense strategies resorting to the most relevant features under different operation states. Protecting impactable branches can effectively avoid triggering failures. However,

during a cascading process, particularly under deliberate attacks or fault propagation, susceptible branches can be easily affected by a fault, which can deteriorate the network functionality due to the enhanced consequences from them. Therefore susceptible branches also need to be protected and considered as well.

Furthermore, we investigated the distribution of removed load for different parametric combinations. Due to space limitations, we offer only the French grid (Figure 10) as an example. In Figure 10, when there are different combinations of  $\lambda$  and  $\tau$ , the load removed from the system is different, which demonstrates the importance of the branches and the adjacent relationships between vertices on determining the vulnerability of a system. In addition, the distribution of the removed load is relatively centralized. For example, the removed load is less when  $\lambda$  and  $\tau$  are 0-1.8 and 0-0.8, respectively. Meanwhile, note that to obtain the optimum of parametric combination, which is analogous to the parameter selection of deep-learning algorithms, some optimization algorithms, such as genetic algorithms, can be employed.

**4.3. Comparison with Existing Methods.** Compared to susceptible branches, because impactable branches can easily spread faults, which cause the grid to collapse faster, impactable branches will be primary targets for deliberate attacks. To verify this, we compared the proposed impactable branches to the critical branches as ranked by the degrees of the CGs, betweenness, electrical betweenness, network efficiency [11] and network ability [19] of system structures on the French grid, and the IEEE 118- and 300- bus systems.

In the three benchmarks, when (1)  $\lambda = 0.6$  and  $\tau = 2$ , (2)  $\lambda = 0.6$  and  $\tau = 1.4$ , and (3)  $\lambda = 0.6$  and  $\tau = 1$ , the rankings of the impactable branches had the optimum values, respectively. Figure 11 shows that the remaining load after the removal of the branches ranked by IM of the branches is generally smaller than the degree of the CG. This

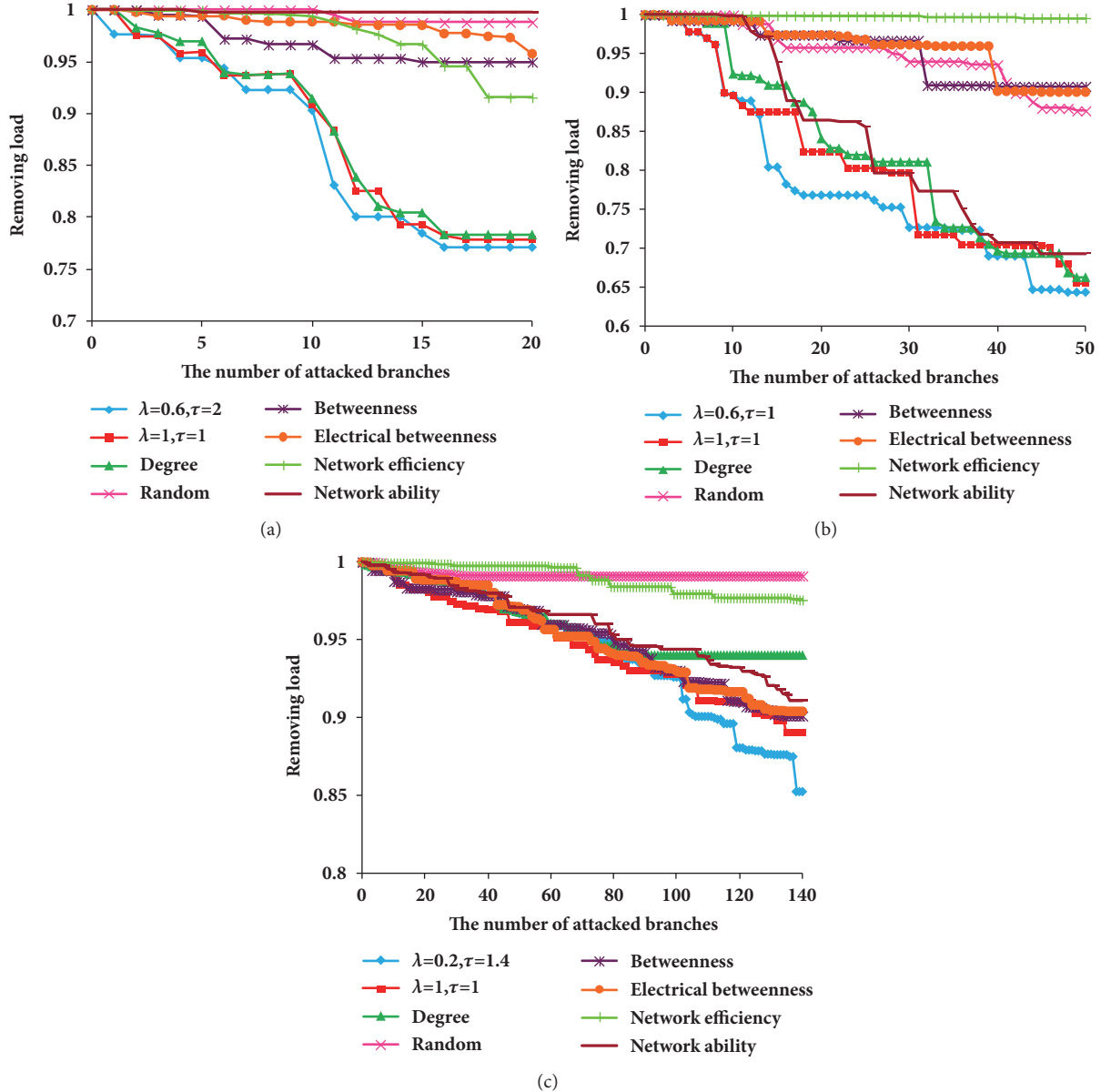


FIGURE 11: The remaining load with the number of the critical branches increasing. (a) IEEE 118-bus system; (b) IEEE 300-bus system; (c) French grid.

indicates that the vulnerability of branches is related not only to the importance of the branches but also to the adjacent relationships between the vertices. Therefore, considering the relationships to construct indices is necessary to improve the accuracy of vulnerability assessments.

Further, we compared our proposed method to other indices and random attacks for the three benchmarks. For random attacks, some branches are randomly selected and successively removed from benchmarks. Figure 11 shows that the remaining load after the removal of the branches identified by our model is smaller, which indicates that our proposed method has better accuracy when identifying vulnerable branches. In addition, in the three benchmarks, when  $\lambda = 1$  and  $\tau = 1$ , our results have also the relatively

better accuracy than other indices even if we did not set the adjustable parameters ( $\tau$  and  $\lambda$ ).

## 5. Conclusions

In this paper, we employed structural, physical, and operational features to construct a CG to analyze the electrical network vulnerability. On this basis, IM and SM are built to distinguish these two vulnerability features. Adjusting the parameters associated with the metrics can dynamically change the load redistribution rules. Simulations based on benchmarks systems proved the validity of the proposed method. Both IM and SM can identify critical branches of a system; however, IM is more effective than SM in most cases.

Thanks to the general features of the proposed method, i.e., from the perspective of the LRM of the CNT, the IM and SM can also be applied to identify the impactable and susceptible vertices of other networks, e.g., water networks and transportation networks. Similarly to the application we conduct in this paper for electric network, the IM and SM are also expected to not only identify the vulnerability in these networks but also reveal their roles in it.

However, there are still some existing problems to overcome. First, it is important to reduce the time complexity so that our proposed method can be applied to an online assessment of large scale electrical networks. Parallel or distributed computing such as PC clusters and, cloud computing may help to combat the time complexity issue. Secondly, even though adjusting the associated parameters of the two metrics can dynamically change the redistribution rules, there is currently no guidance available as to how to choose them.

## Appendix

### Proof of the Maximum of the IM and SM

We are interested to find when the IM and SM reach their maximum, regardless of the values of  $\lambda$  and  $\tau$ . For the IM, supposing that  $N_Q$  and  $D_j^\tau$  are fixed, if (9) reaches its maximum, (A.1) needs to be satisfied by the maximum principle of the entropy. Obviously, the necessary and sufficient condition for (A.1) is (A.2).

$$\frac{l_{j1}^{-\lambda}}{\sum_{r \in Q} l_{jr}^{-\lambda}} = \frac{l_{j2}^{-\lambda}}{\sum_{r \in Q} l_{jr}^{-\lambda}} = \dots = \frac{l_{jN_Q}^{-\lambda}}{\sum_{r \in Q} l_{jr}^{-\lambda}} \quad (\text{A.1})$$

$$l_{j1} = l_{j2} = \dots = l_{jN_Q} = 1 \quad (\text{A.2})$$

From (A.4), it is manifested that if and only if the other vertexes are all neighbors of the vertex  $v_j$ , which means  $v_j$  is located at the central position in a star graph,  $IV_j$  reaches its maximum expressed in (12).

For the SM, supposing  $N_Q$  is fixed, we firstly focus on one of the vertexes affecting the vertex  $v_k$ , say vertex  $v_j$ . If and only if (A.3) and (A.4) are satisfied,  $\Delta\rho_{j \rightarrow k}$  takes the maximum value, which indicates that only degree of  $v_j$  can be greater than or equal to 1 and the degree of other vertexes including  $v_k$  must be equal to 1. Further, we analyze the impact of  $N_Q$  vertexes on  $v_k$ . When  $\Delta\rho_{1 \rightarrow k}, \Delta\rho_{2 \rightarrow k}, \dots, \Delta\rho_{N_Q \rightarrow k}$  simultaneously reaches their maximum values, i.e., (A.3) and (A.4) are satisfied at the same time,  $SV_k$  arrives the maximum. To simultaneously satisfy (A.3) and (A.4) for all  $\Delta\rho_{j \rightarrow k}$  ( $j = 1, 2, \dots, N_Q$ ), if and only if  $D_j = 1$  ( $j = 1, 2, \dots, N_Q$ ) and  $D_k = N_Q$ , i.e.,  $v_k$  is located at the central position in a star graph,  $SV_k$  reaches its maximum value expressed in (13).

$$l_{jk}^{-\lambda} = 1 \quad (\text{A.3})$$

$$D_r = \begin{cases} D_j, & r = j \\ 1, & r \neq j \wedge r \neq k \end{cases} \quad (\text{A.4})$$

## Symbols

### Electrical Network

- L: Set of branches (i.e., lines, transformers) in a transmission network,  $= \{\dots, L_j, \dots\}$ ,  $\dim\{L\} = N_L$
- B: Set of nodes (i.e., buses) in a transmission network,  $\dim\{B\} = N_B$
- S: Critical path.  $S = \{\dots, L'_j, \dots\}$ ,  $S \subseteq L$ ,  $\dim\{S\} = N_S$
- $\alpha_j$ : Loading assessment index of the branch  $j$ ,  $L_j \in L$
- $f_j^0$ : Power flow of the branch  $j$  under normal operation,  $L_j \in L$
- $f_j^x$ : Power flow of the branch  $j$  during the contingency  $x$ ,  $L_j \in L$
- $f_j^M$ : Flow limit of the branch  $j$ ,  $L_j \in L$
- $P_d^x$ : Active load during the contingency  $x$ ,  $d \in B$
- $\delta_z^x$ : Load shedding percentage in the  $z$ th island during contingency  $x$
- $Z^x$ : Number of islands during the contingency  $x$
- $\Lambda$ : Normalized total load shedding percentage ( $0 \leq \Lambda \leq 1$ )
- $\Delta$ : Threshold for total load shedding percentage

### Correlation Graph

- V: Set of vertices in a graph,  $\dim\{V\} = N_L$
- E: Set of edges in a graph,  $\dim\{E\} = N_q$
- G: A correlation graph,  $G = \{V, E\}$
- $V^i$ : Set of vertices in critical path  $i$ ,  $V^i = \{\dots, v_j, \dots\}$ ,  $v_j = L'_j$ ,  $V^i = S^i$ ,  $\dim\{V^i\} = N_S^i$
- $E^i$ : Set of edges in critical path  $i$ ,  $E^i = \{\dots, e_q^i, \dots\}$ ,  $e_q^i = L'_j L'_{j+1}$ ,  $q = j$ ,  $\dim\{E^i\} = N_S^i - 1$
- $G^i$ : Graphic representation of critical path  $i$ ,  $G^i = \{V^i, E^i\}$
- $r$ : Power exponent of cumulative distributions
- $R^2$ : Fitting effect of the power law
- $\rho_j$ : Initial load of the vertex  $v_j$
- $D_j$ : Degree of the vertex  $v_j$
- $\tau$ : Scale factor for initial load,  $\tau > 0$
- $Q_j$ : Set of vertices affected by vertex  $v_j$ ,  $Q_j = \{\dots, v_k, \dots\}$ ,  $Q \subseteq V$ ,  $\dim\{Q_j\} = N_Q$
- $\Delta\rho_{j \rightarrow k}$ : Load variation of vertex  $v_k$  due to failure of vertex  $v_j$
- $l_{jk}$ : Distance between the vertices  $v_j$  and  $v_k$
- $\lambda$ : Portion control factor for load redistribution,  $\lambda \geq 0$
- $\eta$ : Threshold for selection of vertices into  $Q_j$
- $E_j$ : Entropy of the vertex  $v_j$

$IV_j$ : Impactability of the vertex  $v_j$   
 $SV_k$ : Susceptibility of the vertex  $v_k$   
 $\sigma(\cdot)$ : Impulse response function  
 $\psi_y$ : Percentage of remaining load when attacking  $y$  branches  
 $\kappa$ : Descent rate  
 $Y$ : Number of adjacent samples.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

All authors declare that they have no conflicts of interest.

## Acknowledgments

This research was partially funded by grants from the Key Projects of National Natural Science Foundation of China (U1734202), National Key Research and Development Plan of China (2017YFB1200802-12), and the National Natural Science Foundation of China (61703345, 51877181).

## References

- [1] A. M. L. Da Silva, J. L. Jardim, L. R. De Lima, and Z. S. Machado, "A Method for Ranking Critical Nodes in Power Networks Including Load Uncertainties," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1341–1349, 2016.
- [2] J. Yan, Y. Tang, H. He, and Y. Sun, "Cascading Failure Analysis With DC Power Flow Model and Transient Stability Analysis," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 285–297, 2015.
- [3] E. I. Bilis, W. Kröger, and C. Nan, "Performance of electric power systems under physical malicious attacks," *IEEE Systems Journal*, vol. 7, no. 4, pp. 854–865, 2013.
- [4] Å. J. Holmgren, "Using graph models to analyze the vulnerability of electric power networks," *Risk Analysis*, vol. 26, no. 4, pp. 955–969, 2006.
- [5] P. Cructti, V. Latora, and M. Marchiori, "Topological analysis of the Italian electric power grid," *Physica A*, vol. 338, pp. 92–97, 2004.
- [6] V. Rosato, S. Bologna, and F. Tiriticco, "Topological properties of high-voltage electrical transmission networks," *Electric Power Systems Research*, vol. 77, no. 2, pp. 99–105, 2007.
- [7] J. Yan, H. He, and Y. Sun, "Integrated Security Analysis on Cascading Failure in Complex Networks," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 3, pp. 451–463, 2014.
- [8] Y. Cai, Y. Cao, Y. Li, T. Huang, and B. Zhou, "Cascading failure analysis considering interaction between power grids and communication networks," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 530–538, 2016.
- [9] E. Cotilla-Sanchez, P. D. H. Hines, C. Barrows, and S. Blumsack, "Comparing the topological and electrical structure of the North American electric power infrastructure," *IEEE Systems Journal*, vol. 6, no. 4, pp. 616–626, 2012.
- [10] E. Bompard, R. Napoli, and F. Xue, "Extended topological approach for the assessment of structural vulnerability in transmission networks," *IET Generation, Transmission & Distribution*, vol. 4, no. 6, pp. 716–724, 2010.
- [11] E. Bompard, E. Pons, and D. Wu, "Extended topological metrics for the analysis of power grid vulnerability," *IEEE Systems Journal*, vol. 6, no. 3, pp. 481–487, 2012.
- [12] A. Dwivedi and X. Yu, "A maximum-flow-based complex network approach for power system vulnerability analysis," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 81–88, 2013.
- [13] J. Fang, C. Su, Z. Chen, H. Sun, and P. Lund, "Power system structural vulnerability assessment based on an improved maximum flow approach," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 777–785, 2018.
- [14] Y. Wang, X. Li, J. Li, Z. Huang, and R. Xiao, "Impact of Rapid Urbanization on Vulnerability of Land System from Complex Networks View: A Methodological Approach," *Complexity*, vol. 2018, Article ID 8561675, 18 pages, 2018.
- [15] T. Huang, S. L. Voronca, A. A. Purcarea, A. Estebsari, and E. Bompard, "Analysis of chain of events in major historic power outages," *Advances in Electrical and Computer Engineering*, vol. 14, no. 3, pp. 63–70, 2014.
- [16] S. Blumsack, E. Bompard, A. Carbone et al., "Power grids vulnerability: a complex network approach," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 19, no. 1, Article ID 013119, 2009.
- [17] X. Wei, S. Gao, T. Huang, E. Bompard, R. Pi, and T. Wang, "Complex Network Based Cascading Faults Graph for the Analysis of Transmission Network Vulnerability," *IEEE Transactions on Industrial Informatics*, 2018.
- [18] X. Wei, J. Zhao, T. Huang, and E. Bompard, "A Novel Cascading Faults Graph Based Transmission Network Vulnerability Assessment Method," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2995–3000, 2018.
- [19] Y. Zhu, J. Yan, Y. Sun, and H. He, "Revealing Cascading Failure Vulnerability in Power Grids Using Risk-Graph," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 12, pp. 3274–3284, 2014.
- [20] Y. Zhu, J. Yan, Y. Tang, Y. Sun, and H. He, "Resilience analysis of power grids under the sequential attack," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2340–2354, 2014.
- [21] F. Wenli, Z. Xuemin, M. Shengwei, H. Shaowei, W. Wei, and D. Lijie, "Vulnerable transmission line identification using ISH theory in power grids," *IET Generation, Transmission & Distribution*, vol. 12, no. 4, pp. 1014–1020, 2018.
- [22] P. D. H. Hines, I. Dobson, and P. Rezaei, "Cascading Power Outages Propagate Locally in an Influence Graph That is Not the Actual Grid Topology," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 958–967, 2017.
- [23] J. Qi, K. Sun, and S. Mei, "An interaction model for simulation and mitigation of cascading failures," *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 804–819, 2015.
- [24] P. Hines, E. Cotilla-Sanchez, and S. Blumsack, "Do topological models provide good information about electricity infrastructure vulnerability?" *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 20, no. 3, Article ID 033122, 2010.
- [25] E. Bompard, A. Monti, A. Tenconi et al., "A multi-site real-time co-simulation platform for the testing of control strategies of distributed storage and V2G in distribution networks," in



- Proceedings of the 2016 18th European Conference on Power Electronics and Applications (EPE'16 ECCE Europe)*, pp. 1–9, Karlsruhe, September 2016.
- [26] X. Wei, S. Gao, and D. Li, “Cascading fault graph for the analysis of transmission network vulnerability under different attacks,” *Proceedings of the Chinese Society for Electrical Engineering*, vol. 38, no. 2, pp. 465–474, 2018.
- [27] B. Stott, J. Jardim, and O. Alsac, “DC power flow revisited,” *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, 2009.
- [28] B. Stott, O. Alsac, and F. Alvarado, “Analytical and computational improvements in performance-index ranking algorithms for networks,” *International Journal of Electrical Power & Energy Systems*, vol. 7, no. 3, pp. 154–160, 1985.
- [29] L. Guo and X. Cai, “Degree and weighted properties of the directed China Railway Network,” *International Journal of Modern Physics C*, vol. 19, no. 12, pp. 1909–1918, 2008.
- [30] L. Kencl and J. Le Boudec, “Adaptive load sharing for network processors,” in *Proceedings of the IEEE Information Communications Conference (INFOCOM 2002)*, pp. 545–554, New York, NY, USA, 2002.
- [31] H. Wang, K. Liu, and P. Ao, “Magnetic field and specific axial load capacity of hybrid magnetic bearing,” *IEEE Transactions on Magnetics*, vol. 49, no. 8, pp. 4911–4917, 2013.
- [32] D. Duan, J. Wu, H. Deng et al., “Cascading failure model of complex networks based on tunable load redistribution,” *Systems Engineering-Theory & Practice*, vol. 33, no. 1, 2013.
- [33] W. X. Wang and G. Chen, “Universal robustness characteristic of weighted networks against cascading failure,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 77, no. 2, Article ID 026101, 2008.
- [34] D. Duan and R. Zhan, “Evolution mechanism of node importance based on the information about cascading failures in complex networks,” *Acta Physica Sinica*, vol. 63, no. 6, Article ID 068902, 2014.
- [35] L. Zhao, K. Park, and Y.-C. Lai, “Attack vulnerability of scale-free networks due to cascading breakdown,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 70, no. 3, Article ID 035101, 2004.
- [36] D.-H. Kim, B. J. Kim, and H. Jeong, “Universality class of the fiber bundle model on complex networks,” *Physical Review Letters*, vol. 94, no. 2, Article ID 025501, 2005.
- [37] J. Lehmann and J. Bernasconi, “Stochastic load-redistribution model for cascading failure propagation,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 81, no. 3, Article ID 031129, 2010.