Research Article

Node Importance Evaluation of Cyber-Physical System under Cyber-Attacks Spreading

Xin-Rui Liu, Yuan Meng, and Peng Chang

College of Information Science and Engineering, Northeastern University, Shenyang, Liaoning 110819, China

Correspondence should be addressed to Xin-Rui Liu; liuxinrui@ise.neu.edu.cn

Received 14 December 2020; Revised 24 December 2020; Accepted 31 December 2020; Published 16 January 2021

1. Introduction

In recent years, electric power systems are facing unprecedented threat from cyber-attacks due to the rapid development of the information network and the higher integration of critical infrastructure and IED (intelligent electronic device) equipment in power CPS (cyber-physical system) [1–3]. The widely deployed ICT (information and communications technology) system makes the interaction more complex among multiple systems, especially more vulnerable under cyber-attacks. The failure of a single-layer system after being attacked by a cyber-attack will spread through the interdependent network, causing fragmentation and cascading failures [4, 5] and induce a large-scale power flow abnormal transferring. The abovementioned cascading failure process and cyber-attacks will eventually lead to a blackout that led to the collapse of the power system such as large blackouts which occurred in North America in 2003 [6], Rome in 2004 [7], and Ukraine in 2015 [8].

A large number of studies have shown that the power grid has small-world effect and scale-free property [9]. It shows strong robustness under random attacks but very fragile to deliberate attacks. Wang et al. in [10, 11] proposed methods to improve the stability of the power system. Albert et al. in [12] found that the power grid can maintain stable under most disturbances, but when the key power nodes are attacked, the synchronization ability of the grid will be greatly reduced. Therefore, identifying and assessing important nodes in the power grid and performing prevention and control are of great significance for improving the reliability and survivability of the power system. The current node importance evaluation methods include evaluation methods based on local information such as K-shell decomposition method [13], based on node path such as closeness centrality [14], based on feature vectors such as PageRank algorithm [15], and based on node removal [16] and contraction [17]. The authors in [17] applied the node contraction method to the power grid and verified the feasibility of identifying important nodes based on the topological structure of the power grid. A comprehensive evaluation index that takes into account both electrical characteristics and topological structure characteristics is proposed to identify important nodes in the power grid [18].

Based on the node link strength defined by power flow tracking, the author in [19] identifies important nodes in the power system from the perspective of global energy
transmission. In [20], the authors analyze the benefits, losses, costs, and other factors of network attacks and use dynamic Bayesian networks to comprehensively evaluate the attack effects of network nodes. In [21], the improved threat propagation tree is used to evaluate the situation and a CPPS security situation assessment model that considers threat propagation is proposed. The above node importance evaluation method only focuses on the characteristics of a single-layer network and does not consider the method of identifying key nodes of the power system under the interdependent network.

About interdependent networks, Buldyrev first analyzed the cascading failure in 2010 and proposed a cascading failure model based on network topology [22]. The author in [23] proposed electrical distance and node electrical coupling connectivity metric to identify key nodes in complex power grids. The average load balance of adjacent nodes and the parameter of network load rate are combined to measure the impact of the disabled node on the network and judge the importance of nodes in the interdependent network [24]. The above node importance evaluation considers the impact of disabled nodes on the interdependent network from different aspects but does not effectively analyze the composite value of nodes in the power CPS under the cyber-attack propagation scenario.

Whether the above node importance evaluation is based on a single-layer network or an interdependent network, the importance value is only a fixed value calculated based on a certain characteristic of the system. Without considering the potential dangers brought by the spread of cyber-attacks, such fixed indexes cannot meet the system’s need to distinguish important nodes under the attack spread scenarios. In this paper, the analysis of potential dangers caused by the spread of cyber-attacks is shown in Figure 1. Suppose that at a certain time during the interval between two detections, a certain device in the information network is attacked by a cyber-attack (such as worms, and Trojan Horse).

(1) First, before the attacked device is successfully detected by the power CPS and the countermeasures are executed, the spread of the cyber-attack is in the first stage. In this stage, the attack may have spread from the initially attacked device to the remaining devices that have topological connections or information interactions with it. Take the information layer network as an example: the data detection frequency cannot match the information exchange rate, which will cause the cyber-attack to be spread to the new information equipment due to the information exchange before the next detection. If the attack makes the transmitted device in exposed state (the device has been successfully attacked, but the attacker did not perform any attack operations), then the power system cannot successfully detect such exposed device in the next detection.

(2) Subsequently, the power system detects the initial attacked device and executes countermeasures. Although the current detection failure rate can meet the expectations of the dispatch center, it cannot be reduced to zero. Therefore, the response strategy made by the dispatch center can only reduce and block the harm caused by cyber-attacks to a certain extent. And, this strategy does not consider the cyber-attack propagation described in (1). In fact, after the first propagation stage, there are most likely exposed nodes in the power system.

(3) The period from the execution of the strategy to the next detection, the spread of cyber-attacks has reached the second stage. Although the previous strategies dealt with the initially attacked device, the exposed nodes formed in the first propagation stage will continue deepen the scope of cyber-attacks over time.

(4) After the above two propagation stages, most devices have the potential to be attacked. This potential risk may expose the power system to great danger during the interval between two detections.

In order to improve the proposed node importance evaluation method of power CPS and effectively reduce the potential risks caused by the cyber-attack propagation between the detection gap, in this paper, from the information network and cyber-physical interdependent network two aspects, we propose a node importance evaluation method suitable for cyber-attack scenarios. The main contributions of this paper are as follows:

(1) At the information network level, considering the propagation of cyber-attacks among information devices, a compound centrality index and its calculation method of nodes based on adaptive coefficients are proposed. This index improves the one-sidedness of the existing centrality indexes, and the time-varying calculation result better matches the importance of the power CPS node under the cyber-attack propagation scenario.

(2) Establish a cyber-physical interdependent network model to analyze the cascading failure behavior characteristics of the system and give a quantitative analysis of the degree of power grid decoupling under different attack scenarios. Calculate the degree of decoupling caused by cascading failures which caused by exposed nodes under the spread of cyber-attacks and comprehensively reflect the importance of different nodes in power CPS.

In the second section, we proposed the node risk and established a network attack propagation model. The third section improves closeness centrality and proposes a compound centrality algorithm based on adaptive coefficients. In the fourth section, the power system cyber-physical interdependent network is established to quantitatively calculate the degree of cascading failure that the exposed node formed by the spread of cyber-attacks may cause. In the fifth section, a case study is given to prove the advantages of our proposed algorithm and report the conclusion.
2. Cyber-Attack Propagation Model

Cyber-attacks occur more commonly in power communication networks due to the deep integration of communication and measurement devices such as phasor measurement units (PMUs). Attackers use new viruses and worms with autonomous propagation capabilities to falsify or control the resources of the information layer to endanger the security and reliability of the power grid. Therefore, in this section, establish a cyber-attack propagation model for power CPS information layer devices which can apply to all equipped information devices with communication capabilities, such as information acquisition system SCADA (Supervisory Control and Data Acquisition), energy management system EMS (Energy Management System), WAMS (Wide Area Measurement System), and so on.

Since SCADA, EMS, and WAMS have a certain time interval in collecting, monitoring, processing information, and transmitting instructions, while considering the characteristics of information systems based on discrete events and the propagation mechanism of new viruses such as worms, the unit propagation time of the attack \( \tau \) is defined as the data interaction interval of power CPS. According to the propagation stage of the cyber-attack, the information equipment is divided into 5 states. A single device can only be in one state within a unit of propagation time \( \tau \). Under certain conditions, the current state can be transformed into other states. The situation transformation process is shown in Figure 2.

1. **Infected.** The device has been successfully infected by a cyber-attack and has been manipulated by the attacker to perform some actions (for example, tampering with operating data and modifying switch status). It has the ability to infect other devices. The detection system can detect the device.

2. **Exposed.** The device successfully infected by a cyber-attack, but the attacker has not performed any attack operations. It has the ability to infect other devices. The detection system cannot find the device.

3. **Susceptible.** The device directly connected to the infected or exposed device in the communication topology. There are security vulnerabilities that can be infected by malicious code such as worms but have not been infected yet and do not have the ability to infect other devices.

4. **Isolated.** The device that has been attacked is automatically isolated by the detection software or manually by the system operator. No longer have hardware connections or data interactions with other devices. No longer has the ability to infect other devices.

5. **Normal.** Devices are in a normal state of untouched cyber-attacks.

2.1. Probability of Propagation. In this section, we calculate the cyber-attack propagation probability between pairs of individuals who have a connection in the communication topology. Consider a pair of devices which are connected, one of which is infective or exposed and the other is susceptible. Attack spreads through them by contact from \( i \) to \( j \). Then, the rate that \( j \) not being attacked in the continuous time system is as follows [25]:

\[
1 - y_{ij} = \lim_{\delta t \to 0} \left(1 - r_{ij} \delta t \right)^{1/\delta t} = e^{-r_{ij} \tau},
\]

where \( r_{ij} \) is the probability that the cyber-attack successfully spreads from infective or exposed device \( i \) to susceptible device \( j \) under the self-protection mechanism of the device and information system. The infective device \( i \) remains infective for a time \( \tau_i \) that \( i \) has not been detected and disabled by the detection system.

Due to the cyber systems based on discrete events, use discrete time-steps rather than continuous time, in which case instead of taking the limit in equation (1) we simply set \( \delta t = 1 \), giving

\[
y_{ij} = 1 - \left( 1 - r_{ij} \right)^\tau,
\]

where \( \tau \) is the average detection time of the online detection system that is measured in data transmission time-steps.

2.2. Propagation Model. For modeling the attack propagation clearly, we abstract different types of communication devices in the information layer as nodes. According to the
communication topology of the information network, the connection matrix $C = c_{ij}$ is defined. The matrix elements $c_{ij}$ are binary variables which equals 1 if $i$ is kept connected to $j$ in communication and 0 otherwise. We define the probability that a node may be attacked and turn into exposed state as the node risk $P(\cdot)$. Let us assume at the initial time $t_0$ between two detection gaps, one or several nodes in the CPS information layer are compromised by cyber-attacks and turn into infected while the remaining nodes are uncompromised. In the following period of time, the remaining nodes will turn into exposed with a certain probability, accompanied by a time-varying node risk.

We define $\psi = \{1, 2, \ldots, M\}$ as the set of information nodes; $\eta = \{\ldots\}$ is the set of infected nodes; and $T = n\bar{t}$ is the time when the attack spreads through $n$ units of propagation time $\bar{t}$. The attack spreads in stages, and the risk of each node changes with time as follows [26]:

1. At the initial time $t_0$, the risk of infected and other nodes is as follows:

$$P(\alpha_i^{t_0} = 1) = 1, \quad \forall i \in \eta,$$
$$P(\alpha_i^{t_0} = 1) = 0, \quad \forall i \notin \eta,$$  \hspace{1cm} (3)

where $\alpha_i^t$ is a binary variable which equals 1 ($\alpha_i^t = 1$) if the node $i$ has the possibility of being attacked at time $t_0$, and $\alpha_i^t = 0$; otherwise, $P(\alpha_i^{t_0} = 1)$ is the node risk of $i$ at time $t_0$.

2. By time $t_0 + \bar{t}$, cyber-attacks may spread from infected nodes to susceptible nodes via routers and communication connections, turning susceptible nodes into exposed nodes, and the detection software has not detected them yet. The risk of each node is given by the following equation:

$$P(\alpha_i^{t_0+\bar{t}} = 1) = 1 - \prod_{j\in\eta} (1 - \gamma_{ji}), \quad \forall i \notin \eta,$$ \hspace{1cm} (4)

where $\gamma_{ji}$ is the probability that the attack propagates from node $i$ to node $j$ during the time $\bar{t}$, and it is given in equation (2).

3. By time $t_0 + 2\bar{t}$, the range and volume of cyber-attacks further expand over time, and the risk of exposed nodes deepens. The risk of each node is an iterative process over time. According to equations (4) and (5), the node risk by time $t_0 + 2\bar{t}$ is given by the following equation:

$$P(\alpha_i^{t_0+2\bar{t}} = 1) = 1 - \prod_{j\in\eta} (1 - \gamma_{ji}) \times P(\alpha_i^{t_0+\bar{t}} = 1) \times P(\alpha_i^{t_0} = 1) + P(\alpha_i^{t_0} = 1|\alpha_i^{t_0+\bar{t}} = 0) \times P(\alpha_i^{t_0+\bar{t}} = 0), \quad \forall i \notin \eta,$$ \hspace{1cm} (7)
where \( P(\alpha_i^{t+2\tau} = 1|\alpha_i^{t+\tau} = 0) \) is the probability that the node is not attacked at time \( t_0 + 2\tau \) but is attacked at time \( t_0 + \tau \). In equation (7), only \( P(\alpha_i^{t+2\tau} = 1|\alpha_i^{t+\tau} = 0) \) is unknown during the iteration process. This variable is discussed as follows:

\[
P(\alpha_i^{t+2\tau} = 1|\alpha_i^{t+\tau} = 0) = 1 - P(\alpha_i^{t+2\tau} = 0|\alpha_i^{t+\tau} = 0) \\
= 1 - \prod_{i,j \in \eta \in j} \left( 1 - P(\alpha_j^{t+\tau} = 1) \times y_{ji} \right) \quad \forall i, j \in \psi.
\]

Notice. One of the ways the power system responds to cyber-attacks is as follows. After successfully detecting the cyber-attack, the power system will disable infected nodes automatically by the detection software or manually by the system operator. These nodes are no longer connected to the information network (no more data interaction with other nodes), while the system no longer trusts its uploaded data. Namely, this type of node turns into the isolated state. Then, they may be recovered to normal by the system and reconnect to the network, but in the propagation discussed in this paper, such nodes no longer participate in the propagation of cyber-attacks, that is, they cannot be reinfected by cyber-attacks.

3. Node Centrality Algorithm Based on Adaptive Coefficient

The connection form of the node and its position in the topological structure have a vital influence on the promotion or interruption of the spread of cyber-attacks [4]. At present, a variety of centrality algorithms have been applied to node importance evaluation. In this section, we propose a node importance evaluation algorithm that considers cyber-attacks propagation and the potential threats it may cause in the detection gap, which is more flexible and more suitable for attack scenarios.

3.1. Centrality Analysis. According to the connection form of the node and its position in the topology, common node centrality algorithms include degree centrality [27], closeness centrality [28], betweenness centrality [28], and so on, all describe the importance of nodes in the network from different respects, and they are given as follows:

\[
DC(i) = \sum_{j=1, j \neq i}^{N} c_{ij},
\]

\[
CC(i) = \frac{N - 1}{\sum_{j=1, j \neq i}^{N} \text{dis}_{ij}},
\]

\[
BC(i) = \sum_{s=1}^{N} \sum_{t=1}^{N} \frac{\sigma_{st}(i)}{\sigma_{st}},
\]

where \( N \) is the number of nodes in the network, \( c_{ij} \) is a binary decision variable that judges the topological structure or information interaction between nodes \( i \) and \( j \); \( \text{dis}_{ij} \) is the shortest path between nodes \( i \) and \( j \); in this paper, we use the minimum number of nodes in the path from \( i \) to \( j \) as the value of the shortest path. \( \sigma_{st} \) is the number of shortest paths from \( s \) to \( t \); \( \sigma_{st}(i) \) is the number of shortest paths from \( s \) to \( t \) which passing by node \( i \).

Degree centrality indicates the sum of the number of nodes directly connected to the designated node. Closeness centrality is the reciprocal of the average shortest path from the designated node to all other reachable nodes. In general,
the closer a node is to other nodes, the greater its closeness centrality. Conversely, the smaller the node closeness centrality, the more the node is at the edge of the network. Betweenness centrality is the number of times that the designated node is located on the shortest path between any two other nodes.

The three algorithms describe the importance of nodes from the perspectives of local characteristics, global characteristics, and propagation characteristics, while they also have certain limitations. Degree centrality can only one-sidedly reflect the closeness between the designated node and its surrounding nodes. For example, the connection node between two partitioned networks has a smaller degree of centrality but a higher degree of importance. Or for some nodes with high degree centrality, the clustering network it embeds may be at the edge of the system. At this time, the degree centrality is relatively high but the degree of importance is average. However, as the nodes in a complex network, especially in large-node systems such as power systems, the position of the node in the system and some of its functions are far more important than the number of nodes around it. The descending order curve of the normalized values of DC, CC, and BC of each node in the IEEE14 system is shown in Figure 3. The comparative analysis is as follows:

1. The relative DC values of all nodes are arranged in descending order as shown in. It descends in steps, and there are 3 steps and some steps are larger in width. It descends in steps, and there are 3 steps and some steps are larger in width. For example, the DC of 6 nodes in the graph are all 0.4. The importance of these nodes in the same step cannot be distinguished.

2. The descending arrangement curve of relative CC is relatively smooth, but there is little difference among the values. The variable interval of the value is 0.585–1, which is too narrow.

3. The relative BC descending order curve is also a smoother curve with values distributed in the interval 0–1. However, there is a platform segment with a value of 0, which accounts for 28.57% of the total system nodes, at its end. The node in this segment cannot effectively distinguish the importance.

3.2. Centrality Algorithm Based on Adaptive Coefficient. According to the analysis results of the three centrality algorithms in 3.1, we can get the following. The three algorithms have their respective advantages and disadvantages. If only one centrality algorithm is selected as the basis for judging the importance of nodes, the sorting result is not appropriate and accurate. These centrality algorithms make it impossible to effectively distinguish the importance of each node in the system under a specific attack scenario based on the fixed value calculated by the topology structure. Moreover, if the importance of nodes is only distinguished based on the node risk $P$ that proposed in 2.2, the position of the node in the system is ignored, and the mutual influence between nodes in the process of attack propagation is discarded. Therefore, after comprehensively considering the global attributes of CC and the propagation attributes of BC, this section improves the CC algorithm according to the propagation characteristics of cyber-attacks and proposes a compound centrality algorithm that considers both improved CC and BC to weaken the differences and shortcomings of existing centrality algorithms.

Improved closeness centrality algorithm:

$$CC(i) = \sum_{j \in \theta} \text{dis}_{ij},$$

(15)

where $\text{dis}_{ij}$ is the shortest path from $j$ to $i$, $j$ is the node initially infected by the cyber-attack, and $i$ is the remaining nodes in the system.

The compound centrality algorithm proposed in this paper performs a weighted summation of the two indexes of improved CC and BC, but the two are not completely independent indexes. For example, nodes with a high improved CC will have a higher BC under a certain probability. Summing the two indexes with a weight of 0.5, reasons such as duplication of information and redundant factors will cause the calculated composite index to be inaccurate. To solve this problem, this section proposes an adaptive coefficient as the weight of the two centralities.

In the propagation of the power CPS cyber-attack, the propagation probability between nodes $y_{ij}$ as a key factor, has always been defined as a fixed value in previous studies. Considering the differences in the importance of each node in the network, a correction equation is proposed to modify $y_{ij}$ to a certain extent:

$$y_{ij}^* = y_{ij} - X \cdot C_j,$$

(16)

where $X$ is the correction coefficient and $C_j$ is the centrality index of node $j$. The $y_{ij}^*$ of node $j$ changes with the centrality index $C_j$. The higher the $C_j$, the lower the $y_{ij}$. That is, when $C_j$ increases, the probability of a cyber-attack spreading from node $i$ to node $j$ decreases.

The iterative solution process of the adaptive coefficient $\delta$ is given by equations (17)–(23). Equation (22) is the termination condition of the iteration, and equation (23) is the objective function. The appropriate initial values of $\delta_1$ and $\delta_2$ are selected and substituted into formula (17) to start the iteration. In equations (18) and (19), the composite centrality index $C$ of the node and the propagation probability $y_{ij}$ are updated with the update of $\delta$. According to the risk model proposed in 2.2, use the iteratively updated $y_{ij}^*$ to calculate the risk of each node and the average of the overall node risk, which is shown in equation (21). Substituting the average of the overall node risk in the current and previous iterations into equation (17) again will start a new round of iterative correction process. Until $\delta$ satisfies the iteration termination condition $\delta_k - \delta_{k-1} \leq 10^{-5}$, the iteration process stops and jumps out of the iteration. Take the minimum the average of the overall node risk in all rounds of iterations as the objective function and find the optimal solution that satisfies the objective function in the iteration process. That is, at the target time $t = T$ ($T$ is an integer multiple of $\tau$), after correcting with the optimal $\delta$, the average of the overall node
risk in the system is the smallest. This optimal solution is the adaptive coefficient of this paper. It can be seen from the above algorithm flow that if we want to start the iterative process of $\bar{\delta}$, we need to give the first two initial values $\bar{\delta}_1$ and $\bar{\delta}_2$ of $\bar{\delta}$. The rationality of the initial value selection and its influence on the number of iterations and the judgment of the optimal solution are discussed in the case study in Section 5.

$$\bar{\delta}_k = 1 + \frac{P_{t_{k-1}}^t - tP_{t_{k-2}}^t}{P_{t_{k-2}}^t - tP_{t_{k-2}}^t}$$

$$C_i^k = \frac{1}{\bar{\delta}_k}bc_i + \left(1 - \frac{1}{\bar{\delta}_k}\right)\alpha_{i'}^c,$$

$$y_{ij}^k = y_{ij} - X \ast C_i^k,$$

$$P\left(\alpha_{i'}^{t_{n-m}} = 1\right) = P\left(\alpha_{i'}^{t_{n-1}} = 1\right) + \left(1 - \prod_{i,j \notin \eta} \left(1 - P\left(\alpha_{j'}^{t_{n-m}} = 1\right) \times y_{ii}\right)\right) \times \left(1 - P\left(\alpha_{i'}^{t_{n-m}} = 1\right)\right), \quad \forall i, j \in \psi,$$

$$\bar{P}_i^t = \frac{\sum_{i} P_{t_{k}}^t |k}}{N},$$

$$\bar{\delta}_k - \bar{\delta}_{k-1} \leq 10^{-5},$$

$$f = \min_k \bar{P}_k^t,$$

where $\bar{\delta}_k$ is the adaptive coefficient of the kth iteration, $C_i^k$ is the compound centrality of node $i$ when $\bar{\delta} = \bar{\delta}_k$, $y_{ij}^k$ is the probability of successful propagation of the attack from $i$ to $j$ modified according to $C_i^k$, $P_{t_{k}}^t |k$ is when $y_{ij} = y_{ij}^k$, the risk of node $i$ at time $T$, $P_{t_{k}}^t |k$ is when $y_{ij} = y_{ij}^k$, the average risk of all nodes in the system at time $T$, and $f$ is the decision function of $\bar{\delta}$. A certain $\bar{\delta}_k$ in the iterative process minimizes the average risk $P_{t_{k}}^t |k$ of nodes in the system at the target time $T$.

**Figure 3:** Central index descending order of the IEEE14 system.

4. Cascading Failure of Power Interdependent Network

4.1. Interdependent Network Model. Electric power CPS is a multidimensional heterogeneous system that deeply embeds perception, information processing, and control platforms into the power system to meet real-time monitoring and achieve command-driven of the power system. The power network provides power for the information network, and the control and analysis module in information network reversely drives the power network. The interdependence between the two networks enables the power system to be modeled as a power cyber-physical interdependent network. Based on the "undirected" and "disordered" characteristics of general types of cyber-attacks (nondirected attacks) spreading in information systems, both the power network and the information network are equivalent to nonweighted undirected networks. Combining the actual situation of China's power system [5], only considering the characteristics of interconnection between nodes, using complex network theory to simplify the power and information network is as follows:

(1) Ignore the functional differences among the plants and stations, and regard the power generation nodes and substation nodes, dispatching nodes and routing nodes in the information network as equivalent nodes, regardless the difference in the types,
quantities, and deployment modes of devices in various sites.

(2) Ignoring the differences in information protocols and hierarchical structures among information nodes at all levels, it is considered that the lines between nodes can transmit bidirectional information, and multiple information lines in the same direction are combined to eliminate multiple edges and self-loops.

Use \( G_p = (E_p, V_p) \) and \( G_c = (E_c, V_c) \) to represent a power network with \( n \) nodes and \( k \) branches and an information network with \( m \) nodes and \( g \) branches, where \( E = \{ e_{ij} \} \) is the set of edges and \( v = \{ v_1, v_2, \ldots, v_n \} \) is the set of nodes in the network, respectively. Each node in the power network \( G_p \) is connected to an information node in \( G_c \). The power node provides electrical energy support to the information node, and the information node receives the status information sent by the power node and feeds back control instructions. Naturally, each node in \( G_p \) connects and depends on the corresponding node in \( G_c \), and vice versa. In addition, according to the important status of the scheduling nodes, it is set as an autonomous node independent of the power grid. It deploys a complete backup power supply and power generation equipment, which is not affected by power grid energy fluctuations.

China’s power line information network is a dedicated resource for the power system, and most of the communication lines are laid along with high-voltage transmission lines. The geographical similarity of the two layouts makes the topological structure between the information network and the power network highly similar. On the other hand, in order to meet the needs of control and dispatch, the information network also has dispatching nodes. So, the information network has more stations than the power network. At the same time, the optical information network needs to be formed into a ring to protect its self-healing ability, and the structure of the information network is more complicated [29]. Therefore, the dependent network model selects the “part-to-one correspondence” coupling mode, and the established power cyber-physical dependent network model is shown in Figure 4.

Different from the single-layer network with only connectivity link, there are two types of edges in the interdependent network: the connectivity link and the dependency link. The nodes in the single-layer network rely on the internal connectivity link (the black solid lines in Figure 4) to achieve corresponding functions. For example, the power generation nodes, substation nodes, and load nodes in the power network realize the generation, transmission, and consumption of electric energy through the transmission line. The topological structure of the information network is more complicated than that of the power network. In Figure 5, there are connecting links that do not exist in the power network between the information node 2 and the information node 8. The dependency link between the two networks (the red dotted line in Figure 4) is used as a medium for energy or information exchange to realize the mutual influence between the two-layer networks. The dependency matrix \( D_{P-C}, D_{C-P} \) is defined according to the connection relationship of the power cyber-physical dependent network, and the matrix element \( d_{ij} \) uses logic elements of “1” and “0” to indicate whether there is a dependency edge between the power and information nodes.

In summary, establish a power cyber-physical interdependence network that includes power grids, information networks, and the interdependency links \( G_p = G_c = \mathbf{R} (G_{P}, G_{C}, D_{P-C}, D_{C-P}) \), where \( G = (E, V), V = \{v_1, v_2, \ldots, v_n\} \) is the set of nodes in a single-layer network, \( E = \{e_{ij}\} \) is the set of connectivity links in a single-layer network, and \( D_{P-C}, D_{C-P} \) is the set of the dependency links between the power cyber-physical dependent network.

4.2. Cascading Failure Behavior. According to the operating principle of the dependent network, the failure of the power node or information node will lead to the failure of the dependent node in the other network. The power network or information network will be broken into several fragmented networks, and the scope of the attack will expand with the expansion of the network fragmentation and eventually cause cascading failures in the dependent network [29]. Describe the cascading failure behavior caused by the attack as follows:

(1) When one or some information nodes (power nodes) in the information network \( G_c \) (or power network \( G_p \) ) are attacked and fail, the connectivity links and the dependency links on these nodes fail too;

(2) Corresponding nodes in the power network \( G_p \) fail due to the interdependence with the failure node in \( G_c \). Corresponding, the connectivity links and the dependency links on the nodes also fail. After removing all the abovementioned faulty nodes, the connectivity links, and the dependency links, the power and information network is decomposed into several fragmented networks.

(3) (On the basis of 2), identify the connected subgraph in \( G_p \) and \( G_c \), and judge the nodes that do not belong to the connected subgraph as failed nodes. Remove the newly determined failed node and its connectivity links and dependency links. Based on this process, the system will reach stability after a certain number of propagations of the failure in the dependent network.

(4) Identify the set of nodes in the maximal connected subgraph in the stable network, and finally determine the stable state of the decoupled network.

Notice. According to the main research content of this paper, the information layer network attack and its propagation mechanism demand for network connectivity use “the set of nodes in the maximal connected subgraph” [29] to judge the stable state of the dependent network. The set element is the power-cyber node group \( v_p - v_c \), and the dependency links between \( v_p \) and \( v_c \), where \( v_p \) belongs to the maximal.
connected subgraph in the power network and $v_C$ also belongs to the maximal connected subgraph in the information network. In order to ensure the effectiveness of connectivity, after the fragmentation of the network, only the nodes in the set of nodes in the maximal connected subgraph are available, and the remaining nodes fail.

According to the cascading failure behavior, taking the IEEE14 node as an example, the failure decoupling process of the dependent network is shown in Figure 6 when the initial attack node is 4 and 11. The upper network in the figure is the power system $G_P$, and the lower network is the information system $G_C$. Figure 6(a) shows that the
connectivity of $G_C$ is similar to that of $G_P$, but compared with $G_P$, $G_C$ has one more independent green "scheduling node" and green connectivity links between the scheduling nodes and each information node, and three more green connectivity links between 4–14, 3–11, and 5–12. In Figure 6(b), the nodes 4 and 11 failed due to cyber-attacks and turned to red. Correspondingly, the connectivity link connected to it fails and turns to a dashed line, and the dependency link fails and turns to a red line, and the corresponding dependent nodes and connectivity links in $G_P$ are treated in the same way. In Figure 6(c), we delete the failed node and the failed edge in Figure 6(b). The remaining nodes that do not belong to the connected subgraph are turned into blue, and the connecting edges on these nodes turn to blue dashed lines. At this time, it can be seen that both networks are fragmented since 5–12 connecting edge in $G_C$, and the number of nodes in the connected subgraph is greater than that in $G_P$, and the degree of fragmentation is not as serious as $G_P$. Figure 7(d) filters out the set of nodes in the maximal connected subgraph from Figure 7(c). At this time, the power interdependent network finally reaches a stable state.

5. Experimental Results

We use the connection relationship of the IEEE57-bus test system to simulate the information layer network connection relationship, ignoring the weight of each link; it is shown in Figure 5.

5.1. Compound Centrality Algorithm

Notice. In the information layer cyber-attack propagation model described in Section 1, the initial time $t_0$ should be the time when the cyber-attack occurred. However, at present, the power system cannot accurately analyze which time the cyber-attack occurred between the two detection gaps. Therefore, in the experiment, the moment when the cyber-attack is detected is regarded as the initial moment in the risk $P(\cdot)$ calculation. And, set the power system to make a response strategy to disable the attacked device at time $t + \bar{t}$.

Establish the attack scenario with attack the information device equipped on node 18. The BC and improved CC
indexes of the remaining nodes calculated according to formulas (13) and (14) are shown in Table 1, and the two indexes are ranked according to importance. It can be seen from the table:

1. The improved CC index mainly serves the attack scenario. The closer the node to the initial infected node, the lower the node’s improved CC value. At the same time, the indexes have a certain degree of repeatability; for example, the value of the improved CC index for 7 nodes is equal to 4.

2. The importance ranking results of BC and improved CC are obviously different. Some nodes have higher improved CC and lower BC, such as node 9. Some nodes have lower improved CC and higher BC, such as node 37.

Perform log normalization processing \( x_{ij} = \frac{\log_{10}(x)}{\log_{10}(\max(x))} \) on the BC and improved CC indexes in Table 1, and use them as the input data to execute the compound centrality algorithm proposed in this paper. The iterative results \( \varphi \) and iteration times \( N \) of the adaptive coefficients under different correction coefficients \( X \) and different target times \( T \) are shown in Table 2.

According to the iterative results in Table 2, it can be seen that in the initial stage of attack propagation, there is a certain proportion of improved CC in the compound centrality. That is, in the initial stage of propagation, mainly modifying \( y_{ij} \) of node \( j \) which is near the infected node has a better effect on reducing the overall risk. However, as the attack spreads, the proportion of BC in the compound centrality gradually increases. That is, the propagation characteristics possessed by nodes in the topological structure in the later stage of propagation have a more obvious effect on reducing risk which caused by the spread of the attack. In addition, it can be seen that the iteration times of this algorithm are less than 10, and the convergence speed of the algorithm is fast, which can meet the demand for timeliness of power system scheduling. And, we verified that the selection of the initial value of \( \varphi \) has no effect on the convergence value of the iteration and the optimal solution. It has a slight impact on the iteration times, but the number of iterations can also be guaranteed within 10 times.

In the case of \( X = 0.03, T = 15 \) and \( X = 0.03, T = 30 \), using compound centrality, improved CC, BC, and DC to modify \( y \), the risk of the remaining nodes of the system is shown in Figures 7 and 8.

Figure 7 shows the following:

1. In the initial stage of attack propagation, the risk of several nodes that are topologically close to the initial infection node 18 rises fastest (for example, node 4, 19, 5, 6, 20)

2. In the initial stage of attack propagation, in view of the strong correlation between the improved CC and
Table 2: Iteration results of the adaptive coefficients under the IEEE57 attacking the 18th node.

| δ   | T = 12t | T = 15t | T = 18t | T = 21t | T = 24t | T = 27t | T = 30t |
|-----|---------|---------|---------|---------|---------|---------|---------|
| 0.03| 1.479   | 1.433   | 1.389   | 1.346   | 1.306   | 1.268   | 1.234   |
| N = 6|        |         |         |         |         |         |         |
| 0.04| 1.429   | 1.385   | 1.342   | 1.301   | 1.263   | 1.227   | 1.194   |
| N = 7|        |         |         |         |         |         |         |
| 0.05| 1.330   | 1.292   | 1.256   | 1.221   | 1.189   | 1.159   | 1.132   |
| N = 6|        |         |         |         |         |         |         |

Figure 7: Risk of node under compound C, improved CC, BC, and DC correction at T=15.

Figure 8: Risk of node under compound C, improved CC, BC, and DC correction at T=30.
the cyber-attack scenario, the minimum risk after the use of the improved CC indicator to modify γ is slightly better than the effect of using the composite centrality indicator to modify.

Figure 8 shows the following:

(1) Node 4 is located in the direct connection position of the infected node 18, and node 6 is in a special position in the connection relationship (with a high BC and DC index). When the conventional centrality is used to modify c, because the index does not consider the characteristics of each attack scenario, the risk of node 4 and node 6 rises faster, which cannot effectively reduce the risk in the system. Note that the correction effect of the BC is better than that of the DC; in other words, during the attack propagation process, the propagation attribute of the node is more important than the local attribute.

(2) When using the improved CC and compound centrality to modify, since the characteristics of the attack scenario is considered in the index, the attack spread in the system is effectively curbed and the risk of node 4 and node 6 is reduced. Note that the improved CC only considers the actual situation of the attack, which is really effective in reducing the risk of node 4 and node 19 which are close to infected node 18, but does not consider the role of nodes in the subsequent attack propagation. Therefore, the effect of risk reduction for other nodes is worse than that of the compound centrality index.

Figures 7 and 8 show the following:

(1) As the attack spreads, the compound centrality is used to modify γ, which effectively reduces the impact of the attack on the 19th node with a higher initial risk.

(2) Although in the initial stage of attack propagation, the effect of using the improved CC to modify γ is slightly better than that of the compound centrality modification; considering the overall attack propagation, the effect of compound centrality correction γ is better.

In summary, the compound centrality considers both the characteristics of attack scenarios and the role of nodes in

Table 3: Compound centrality and S value of node under the IEEE57 attacking the 18th node.

| Node | Compound C | S    | Node | Compound C | S    |
|------|------------|------|------|------------|------|
| 1    | 0.481938925| 0.912280702| 30   | 0.310493  | 0.894736842|
| 2    | 0.485281429| 0.964912281| 31   | 0.260317  | 0.877192982|
| 3    | 0.727546619| 0.947368421| 32   | 0.267126  | 0.877192982|
| 4    | 0.932043014| 0.877192982| 33   | 0      | 0.982456144|
| 5    | 0.480491466| 0.964912281| 34   | 0.303463  | 0.877192982|
| 6    | 0.727907074| 0.947368421| 35   | 0.351337  | 0.877192982|
| 7    | 0.629168326| 0.964912281| 36   | 0.408994  | 0.877192982|
| 8    | 0.64029582 | 0.964912281| 37   | 0.460605  | 0.929824561|
| 9    | 0.594675807| 0.894736842| 38   | 0.546372  | 0.929824561|
| 10   | 0.409535734| 0.929824561| 39   | 0.308133  | 0.947368421|
| 11   | 0.50441267| 0.947368421| 40   | 0.309953  | 0.964912281|
| 12   | 0.466716998| 0.929824561| 41   | 0.436589  | 0.929824561|
| 13   | 0.602369165| 0.964912281| 42   | 0.127394  | 0.964912281|
| 14   | 0.489843246| 0.929824561| 43   | 0.170842  | 0.964912281|
| 15   | 0.644330887| 0.929824561| 44   | 0.435195  | 0.947368421|
| 16   | 0.321392772| 0.964912281| 45   | 0.494724  | 0.947368421|
| 17   | 0.321392772| 0.964912281| 46   | 0.376085  | 0.947368421|
| 18   | 0.213929721| 0.92820702  | 47   | 0.331809  | 0.947368421|
| 19   | 0.878290586| 0.92820702  | 48   | 0.381024  | 0.929824561|
| 20   | 0.694218585| 0.92820702  | 49   | 0.526516  | 0.964912281|
| 21   | 0.59712313 | 0.92820702  | 50   | 0.349656  | 0.964912281|
| 22   | 0.58228702 | 0.877192982| 51   | 0.307274  | 0.947368421|
| 23   | 0.49305805 | 0.964912281| 52   | 0.399481  | 0.912280702|
| 24   | 0.450285109| 0.789473684| 53   | 0.342019  | 0.912280702|
| 25   | 0.370397806| 0.877192982| 54   | 0.377108  | 0.912280702|
| 26   | 0.365271476| 0.929824561| 55   | 0.458715  | 0.912280702|
| 27   | 0.405152159| 0.929824561| 56   | 0.379537  | 0.894736842|
| 28   | 0.46248975  | 0.929824561| 57   | 0.24699  | 0.964912281|
| 29   | 0.549598151| 0.842105263| 58   | 0.379537  | 0.894736842|
attack propagation, and using it to modify \( \gamma \) can effectively reduce the risk of the remaining nodes in the system. The ranking result of compound centrality is more suitable for identification requirements of key nodes in the power system under cyber-attack scenarios. In the case of \( X = 0.03 \) and \( T = 15f \), the importance ranking of the remaining nodes in the system is shown in Figure 9 when the compound centrality, improved CC, BC, and DC indexes are used to modify \( \gamma \).

5.2. Cascading Failure Analysis. The node importance evaluation in this paper focuses on the distinction of the importance of a single node. Therefore, according to the cascading failure model established in Section 3, each time a single node is selected for attack and the maximal connected subgraph and the node survival rate \( S \) under the stable state are obtained. Analyze the difference of decoupling results when attacking different single nodes. The experiment uses the IEEE57 system "one-to-one correspondence" coupling method. Sequentially attacking a single node of the information network, the power CPS decoupling steady state results are shown in Table 3.

It can be seen that the decoupling caused by the attack on the 24th node is the greatest. After the decoupling, there are only 45 nodes left in the stable system, accounting for 78.9% of the total number of nodes. In the case of cascading failures caused by a cyber-attack on the power CPS, the importance of nodes based on the node survival rate \( S \) is ranked as follows: \( \{24,29,4,22,25,31,32,34,56,9,30,56,1,18,19,20,21,52,53,54,55,10,12,13,14,26,27,28,37,38,41,48,36,11,39,44,45,46,75,21,2578,136,1723,40,42,43,49,50,57\} \). There are some nodes with the same \( S \) in the sorting result, so the importance of these nodes cannot be accurately distinguished. The centrality of each node calculated by the compound centrality algorithm of the one-sided network can just make up for this shortcoming. Considering two indexes at the same time, the importance of nodes can be distinguished smoothly.

In summary, the experiment gives indexes that can describe the importance of IEEE57 nodes from two aspects: the compound centrality of the single-layer network and the value \( S \) that reflects the degree of cascading failure decoupling, as shown in Table 3. Both indexes can be used as a reference basis for operators and operation centers in the operation of the power system. The dispatch center can select a certain index according to the actual situation of the cyber-attack to distinguish the importance of the node and guide the follow-up strategy.

Data Availability
The data used to support the findings of this study are included within the article.

Conflicts of Interest
The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments
This work was supported by the National Key R&D Program of China (2018YFA0702200) and the Fundamental Research Funds for the Central Universities (N2004013).

References
[1] Q. Guo, S. Xin, H. Sun et al., “Power system cyber-physical modelling and security assessment: motivation and ideas,” Proceedings of the CSEE, vol. 36, no. 6, pp. 1481–1489, 2016.
[2] Y. Liu and P. Ning, False Data Injection Attacks against State Estimation in Electric Power Grids, Association for Computing Machinery, New York, NY, USA, 2009.
[3] S. Cui, Z. Han, S. Kar et al., “Coordinated data-injection attack and detection in the smart grid: a detailed look at enriching detection solutions,” IEEE Signal Processing Magazine, vol. 29, no. 5, pp. 106–115, 2012.
[4] Y. Wang, K. Gao, T. Zhao et al., “Assessing the harmfulness of cascading failures across space in electric cyber-physical system based on improved attack graph,” Proceedings of the CSEE, vol. 36, no. 6, pp. 1490–1499, 2016.
[5] X. Yi, B. Wang, D. Chen et al., “Review on interdependent networks theory and its applications in the structural vulnerability analysis of electrical cyber-physical system,” Proceedings of the CSEE, vol. 36, no. 17, pp. 4521–4532, 2016.
[6] A. Vespignani, “The fragility of interdependency,” Nature, vol. 464, no. 7291, pp. 984–985, 2010.
[7] X. Yi, Research on Structural Vulnerability of Power Information-Physical System Based on Interdependence Network Theory, Wuhan University, Wuhan, China, 2016.
[8] Q. Guo, S. Xin, J. Wang et al., “Comprehensive security assessment for a cyber physical energy system a lesson from Ukraine’s blackout,” Automation of Electric Power Systems, vol. 40, no. 5, pp. 1–3, 2016.
[9] Z. Meng, Z. Lu, and J. Song, “Comparison analysis OF the small-world topological model of chinese and american power grids,” Automation of Electric Power Systems, vol. 28, no. 15, pp. 21–24, 2004.
[10] R. Wang, Q. Sun, P. Zhang et al., “Reduced-order transfer function model of the droop-controlled inverter via Jordan continued-fraction expansion,” IEEE Transactions on Energy Conversion, vol. 35, no. 3, pp. 1585–1595, 2020.
[11] W. Hu, C. Ruan, H. Nian, and D. Sun, “Zero-sequence current suppression strategy with common mode voltage control for open-end winding PMSM drives with common DC bus,” IEEE Transactions on Industrial Electronics, vol. 99, 2020.
[12] R. Albert, I. Albert, and G. L. Nakarado, “Structural vulnerability of the North American power grid,” Physical Review E, vol. 69, no. 2, Article ID 25103, 2004.
[13] M. Kitsak, L. K. Gallos, S. Havlin et al., “Identification of influential spreaders in complex networks,” Nature Physics, vol. 6, no. 11, pp. 888–893, 2010.
[14] S. Dolev, Y. Elovici, and R. Puzis, “Routing betweenness centrality,” Journal of the ACM, vol. 57, no. 4, pp. 1–27, 2010.
[15] A. N. Langville and C. D. Meyer, “Deeper inside page rank,” Internet Mathematics, vol. 1, no. 3, 2003.
[16] Y. Chen, A. Hu, and X. Hu, “Evaluation method for node importance in communication networks,” Journal of China Institute of Communications, vol. 25, no. 8, pp. 129–134, 2004.
[17] Y. Tan, J. Wu, and H. Deng, “Evaluation method for node importance based on node contraction in complex networks,” Systems Engineering Theory & Practice, vol. 26, no. 11, pp. 79–83, 2006.
[18] Q. Xie, C. Deng, H. Zhao et al., “Evaluation method for node importance of power grid based on the weighted network model,” *Automation of Electric Power Systems*, vol. 33, no. 4, pp. 21–24, 2009.

[19] J. Wang, X. Gu, T. Wang et al., “Power system critical node identification based on power tracing and link analysis method,” *Power System Protection and Control*, vol. 45, no. 6, pp. 22–29, 2017.

[20] Z. P. Network, “Security situation analysis based on a dynamic bayesian network and phase space reconstruction,” *Journal of Supercomputing*, vol. 76, no. 2, pp. 1342–1357, 2020.

[21] G. Li, P. Huang, Y. Chen et al., “Security situation assessment method for cyber physical power system considering threat propagation characteristics,” *Electric Power Construction*, vol. 40, no. 5, pp. 29–37, 2019.

[22] V. Buldyrev Sergey, R. Parshani, G. Paul, H. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, 2010.

[23] Y. Tan, X. Li, Y. Cai et al., "Critical node identification for complex power grid based on electrical distance," *Proceedings of the CSEE*, vol. 34, no. 1, pp. 146–152, 2014.

[24] R. Wu, B. Zhang, and L. Tang, "A cascading failure based nodal importance evaluation method applied in dual network coupling model," *Power System Technology*, vol. 4, pp. 1053–1058, 2015.

[25] M. E. J. Newman, “The spread of epidemic disease on networks,” *Physical Review E Statistical Nonlinear & Soft Matter Physics*, vol. 66, no. 1, Article ID 16128, 2002.

[26] S. Mousavian, J. Valenzuela, and J. Wang, "A probabilistic risk mitigation model for cyber-attacks to PMU networks," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 156–165, 2015.

[27] D. Krackhardt, "Assessing the political landscape: structure, cognition, and power in organizations," *Administrative Science Quarterly*, vol. 35, no. 2, pp. 342–369, 1990.

[28] L. C. Freeman, "Centrality in social networks: conceptual clarification," *Social Networks*, vol. 1, no. 3, 1979.

[29] X. Ji, B. Wang, D. Liu et al., "Improving interdependent networks robustness by adding connectivity links," *Physical A: Statistical Mechanics and its Applications*, vol. 444, pp. 9–19, 2016.