Modeling and Analysis of Cascading Failures in Cyber-Physical Power Systems Under Different Coupling Strategies

HUAN PAN1,2, (Senior Member, IEEE), XIAO LI1, CHUNNING NA1,2, AND RUIJIA CAO1

1School of Physics Electronic-Electrical Engineering, Ningxia University, Yinchuan 750021, China
2Ningxia Key Laboratory of Electrical Energy Security, Yinchuan 750021, China

Corresponding author: Huan Pan (pan198303@gmail.com)

This work was supported in part by National Natural Science Foundation (NNSF) of China under Grant 61763040, in part by the Natural Science Foundation of Ningxia under Grant 2022AAC05001, in part by the Western Light Talent Training Program of Chinese Academy of Sciences, and in part by the Third Batch of Ningxia Youth Talents Supporting Program under Grant TJGC2018095.

ABSTRACT The cyber-physical power systems (CPPS) may experience cascading failures due to damage to power transmission lines. This paper focuses on modeling and analysis of CPPS with grid-line random failure under different coupling strategies to reveal the coupling mechanisms between the power grid and the information network. The CPPS mainly includes the power grid, information network, and the coupling connections between them. An uncertain failure mechanism is introduced in the cascading failures model to determine the failed power lines. The DC optimal power flow calculates and distributes the power flows to determine the tripping rates of power lines. A new comprehensive metric is proposed to assess the critical nodes, which considers both topology and state information of the power grid. According to some assessment metrics, the nodes in the power grid and information network are ranked. The CPPS model is constructed by coupling the power and information nodes in specific sequences. The cascading failures of CPPS are simulated when the power lines suffer random attacks to obtain the most robust coupling method for typical coupled CPPSs.

INDEX TERMS Critical nodes assessment metrics, cyber-physical power systems (CPPS), information network, power grid, robustness, stochastic cascading failures.

I. INTRODUCTION

The smart grid is the energy sector’s cyber-physical system (CPS) [1]. It provides adequate information and data support for planning and construction, production and operation, comprehensive management and services, new business and model development, and enterprise ecosystem construction by widely applying information and intelligent technologies such as big data, cloud computing, Internet of Things, mobile Internet, artificial intelligence, and so on. The scale, function, and intelligence of the power CPS (CPPS) are getting higher and higher, but the power grid’s vulnerability also keeps presenting itself. In recent years, many countries and regions have experienced power outages in the power grid, for example, Brazil blackout on November 10, 2009 [2], Venezuela blackout on March 7, 2019 [3], and UK blackout on August 9, 2019 [4]. The issue of grid security is receiving increasing attention from governments [5], [6].

The major blackout events revealed that the failures of power components or the abnormal working condition of the information system might cause the large-scale irregular transfer of power flow in the power system. This situation resulted in power system cascade failure, which triggered significant power paralysis and caused great economic losses.

Most scholars applied the interdependence model to establish the CPPS model. They portrayed the power grid and the information network as two mutually coupled and interdependent complex networks and investigated their vulnerability.
However, there were differences in the modeling of CPPS, and different modeling approaches would yield different analysis results. Cascading failures in the Italian power-information coupled network were simulated in [7] using percolation theory, assuming that the collapse of the information node led to the failure of the power node and the failure of the power node similarly led to the destruction of the information node. An interdependency model was developed to study cascading failures in CPPS. The loss of an information node was determined by judging whether the information node belonged to the most significant component of the remaining information network; the failure of the power node led to the grid’s power flow reallocation [8]. Compared with the actual CPPS, there was no dispatch center in this model, which did not characterize the situation under attack in the existing grid well. The grid model took less account of electrical factors. A new model based on binary optimization was proposed to assess the vulnerability of the power system, which was a geographical network interdependent with the communication network. The numerical results showed that the power system vulnerability might increase considering the physical network dependencies, and the location of the dispatch center can mitigate the impact of certain attacks [9]. Reference [9] lacked an accurate description of the information network and used two weighting coefficients to describe the delay for possible communication delays. But in practical applications, the communication delay problem should be simulated in the whole communication network. To better model CPPS, dispatch network data was used to model communication networks, using examples of the IEEE 39 node system and the Guangdong 500 kV system in China. It was shown through simulation that in the case of random attacks on interdependent systems, the probability of catastrophic failures in the grid combined with a binary star dispatch data network was lower than having a mesh structure [10]. The vulnerability of CPPS under community attack styles was analyzed where topology consistency and scale-free cyber networks were considered, respectively [11]. The uncertainty of the physical component failure mechanism was not considered in any of the above literature, and the physical component failure mechanism did not match the actual situation. Most literature assumed the grid as a betweenness-capacity model, where the power line was set to be failed if its power flow crossed the capacity [12]. However, in the reality of the grid, because of environmental or other factors, line interruptions often do not occur instantaneously. Still, they are disconnected after a period of overload because they cannot withstand the continuous overload pressure.

Recently, a stochastic cascading failures model was proposed, which was found to be a more realistic response to the impact of the power grid when subjected to a stochastic attack by comparing the actual grid cascading failures simulation diagram [13]. Subsequently, Lai and Guo introduced the stochastic model from [13] into the modeling of CPPS cascading failures and investigated the effect of capacity factor changes and different coupling modes on CPPS robustness under intentional and random attacks [14], [15]. References [14] and [15] did not do an in-depth study on the coupling modes and connectivity of the information network. A novel interdependence CPPS model was proposed to identify essential nodes in the grid by the proposed electrical degree metrics and to couple power nodes with information nodes based on degree ascending order using one-to-one correspondence [16]. The surviving nodes in the cascading failures model only considered the maximum connected subgraph in this literature. In reality, not all other subgraphs in the accurate grid failed collectively after disconnecting from the maximum connected subgraph. Some of them may switch to the islanded operation mode and continue to work. The CPPS was constructed by considering the islanded operation case of the power grid in [17], which made the CPPS model only a single degree coupling case: degree-degree coupling, stochastic degree coupling, and inverse degree-degree coupling.

The References [18] and [19] applied degree centrality, betweenness centrality, and closeness centrality as local metrics to analyze the power grid’s topological vulnerability and found that these metrics can play critical roles in identifying critical nodes. Meanwhile, betweenness centrality also enabled the identification of important lines. Zhang, Peng, et al. utilized the inverse of the line impedance from the shortest path between two nodes to indicate the difficulty of the connection between two nodes [16]. The node load was used as an indicator of state vulnerability [14]. However, neither the structural nor the state information can fully reflect the grid node information.

The above analyses revealed some shortcomings in the existing literature on the identification of critical nodes in power grids and the study of CPPSs. The primary manifestation was that only the topological vulnerability of the grid had been considered in the study of the power grid critical nodes identification, and the operation state vulnerability of the grid has not been addressed. Most of the CPPS models did not consider the uncertainty of power component failures. Some only used betweenness-capacity models to simulate power line overloads, lacking the portrayal of the power grid flows.

To address the above problems, we would like to conduct research on modeling and analysis of cascading failures of CPPS under different coupling strategies considering the situation of isolated islands and non-dispatchable power nodes in this paper. The main contribution of our work is described as follows:

- The gravitational model (GM) is introduced to measure to topology vulnerability of power nodes. A new critical nodes identification metric is proposed combining the topological and state vulnerabilities of the grid. The correlation analysis is done to obtain the weighting factors of topology and state vulnerability metrics.
- A stochastic power grid cascading failures model is developed using a stochastic physical component uncertainty failure mechanism to simulate the uncertainty of power component failures and to overcome the...
unrealistic power grid portrayal by the betweenness-capacity model.

- The CPPS model is established by identifying the critical nodes in the power grid and information network and coupling power nodes and information nodes in different sorting. A CPPS model is formed by coupling the IEEE-118 bus system with a BA-scale free network. The cascading failures simulation is performed to determine the best grid-information network coupling strategy using the start time as the robustness evaluation index.

The rest of this paper is organized as follows. Section II introduces the CPPS model, including the power grid, information network, and stochastic failure model. In Section III, the DC optimal power flow (DCOPF) model and the robustness assessment indicator are given. Section IV shows the traditional critical nodes identification metrics and provides a new comprehensive metric. The robustness of stochastic CPPS under different coupling strategies is analyzed in Section V. Finally, a conclusion is provided in Section VI.

II. CPPS MODEL

The power grid and the information network in CPPS coordinate to achieve dynamic closed-loop control. The nodes in the power grid, such as power stations and substations, provide energy support for nodes in the information network. The nodes in the information network include routers, monitors, controllers, etc., which monitor power nodes and send control commands in real-time to them. A typical CPPS topology consists of power nodes, information nodes, power lines, communication lines (or information edges), and coupled edges of power nodes and information nodes as shown in Figure 1, which can be expressed as an adjacency matrix $A$ [11]

$$A = \begin{bmatrix} A_P & A_{P-C} \\ A_{P-C}^T & A_C \end{bmatrix}$$

where $A_P$ denotes the power grid adjacency matrix; $A_C$ denotes the information network adjacency matrix; $A_{P-C}$ is the correlation matrix between the power grid and the information network; the elements of $A_{P-C}$ indicate whether there are connected edges between the information network and the power grid.

The correlation matrix is written as $A_{P-C} = [a_{P-C}(ij)]_{N \times M}$, where $N$ represents the number of power nodes, and $M$ represents the number of information nodes. If there is an edge between power node $P_i$ and information node $C_j$, then $a_{P-C}(i, j) = 1$, otherwise it is 0. Figure 1 shows a typical interdependence CPPS topology, where the power and information nodes are one-to-one coupled. The red node in the information network indicates the dispatch center, which is not directly coupled with the power grid. The classical CPPS model assumes that if a power node fails, the connected information node fails too, and vice versa. However, when an information node fails, the dispatch center will not send the control command to its corresponding power node. The power node will not be invalidated immediately but become a non-dispatchable node. The dispatch center for load shedding cannot control non-dispatchable nodes, and the failed power nodes will not belong to the set of non-dispatchable nodes.

A. POWER GRID MODEL

The power grid can be abstracted as a complex network $G = (V, E)$, where $V$ denotes the set of the power nodes, including power stations, substations, and loads; $E$ represents the set of transmission lines between the power nodes. The adjacency matrix of the power grid is expressed as $A_P = [a_P(ij)]_{N \times N}$, if a transmission line exists between nodes $i$ and $j$, then $a_P(i, j) = 1$, otherwise $a_P(i, j) = 0$.

When some transmission lines are tripped in the grid, we assume that one of them is randomly selected to be disconnected based on the probability of component failure instead of instantaneously disconnecting all the overload transmission lines. The higher the overload pressure, the higher the likelihood of disconnection. Only one transmission line is disconnected before the next state transition of the grid. After disconnecting the transmission line for the power flow redistribution, a new slack node needs to be assigned as a benchmark for each island due to the possibility of creating an islanded grid. For islands with no generation node, the nodes therein are considered comprehensively failed.

B. INFORMATION NETWORK MODEL

The structure of the information network in CCPS mainly includes three different networks: artificial scale-free network (BA model) [20], small-world network (SW model) [21], and random network (ER network) [22].

The adjacency matrix of the information network is denoted as $A_C = [a_C(ij)]_{M \times M}$, if a connection exists between two nodes $i$ and $j$, then $a_C(i, j) = 1$, otherwise $a_C(i, j) = 0$. The dispatch center with an uninterrupted power supply collects the status of grid nodes through information nodes and...
sends unified control commands to realize the dispatching of power stations and loads. When an information node fails, all the edges connected to it fail. If there is no communication path between the information node and the dispatch center, this information node is also considered invalid. In the simulation process, the Dijkstra algorithm [23] is used to judge the shortest path between each node in the information network and the dispatch center, and if there exists no direct route between any node and the dispatch center, then this information node is considered to be invalid.

C. STOCHASTIC FAILURE MODEL FOR POWER GRID

Zhang, Zhan, and Chi simulated the cumulative number of trips in a July 1996 blackout in the western North American power grid [13]. The stochastic cascading failures model was closer to the actual power grid outage by comparison. Therefore, to simulate the uncertain failure mechanism of the physical components of the power grid more realistically, a stochastic cascading failures model is developed in this section. For a given power grid, the deterministic power flow equation generates a sequence of failures and their locations. An interruption of a transmission line is a transition of the grid states, and the propagation of the power grid’s cascading failure can be considered a change of states of the grid.

Let \( \Xi_i(t) \in [0, 1] \) be the state of a given power transmission line \( i \), where \( \Xi_i(t) = 0 \) indicates normal operation of line \( i \) at time \( t \); \( \Xi_i(t) = 1 \) indicates that line \( i \) is removed (tripped or open) at time \( t \). Let \( \lambda_i \) denote the transition rate of line \( i \) from the state 0 to 1. Transmission line disconnection rate in the grid is caused by natural equipment failure or tripping of its protection equipment, that is

\[
\lambda_i(t) = \lambda^0_i(t) + \lambda^1_i(t),
\]

where \( \lambda^0_i(t) \) is the equipment failure rate when there is no load pressure and its value is constant, \( \lambda^1_i(t) \) is the removal or tripping rate of the relay protection device and is determined by the load conditions and transmission line capacity.

When the load on transmission line \( i \) is within its capacity, it is assumed to operate under normal conditions and would not be removed or tripped by the protective relay [24], i.e. \( \lambda^1_i(t) = 0 \). When a transmission line is loaded beyond its capacity, it would be delayed before it is removed or tripped. Since the tripping rate is related to the degree of overload, the tripping rate of transmission line \( i \) can be written as [13]

\[
\lambda^1_i(t) = \begin{cases} 
\alpha_i \left( \frac{L_i(t) - C_i}{C_i} \right), & L_i(t) > C_i, \\
0, & L_i(t) \leq C_i,
\end{cases}
\]

with

\[
C_i = (1 + \alpha)L_i(0),
\]

where \( L_i(t) \) is the power load of transmission line \( i \), obtained from the power flow calculation; \( C_i, \alpha, \) and \( L_i(0) \) represent the capacity, redundancy factor and initial flow of transmission line \( i \), respectively; \( \alpha_i \) is the introductory unit rate. The larger the \( \alpha \) means that the line power is less prone to overload, generally, we set \( \alpha = 1.2, \alpha = 0.035 \) [13]. The simulation is conducted according to the above parameters if not specified in the subsequent text.

For normal operating conditions, during a cascading failure, \( \lambda^1_i(t) \gg \lambda^0_i(t) \). For the universal case, it is assumed that in the analysis of power system cascading failures \( \lambda_i(t) \approx \lambda^1_i(t) \) [25].

According to the transmission line state transition, the trip probability of each transmission line is integral to the trip rate concerning time. It is assumed that the system always reaches a steady state when a trip occurs and that the transient before the system reaches the next steady-state is short enough to make the cumulative effect negligible. As far as the propagation of cascading failures is concerned, it is possible to consider outages caused by overloads, ignoring the circuit elements’ nonlinear characteristics and likely oscillatory behavior.

Suppose the power grid \( G \) has \( n \) transmission lines and the state vector of \( G \) is defined as \( \Xi(t) = \{ \Xi_1(t), \Xi_2(t), \ldots, \Xi_n(t) \} \). The dynamic propagation of cascading failures in grid \( G \) is equivalent to the dynamic evolution of \( \Xi(t) \). Given the grid’s current state, state transitions can be described by the time of the following state transition, and the subsequent transmission line being tripped is identified by probability.

The transition rate of the power grid states is defined as the sum of the tripping rates of the overloading lines:

\[
\lambda^*(t_1) = \sum_{i \in \Omega_0} \lambda_i(t_1), \tag{5}
\]

where \( \Omega_0 \) indicates the set of overloading transmission lines; \( \lambda_i(t_1) \) denotes the tripping rate of transmission line \( i \) at the moment \( t_1 \); \( \lambda^*(t_1) \) represents the sum of the state transition rates of all overloaded transmission lines in the grid, and its physical meaning is interpreted as the overloading stress of the entire power grid.

Let \( \tau \) denote the time interval between two adjacent power grid state transitions. It is calculated from \( \lambda^*(t_1) \) [13]

\[
\tau = \frac{\ln(1 - Z_1)}{\lambda^*(t_1)} \tag{6}
\]

where \( Z_1 \) is a random number uniformly generated in \((0, 1)\). When \( Z_1 \) is larger, the transition time interval \( \tau \) is shorter, and state transitions (cascading failures) in the power grid would occur quickly.

In the stochastic model, any overloaded transmission line in \( \Omega_0 \) may be triggered first. From a probabilistic point of view, transmission lines with larger \( \lambda_i(t_1) \) are more likely to be tripped. Therefore, the relative probability of transmission line \( i (i \in \Omega_0) \) tripping first is

\[
f_i = \frac{\lambda_i(t_1)}{\lambda^*(t_1)} \tag{7}
\]
The \( j \)th transmission line in \( \Omega_0 \) is selected to be tripped according to
\[
\sum_{k=0}^{j-1} \frac{\lambda_k}{\lambda^2} \leq Z_2 \leq \sum_{k=0}^{j} \frac{\lambda_k}{\lambda^2},
\]
where \( Z_2 \) is a random number uniformly generated in \((0, 1)\). With equation (8), it can be specified that only one transmission line is disconnected due to overload in each time interval \( \tau \).

III. CASCADING FAILURES MODEL

A. DC OPTIMAL POWER FLOW MODEL

The linearized DC flow model is extensively adopted to approximate the AC flow model for engineers in large power grids. The DC flow model has advantages over the AC model in simplicity, robustness, and real-time operations. With these advantages mentioned above, the DCOPF model used in this paper is as follows: the objective function is minimizing generation cost and load shedding.

\[
\text{min} \sum_{k \in G} b_k P_{gk} + c \sum_{i \in \Gamma} LS_i
\]

s.t. \( F_{ij} = \frac{1}{x_{ij}}(\theta_i - \theta_j) \)
\[
P = B\Theta
\]
\[
P_{out}^k - P_{in}^k - P_{gk} + LS_k = P_k, \quad k \in \mathcal{V}
\]
\[
P_{gk}^{min} \leq P_{gk} \leq P_{gk}^{max}, \quad k \in \Gamma \cap G
\]
\[
\Delta P_k = 0, \quad \Delta P_k = 0, \quad k \in \mathcal{F},
\]

where (9) is the objective function, \( P_{gk} \) represents the output power of power station \( k \), \( G \) denotes the set of power stations, \( b_k \) represents the generation cost factor of power station \( k \), \( LS_i \) represents the load reduction of node \( i \), \( \Gamma \) is the set of dispatchable power nodes, \( c < 0 \) represents the load shedding cost factor, the absolute value of \( c \) is set large enough to ensure that load shedding is performed only when the constraint cannot be satisfied under the generation station dispatch; (10) is the DC flow equation, \( F_{ij} \) is the power flow between nodes \( i \) to \( j \), \( x_{ij} \) is the normalized inductance of the transmission line between nodes \( i \) and \( j \), \( \theta_i \) denotes the voltage phase angle vector of node \( i \); (11) is the matrix form of the power grid flows, \( P \) is the power vector of the grid nodes, \( B \) is the conductance matrix of the grid, \( \Theta \) is the vector of node-voltage angles; (12) represents the nodal power balance equation, \( P_{out}^k \) and \( P_{in}^k \) are the outflow power and inflow power of node \( k \), respectively, \( P_k \) denotes the load of node \( k \); if there is no power station at node \( k \), then \( P_{gk} = 0 \); (13) denotes the constraint on the output power of node \( k \), \( P_{gk}^{max} \) and \( P_{gk}^{min} \) denote the upper and lower limits of \( P_{gk} \), respectively; (14) shows that the load reduction cannot exceed the load of the node; (15) indicates that a power node that loses its information node will not be dispatchable. \( \mathcal{F} \) denotes the set of non-dispatchable nodes, \( \Delta P_{gk} \) and \( \Delta P_k \) denote the amounts of power change and load change at node \( k \), respectively. The dispatching function of the information network to the power grid is realized by (9)-(15).

B. ROBUSTNESS ASSESSMENT INDICATORS

The robustness assessment metrics of the power grid usually include the percentage of remaining load on the grid, the number of failed nodes, etc. Besides, the growth rate of the number of tripped lines in the process of power grid cascading failures, i.e., removing overloaded lines or trip frequency, is also significant, and the rapid rise in trip frequency of overloaded lines is a harbinger of substantial outages. The point in time when the grid transition rate or overload stress is very high across the grid is just the time when the cascade failures start to be widespread. This point in time is defined as the time when the cascading failure propagation rate increases rapidly and is called the start time. The start time is a critical point where the grid can avoid a significant outage if remedial measures are taken to protect the grid before the start time. Therefore, the start time will be used as a robustness assessment indicator of the power grid. If the start time is large enough, the grid is considered more robust, and if it is short, the grid is less robust.

For simplicity, the start time \( t_{set} \) is defined as the time at which the transition rate of the power grid state reaches its maximum, i.e.,
\[
t_{set} = \arg \max \left\{ \lambda^*(t_1), \lambda^*(t_2), \ldots, \lambda^*(t_h) \right\},
\]
where \( h \) represents the number of cascading failures cycles.

IV. CRITICAL NODES IDENTIFICATION

A straightforward solution to avoiding catastrophic cascading failures in the power grid is to identify the critical nodes [18], [19], [26] for additional protection and prevention. In general, the physical characteristics of a power system are determined by two aspects: the topology and the operating state. Power nodes’ importance can be identified by traditional state vulnerability analysis and topological vulnerability analysis [27].

Neither the grid structure nor the state information used in the existing literature can fully reflect the power nodes’ importance. This section combines the topological vulnerability metric and the state vulnerability metric to propose a new comprehensive evaluation metric to identify the critical nodes in the power grid.

A. NODE VULNERABILITY METRICS

Considering both topological vulnerability and state vulnerability of the power grid, the topological vulnerability metric and state vulnerability metric are combined in a linear weighting manner to generate a new comprehensive assessment metric, and the optimal weights are sought using simulation to ensure the validity of the selected weight values.

1) TOPOLOGICAL VULNERABILITY

Standard complex network topology vulnerability metrics are mainly based on degree, betweenness, and closeness
centrality; besides, GM is introduced compared to traditional topology metrics. For a complex network $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ are sets of nodes and edges, respectively. The definitions of the above vulnerability metrics are shown as follows [28].

Degree centrality (DC) is defined as
\[
D_C(i) = \sum_{j \in \mathcal{N}_i} a_{ij},
\]
where $\mathcal{N}_i$ denotes the set of neighboring nodes of node $i$, $a_{ij}$ is the element of $G$’s adjacent matrix $A = [a_{ij}]$.

The expression of betweenness centrality (BC) is given by
\[
B_C(i) = \sum_{i \neq j \neq k \in \mathcal{V}} \frac{\sigma_{jk}(i)}{\sigma_{jk}},
\]
where $\sigma_{jk}$ denotes the number of the shortest paths from nodes $j$ to $k$, $\sigma_{jk}(i)$ denotes the number of the shortest paths from nodes $j$ to $k$ which pass node $i$.

Closeness centrality (CC) is defined as
\[
C_C(i) = \frac{1}{\sum_{j \in \mathcal{V}/i} d_{ij}},
\]
where $d_{ij}$ represents the shortest path from node $i$ to node $j$.

Recently, Ma, Ma, Zhang, et al. were inspired by the law of gravity and proposed a variant algorithm called gravitational model (GM) by considering neighborhood information and path information [29]. GM believes that a node with larger degree (neighborhood information) and with an average shorter distance to other nodes (path information) is more influential [30]. The definition of GM is represented as
\[
G(i) = \sum_{i \neq j \in \mathcal{V}} \frac{D_C(i) \times D_C(i)}{d_{ij}^2}.
\]

2) STATUS VULNERABILITY

This paper quantifies the state vulnerability using node loads (NL) and electrical connection efficiency (ECE) of nodes to investigate the effect of operating state on CPPS robustness. NL is often used to describe the state vulnerability of the power grid [14] to evaluate the impact of power nodes on the overall power grid operating status.

\[
L(i) = \text{load}(i)
\]
where $\text{load}(i)$ is the loads of node $i$.

The ECE of a power node is defined as the sum of the inverse of the shortest path impedance between the nodes to account for the connecting difficulty among power nodes [16].

\[
E(i) = \frac{1}{|\mathcal{V}|} \sum_{i \neq j \in \mathcal{V}} d_{ij},
\]
where $|\mathcal{V}|$ represents the number of power nodes; $d_{ij}$ is the electrical distance between nodes $i$ and $j$ and it is expressed by the inverse of line impedance of the shortest path from $i$ to $j$:
\[
d_{ij} = 1/Z_{ij}.
\]

B. A COMPREHENSIVE ASSESSMENT METRIC

The topological vulnerability and state vulnerability metrics are combined in a linearly weighted manner to obtain a new comprehensive metric that considers both the impact of the grid structure when it is disrupted and the impact of the operational state. The new metric is expressed as
\[
TS(i) = \omega \frac{T(i)}{\sum_{j \in \mathcal{V}} T(j)} + (1 - \omega) \frac{S(i)}{\sum_{j \in \mathcal{V}} S(j)},
\]
where $T(i)$ and $S(i)$ denote a topology and a status vulnerability metrics, respectively, $\omega$ is a weighting factor indicating the weight of the topological structure and state information in the evaluation of nodes importance.

C. GRID CASCADE FAILURES MODEL

To demonstrate the validity of the comprehensive metric TS and to determine the optimal weight values, simulations are performed with the IEEE-118 standard grid model, which contains 118 nodes, 186 edges, and 54 nodes for power stations. The hardware computational environment used for simulations is Inter(R) Xeno(R) E3-1235L v5 CPU @2.00GHz 2.00GHz, 8.00GB RAM. The software environment is Win. 10, MATLAB2020b.

A grid topology graph $G$ is generated from the IEEE-118 standard power grid model. When a physical component of the grid fails, the power flow is recalculated according to (10) and (11). The line will disconnect if the transmission line flow exceeds its flow limit after the calculation.

The residual load ratio (RLR) of the grid is adopted to evaluate the effectiveness of the node importance metric (24).\[
R_L = \frac{P_{suv}}{P},
\]
where $P_{suv}$ is the power grid load after a cascading failure, $P$ is the initial load of the grid.

The failure node ratio (FNR) indicates the degree of damage to the entire grid structure due to the failure of any nodes and can be used to evaluate the effect of the comprehensive metric TS, expressed as
\[
R_n = \frac{N_{\text{failure}}}{N},
\]
where $N_{\text{failure}}$ is the number of failed power nodes after a cascading failure, $N$ is the number of nodes in the entire grid.

The DC flow equations (10) and (11) are used to calculate the grid power flow, and (4) is used to determine whether the grid line is disconnected. The IEEE-118 standard grid is simulated for cascading failures, and the effectiveness of the metric (24) is assessed using RLR and FNR. There are four different topological vulnerabilities: GM, DC, BC, and CC, and two different state vulnerabilities: NL and ECE. Therefore, a total of eight combinations to obtain separate TSSs. The specific simulation flow is shown in Figure 2, and the detailed simulation steps are as follows:

Step 1. Initialize the power grid and model the grid topology graph; Calculate the grid initial power flow according
to (10)-(11), and let $k = 1, t = 0$, where $t$ represents the time interval between each topology change of the grid.

Step 2. Select the $k$th node in the vulnerability sequence as the initial failure node and calculate the grid’s power flow.

Step 3. Remove all lines that have tripped for their power flow exceeds the limits, and modify the topology graph based on the overhead transmission lines disconnection; Do islands detection, determine whether there exist power stations on each island. If it is, go to step 4, or else, there do not exist any power stations on some islands, set the power flow of these islands to be zero (These islands have all failed due to lack of electricity.).

Step 4. Recalculate the power flow of the power grid according to (10)-(11), and obtain the power flow of each line; Compare the power flow of $i$th line $L_i(t)$ with $C_i$, and judge whether $L_i(t) > C_i$. If it is, go to step 4; or else, go to step 5.

Step 5. The grid enters a steady-state and calculating $R_L$ and $R_n$. Judge whether $k > 117$, if it is, go to step 6, or else, go to step 2.

Step 6. The simulation is finished.

D. COMPREHENSIVE METRIC VALIDITY ANALYSIS

RLR and FNR after the IEEE-118 standard grid cascading failures are calculated separately, and their correlation analyses with the proposed comprehensive metric TS are done to prove the effectiveness of TS. The correlation analysis is calculated using Spearman’s rank correlation coefficient [31], which has the following expression:

$$C_{coeff} = \frac{\sum (r_i - \bar{r})(y_i - \bar{y})}{\sqrt{\sum (r_i - \bar{r})^2 \sum (y_i - \bar{y})^2}}, \quad i \in V,$$

where $r_i$ denotes the value corresponding to a power node calculated according to a specific metric and arranged in ascending order; $\bar{r}$ denotes the average value of the sequence $r_i$; $y_i$ indicates the RLR or FNR after attacking the power nodes in turn, $\bar{y}$ is the average value of the sequence $y_i$.

The Spearman correlation coefficient indicates the correlation between $r_i$ (independent variable) and $y_i$ (dependent variable). If $y_i$ tends to increase when $r_i$ increases, $C_{coeff} > 0$; If $y_i$ tends to decrease when $r_i$ increases, $C_{coeff} < 0$.

Figure 3 plots the correlation coefficient curves between the RLR and the comprehensive metric TS composed of the four topological vulnerability metrics combined with ECE, where the horizontal coordinate represents the weights, and the vertical coordinate represents the correlation coefficients between TS and RLR. From the definition of the Spearman coefficient, if TS shows a negative correlation with RLR, and the higher the degree of negative correlation, the better the effect of the new indicator TS. In Figure 3, TSs consisting of GM, DC, and ECE correlate best with RLR, followed by BC and the worst CC. Besides, the effects of GM and DC are very close to each other, with only a tiny difference in the weights corresponding to the minimum values of the two curves. For the four “U” curves, when $w = 0$ and $w = 1$, the corresponding correlation coefficients are more prominent, indicating lower correlations. It suggests that the assessment metrics, including topological and state vulnerability, have better results than single metrics (either topological or state metrics).

Figure 4 portrays the correlation coefficients between the FNR and the metric TSs composed of the four topological vulnerability metrics combined with ECE. With the same conclusion as Figure 3, the four “n” curves in Figure 4 show that the assessment metric that considers both topological and state vulnerability performs better than others.
vulnerability and state vulnerability is better than a single metric to determine nodes’ importance.

Figures 5 and 6 depict the correlation coefficients between RLR, FNR and the metrics TSs, respectively, where TSs are four topological vulnerability metrics each combined with NL. The two figures show that the TS consisting of GM and NL has the best performance, followed by DC, BC, and CC. The curves in Figure 5 are still “U” shaped, while the curves in Figure 6 are “n” shaped, and we can also obtain similar conclusions as in Figures 3 and 4.

Comparing Figures 3, 4, 5, and 6, it is clear that the comprehensive metric TS with GM has the most significant Spearman correlation coefficient. It means that the TS built with GM has better performance in identifying critical nodes. Except for Figure 6, the horizontal coordinates corresponding to the extreme values of the curves in the other three plots are basically at [0.4, 0.6], which indicates that both topological vulnerability and state vulnerability have essential effects on the power grid. Taken together, the power node comprehensive metric TS constructed by choosing GM and NL at \( w = 0.55 \) is the best for critical node identification.

V. ROBUSTNESS ANALYSIS OF CPPS UNDER DIFFERENT COUPLING STRATEGIES

In this section, a moderately connected scale-free network is set as the information network to study the variation of CPPS robustness under different coupling methods between the power grid and the information network. The coupling manner means that specific metrics are used to rank the importance of power nodes and information nodes, and connections are established between power nodes and information nodes according to particular rules.

A. COUPLING METHOD

The standard coupling method ranks the importance of power nodes and information nodes by degree or betweenness and then to couple the corresponding power nodes and information nodes in the positive sequence coupling, inverse sequence coupling, disorderly coupling, or utterly random coupling. Positive sequence coupling is the first power node coupling with the first information network node, the second power node coupling with the second information network node, and so on until all nodes are coupled; inverse sequence coupling implies that the first power node is coupled to the last information node, the second power node is related to the penultimate information node, and so on until all nodes are connected; the disorderly coupling means random coupling between power and information nodes sequences until all nodes are coupled; completely random coupling means power nodes and information nodes are randomly associated with a certain probability.

In this section, we use the four metrics TS, GM, degree, and betweenness, respectively, to rank the power nodes, where the weight \( w = 0.55 \) is chosen for TS. Information nodes' importance is ranked by degree, betweenness, and GM. Thus, there are 12 ways of coupling power nodes to information nodes: B-B, B-D, B-GM, D-B, D-D, D-GM, GM-B, GM-D, GM-GM, TS-B, TS-D, and TS-GM, where B and D denote betweenness and degree, respectively.

B. CPPS CASCADING FAILURES SIMULATION FLOW

The IEEE-118 standard grid is used as the power grid model. The BA scale-free network with 119 nodes is generated by BA construction algorithm [20] to obtain the information...
network. The dispatch center is the node with the largest betweenness in the information network. The grid couples with the information network in 12 ways to constitute different CPPS models, and the following specific simulation process of the failures model is established:

Step 1. Initialize the power grid and information network, calculate initial grid power flow using DCOPF through MATPOWER toolkit in MATLAB, and record initial loads.

Step 2. The number of attacking lines of the power grid varies from 1 to 10.

Step 3. Detect isolated islands in the grid and balance the nodes’ power in isolated islands.

Step 4. Check the information nodes coupled to the power nodes. If there is a failed power node, the corresponding information node failed; check whether all information network nodes are connected to the dispatch center. If there is no router between an information node and the dispatch center, it is also failed, and all edges are severed.

Step 5. Check the power nodes coupled to the information nodes. If there is a failed information node, the corresponding power node is recorded as a non-dispatchable node, which cannot be controlled by the dispatch center for load shedding or power generation.

Step 6. The DCOPF is used again to reallocate the power flow, and if a no-solution situation occurs, the DCPF is used to calculate the power flow.

Step 7. Calculate (5), if the result is not 0, select the tripped power lines according to (8) and return to step 4; if the calculation result is 0, there is no overload line at this time, the cascading failure simulation ends and the robustness indicator start time is calculated.

The flow chart of the CPPS cascading failures simulation is shown in Figure 7.

C. ANALYSIS OF SIMULATION RESULTS

Rank the information and power nodes according to node importance metrics, respectively, and build different CPPS models with varying coupling methods. Randomly attack the power lines, the number of attacking lines is \( x \), and \( x \in \{1, 2, \cdots, 10\} \). CPPS cascading failures simulation is done by using MATLAB, according to Figure 7. Since the stochastic CPPS model is used and the start time values for each simulation are not representative, we performed 100 simulations and calculated the mean value of the start time to plot the simulation results. BC is used for the critical power nodes ranking, and BC, DC and GM are adapted to sort the information nodes, respectively. The cascading failure simulation results of CPPS under different coupling methods are shown in Figures 8 to 10.

The start time gradually gets smaller as the number of failed lines increases in Figures 8, 9, and 10. As more lines are attacked, the power flow redistribution leads to more power nodes failure, making information nodes fail. In turn, failed information nodes make power nodes non-dispatchable. With such a cascading failure, the power grid faces burdensome overload stress, and the cascade failure begins to spread widely.

By comparing the start time in Figures 8, 9, and 10, we find that the robustness of CPPS is ranked as follows for different coupling methods: positive sequence coupling > disorderly coupling > inverse sequence coupling > completely random coupling. Both the power grid and the information network have scale-free network characteristics. In the case of positive sequence coupling, random attacks that fail to disable important nodes have limited damage to CPPS. In the case of disorderly coupling, some important power nodes may be...
coupled with a relatively large number of information nodes that are more prone to failure and affect important power nodes through the interdependence model. Inverse sequence coupling is the extreme case of disorderly coupling. Essential nodes are necessarily connected to low-importance nodes, therefore have a higher failure probability and are more prone to failure. The best CPPS robustness is achieved with GM as the information node ranking, D as the second-best, and B as the worst by changing the orders of the information nodes coupled with the power nodes.

The start time curves of CPPSs under D-B, D-D, D-GM coupling manners are described in Figures 11, 12, 13, respectively. We can obtain the similar conclusions as in Figures 8, 9, 11. Comparing Figures 8, 9, 11, and 12, we see among the four conventional coupling methods, the robustness of CPPS under the B-D coupling is the best, and the robustness under the B-B, D-B, and D-D coupling methods decreases in order. It indicates that the power grid focuses more on the connectivity represented by the betweenness. At the same time, the degree measures the neighborhood information of the power nodes and thus appears less effective than the betweenness. Due to the existence of an independent dispatch center in the information network, information nodes coupled with high betweenness power nodes need to be guaranteed to be connected to the dispatch center. Therefore a large number of information nodes with a high degree are required to ensure the integrity of the path. B-D works best among the four coupling methods, B-D, D-D, B-B, and D-B. The results obtained in [32] which considered a one-to-one coupling CPPS model, are essentially the same as those concluded in Figures 8, 9, 11, and 12.

When the coupling methods for power and information nodes in CPPS are GM-B, GM-D, and GM-GM, the simulation results of cascading failure are presented in Figures 14, 15, and 16. D and GM are used as the power nodes’ importance ranking, respectively, and the robustness
rankings of CPPS after ranking the importance of information nodes with different metrics are both $\text{GM}>\text{B}>\text{D}$.

When TS is used to rank the power nodes’ importance, the start time curves of CPPS suffering from cascading failures are shown in Figures 17, 18, and 19. We can obtain the same the robustness rankings conclusion as before by comparing the start time in Figures 17 to 19.

In summary, GM works best for organizing the importance of information nodes. The degree mainly considers the neighborhood information of the network, the betweenness believes the path information of the network, and the GM includes both the neighborhood and the path information. In Figures 10, 13, 16, and 19, using GM as the importance metric of information nodes, the importance rankings of power nodes are carried out in four ways, respectively. We observe that the CPPS robustness is be ranked as $\text{TS}>\text{GM}>\text{B}>\text{D}$. The same results are concluded in Figures 8, 11, 14, 17, and 9, 12, 15, 18.

Through the analyses of the above simulation results, CPPS shows the best robustness in the face of random attacks with TS-GM positive sequence coupling compared to the
traditional coupling methods, and TS and GM have better performance in identifying the importance of power and information nodes, respectively.

VI. CONCLUSION

The cascading failures of CPPS under different coupling strategies have been analyzed where the uncertain failure mechanism is applied to simulate the cascading failures of the power grid. DCOPF is used to calculate the power grid flow, and the start time is utilized as the robustness assessment indicator. In the process cascading failures, when a power node fails, the information nodes coupled with it would be invalid; when a information node fails, the corresponding power node would become a non-dispatchable node. A comprehensive metric has been proposed to measure vulnerability of power nodes, which integrates the topological and state vulnerabilities. Different vulnerability metrics have been applied to rank the power nodes and information nodes, and different coupling strategies are obtained by coupling these two kinds of nodes in four manners. Let’s take the IEEE-118 bus system as an example for simulation analysis. The results have revealed the CPPS with TS-GM positive sequence coupling strategy is better for power grid or information network, a random failure shows the best robustness in the face of random attacks. The mechanism is applied to simulate the cascading failures of the power grid and the information network, a random failure of a power node does not necessarily make an important information node fail under the positive sequence coupling for the CPPS and vice versa. In addition, TS can evaluate the power nodes comprehensively, which considers the power grid’s topology and state information; GM can better reflect the importance of information nodes, including the neighborhood and the path information of the information nodes.

This paper adopted a one-to-one coupling CPPS model, but a power node may supply electricity for multiple information nodes, and an information node may monitor many power nodes. So, the power grid or information network may be one-to-many coupling or many-to-many coupling in the CPPS nodes. And an information node may monitor many power nodes, which integrates the topological and state vulnerabilities; GM can better evaluate the power nodes comprehensively, which considers the power grid and the information network, a random failure shows the best robustness in the face of random attacks.

REFERENCES

[1] X. H. Yu and Y. S. Xue, “Smart grids: A cyber-physical systems perspective,” Proc. IEEE, vol. 104, no. 5, pp. 1058–1070, May 2016.

[2] O. P. Veloza and F. Santamaria, “Analysis of major blackouts from 2003 to 2015: Classification of incidents and review of main causes,” Electr. J., vol. 29, no. 7, pp. 42–49, Sep. 2016.

[3] D. Ryan and M. Castro, “Venezuela blackout plunges millions into darkness,” Wall Street J., Mar. 2019.

[4] J. Carpani, A. Shipman, and K. OagNeill. (Aug. 9, 2019) Major Power Cut Across Country as London Goes Dark After National Grid Failure. UK News. [Online]. Available: https://www. telegraph.co.uk/news/2019/08/09/majorpowercut-across-country-london-goes-dark-national-grid/

[5] S. Sridhar, A. Hahn, and M. Govindaraju, “Cyber-physical system security for the electric power grid.” Proc. IEEE, vol. 100, no. 1, pp. 210–224, Jan. 2012.

[6] D. L. Liu, X. Liu, H. Yu, W. Wang, X. Zhao, Y. Zhao, H. Zhang, L. Ma, D. Li, and Q. Li, “Research on technology application and security threat of Internet of Things for smart grid,” in Proc. 5th Int. Conf. Inf. Sci. Control Eng. (ICISCE), Jul. 2018, pp. 496–499.

[7] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, “Catastrophic cascade of failures in interdependent networks,” Nature, vol. 464, no. 7291, pp. 1025–1028, Apr. 2010.

[8] Y. Zhang and O. Yagan, “Modeling and analysis of cascading failures in interdependent cyber-physical systems,” in Proc. IEEE Conf. Decis. Control (CDC), Dec. 2018, pp. 4731–4738.

[9] M. Zeraati, Z. Aref, and M. A. Latifi, “Vulnerability analysis of power systems under physical deliberate attacks considering geographic-cyber interdependence of the power system and communication network,” IEEE J. Sel. Areas in Commun., vol. 31, no. 6, pp. 1111–1120, Aug. 2013.

[10] Y. Cai, Y. Cao, Y. Li, T. Huang, and B. Zhou, “Cascading failure analysis considering interaction between power grids and communication networks,” IEEE Trans. Smart Grid, vol. 7, no. 1, pp. 530–538, Jan. 2016.

[11] H. Pan, H. Lian, C. Na, and X. Li, “Modeling and vulnerability analysis of cyber-physical power systems based on community theory,” IEEE Syst. J., vol. 14, no. 3, pp. 3938–3948, Sep. 2020.

[12] Y. Zhang, J. Wu, W. Ren, C. K. Tse, and C. Zheng, “Sequential restorations of complex networks after cascading failures,” IEEE Trans. Syst., Man, Cybern., Syst., vol. 51, no. 1, pp. 400–411, Jan. 2021.

[13] X. Zhang, C. Zhan, and C. K. Tse, “Modeling the dynamics of cascading failures in power systems,” IEEE J. Emerg. Sel. Topics Circuits Syst., vol. 7, no. 2, pp. 192–204, Jun. 2017.

[14] R. Lai, X. Qiu, and J. Wu, “Robustness of asymmetric cyber-physical power systems against cyber attacks,” IEEE Access, vol. 7, pp. 61342–61352, 2019.

[15] H. Guo, S. S. Yu, H. H. C. Iu, T. Fernandez, and C. Zheng, “A complex network theory analytical approach to power system cascading failure—From a cyber-physical perspective,” Chaos, Interdiscip. J. Nonlinear Sci., vol. 29, no. 5, May 2019, Art. no. 053111.

[16] H. Zhang, M. Peng, J. M. Guerrero, X. Gao, and Y. Liu, “Modelling and vulnerability analysis of cyber-physical power systems based on interconnected networks,” Energies, vol. 12, no. 18, p. 3439, Sep. 2019.

[17] L. Chen, D. Yue, C. Dou, Z. Cheng, and J. Chen, “Robustness of cyber-physical power systems in cascading failure: Survival of interdependent clusters,” Int. J. Electr. Power Energy Syst., vol. 114, Jan. 2020, Art. no. 105374.

[18] T. A. Ernest and A. K. Srivastava, “Power system vulnerability analysis—towards validation of centrality measures,” in Proc. IEEE PES T&D, May 2012, pp. 1–6.

[19] H. Cetinay, K. Devriendt, and P. Van Mieghem, “Nodal vulnerability to targeted attacks in power grids,” Appl. Netw. Sci., vol. 3, no. 1, pp. 1–19, Dec. 2018.

[20] A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” Science, vol. 286, no. 5439, pp. 509–512, 1999.

[21] D. J. Watts and S. H. Strogatz, “Collective dynamics of small-world networks,” Nature, vol. 393, no. 6684, pp. 440–442, 1998.

[22] P. Erdös and A. Rényi, “On random graphs,” Publ. Math., Debrecen, vol. 6, pp. 290–297, Jan. 1959.

[23] E. W. Dijkstra, “A note on two problems in connexion with graphs,” Numer. Math., vol. 1, no. 1, pp. 269–271, Dec. 1959.

[24] P. Rezaei, P. D. H. Hines, and M. J. Eppstein, “Estimating cascading failure risk with random chemistry,” IEEE Trans. Power Syst., vol. 30, no. 5, pp. 2726–2735, Sep. 2014.

[25] I. Chen, J. S. Thorp, and M. Parashar, “Analysis of electric power system disturbance data,” in Proc. 34th Annu. Hawaii Int. Conf. Syst. Sci., Jan. 3, 2001, p. 2020.

[26] R. Ghanbari, M. Jalili, and X. Yu, “Correlation of cascade failures and centrality measures in complex networks,” Future Gener. Comput. Syst., vol. 83, pp. 390–400, Jun. 2018.

[27] E. Bompad, R. Napoli, and F. Xue, “Stochastic models for the assessment of structural vulnerability in transmission networks,” IET Gener. Transmiss. Distrib., vol. 4, no. 6, pp. 716–724, Jun. 2010.

[28] R. Albert, H. Jeong, and A.-L. Barabási, “Error and attack tolerance of complex networks,” Nature, vol. 406, no. 6804, pp. 378–382, Sep. 2000.

[29] L.-L. Ma, C. Ma, H.-F. Zhang, and B.-H. Wang, “Identifying influential spreaders in complex networks based on gravity formula,” Phys. A, Stat. Mech. Appl., vol. 451, pp. 205–212, Jul. 2016.

[30] Z. Li, T. Ren, X. Ma, S. Liu, Y. Zhang, and T. Zhou, “Identifying influential spreaders by gravity model,” Sci. Rep., vol. 9, no. 1, p. 8387, 2019.
[31] E. C. Fieller, H. O. Hartley, and E. S. Pearson, “Tests for rank correlation coefficients,” Biometrika, vol. 44, pp. 470–481, Jan. 1957.

[32] Z. C. Dong, Y. J. Fang, and M. Tian, “Influences of various coupled patterns and coupling strength on power-communication coupled networks,” High Voltage Eng., vol. 41, no. 10, pp. 3464–3469, 2015.

HUAN PAN (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees from Northwest Normal University, in 2005 and 2008, respectively, and the Ph.D. degree from Central South University, in 2012.

He is a Professor with the School of Physics and Electronic-Electrical Engineering, Ningxia University. His research interests include modeling and analysis of cyber-physical systems, cooperative control of microgrids, and grid-connection control of distributed generations.

XIAO LI received the B.Sc. degree from the Hefei University of Technology, in 2017. He is currently pursuing the M.Sc. degree with Ningxia University. His research interest includes vulnerability of cyber-physical systems.

CHUNNING NA received the Ph.D. degree from North China Electric Power University, China, in 2017. She is an Associate Professor with Ningxia University. Her research interests include electricity markets, generation schedule optimization, and power system planning.

RUIJIA CAO received the B.Sc. degree from the Qindao University of Technology, in 2020. She is currently pursuing the M.Sc. degree with Ningxia University. Her research interest includes cyber security of smart grids.