

Operational Risk-averse Routing Optimization for Cyber-physical Power Systems

Baozhong Ti, Jianxiao Wang, *Member, CSEE, Member, IEEE*, Gengyin Li[✉], *Senior Member, CSEE, Member, IEEE*, and Ming Zhou, *Senior Member, CSEE, Senior Member, IEEE*

Abstract—The extensive application of modern information and communication technology in the power system through the in-depth integration of the information system and the power system has led to the gradual development of the cyber-physical power system (CPPS). While advanced information technology increases the safety and reliability of power system operations, it also increases the risks of fault propagation. To improve the reliability of CPPS from the perspective of power communication routing, it is proposed that the CPPS model and vulnerability assessment of power node reflect the correlation between information and energy flows with the service impact on power grid operation, which is an important index for evaluating communication services. According to the distribution of services at the different important levels on the links, the importance of the cross-layer link is established as the vulnerability evaluation index of the communication network. Then, the routing optimization model is proposed in combination with the service transmission risk under cyber-attack and the operating characteristics of the information system, which is solved through an improved fast-convergent genetic algorithm. The simulation results show that the proposed method allocates the alternate route to the low-risk link without significantly increasing the delay of the main route, which effectively improves the power supply reliability of CPPS in extreme cyber-attack scenarios.

Index Terms—Comprehensive balanced degree, cross-layer link importance, cyber-attack, cyber-physical power system, routing optimization model, service importance.

I. INTRODUCTION

UNDER the powerful drive for the construction of smart grids, the application of advanced computer and communication technology in power systems continues to deepen. The development of cyber-physical power systems (CPPS), which result from the in-depth integration of information systems and power systems, is an inevitable development [1]–[3]. This deep coupling has enhanced the observability and controllability of the power system, but it has also increased the potential risk of fault propagation and expanded the scale of accidents in the

power system [4]. For example, in 2015, the Ukrainian State Grid Information System was attacked by hackers, leading to the control servers of the underlying generators and substations to shut down, thereby causing large-scale power outages [5].

As the infrastructure that handles the energy information interaction in the smart grid, the power communication network provides a high degree of reliability for the safe operation of the whole system. Thus, the reliability and risk research of this system is of important significance [6], [7]. Routing optimization is an effective method to reduce the operational risks of the communication network without changing the network topology. To improve the safe operation of the power communication network, many experts and scholars have conducted research on routing optimization and protection from different levels of threats. The common methods of service transmission are the shortest path model, availability-aware routing (AAR), and risk balancing (Load Balancing, LB). In [8], the author selects one of the k shortest links as the transmission path. In [9], the shortest path algorithm is proposed to build a general graph model suitable for static service grooming in heterogeneous wavelength-division multiplexing (WDM) mesh networks. The AAR is proposed under different failure probability conditions by defining the minimum cross-layer cut set and minimum spanning tree set in [10]. In [11], by defining the concept of available boundary and assuming that routing has a dynamic failure arrival rate and fixed repair time, a method for evaluating routing ability and allocation algorithm is proposed. The shortest path method and the availability routing method consider the service transmission delay and the state of the transmission channel, respectively, but they cannot avoid the excessive concentration of important communication services. At the same time, the above methods do not consider the service importance in routing planning, making them unsuitable for solving the service distribution problem related to power communication network operation risks. Simply considering balancing of services may also lead to the transmission of some important services being delayed or slow. In [12], a routing method is proposed by using the index of risk balance degree and average service risk as optimization targets. A fast non-dominated genetic algorithm is utilized with an elite strategy to achieve overall path planning. In [13], a load balancing mode is established based on the bandwidth resource of nodes and links. Considering the attributes of service types and the joint balance of link load and service risks, a route optimization method is proposed. A balanced service routing determination algorithm

Manuscript received January 12, 2021; revised April 15, 2021; accepted July 13, 2021. Date of online publication January 5, 2022; date of current version March 10, 2022. This work was supported by the National Key Research and Development Program of China under Grant 2016YFB0901100.

B. Z. Ti, J. X. Wang, G. Y. Li (corresponding author, e-mail: ligy@ncepu.edu.cn; ORCID: <https://orcid.org/0000-0002-2639-4875>), and M. Zhou are with State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China.

DOI: 10.17775/CSEEJPES.2021.00370

is proposed by using the analytic hierarchy process to assess the degree of network risk from the perspectives of equipment, service load and service characteristics in [14]. The above-mentioned methods for service evaluation are mainly based on the communication characteristics of the service, and don't fully consider its impact on the safe operation of the power system. There is no clear distinction between the same kind of communication services from different power infrastructure.

Most of the existing research on routing optimization start from the perspective of information systems but don't consider the interdependence between the communication network and power network. Therefore, we face the following two challenges that need to be resolved. Firstly, the assessment of importance amongst the power communication services needs to fully consider the service impact on the reliable operation of the power grid. The safe transmission of communication services provides reliable guarantees for the stable operation of the power system. In actual work, the service distribution is mostly from the perspective of performance indicators of services without considering its impact on grid operation, which leads to some services that have a large impact on power supply reliability not being given priority. Secondly, the research on service transmission lacks comprehensive consideration of the transmission risks. This leads to a mismatch between the reliability of communication links and the importance of the services. In other words, some communication services that have a large impact on the power system may be concentrated on high-risk links, while other services that have less impact on the power system are reliably transmitted. The above-mentioned challenges increase the risk of large-scale power outages in some extreme scenarios such as cyber-attacks and extreme weather. By optimizing the paths of information flow in communication networks to accurately perceive and control the current risks of the power grid, the overall robustness of CPPS can be effectively improved.

To guarantee the reliable operation of the CPPS, a risk-averse routing optimization model considering the interdependences of CPPS and network attacks is proposed in this paper. The main contributions are as follows:

1) A service evaluation method is proposed based on the service impact on power grid operation and the correlation between information flow and energy flow. Considering operating status and network structure of the power system, a vulnerability assessment index is established as the reference basis for service importance evaluation.

2) According to the network attack mode and the distribution of services with different importance degrees, a quantification index is defined for the cross-layer link importance. Then, an assessment index is proposed by utilizing information entropy for the overall risk estimation.

3) Based on the above evaluation approach, a risk-averse routing model is built to make the main-alternate routing strategy consider transmission delay and operation reliability comprehensively. The integer non-linear programming problem is solved through an improved adaptive genetic algorithm.

This paper is organized as follows: In Section II, we propose the CPPS network structure model and cyber-attack patterns. Section III puts forward the service importance

evaluation method, the operational risk assessment index, and the route optimization method of the communication network. Section IV provides an improved fast genetic algorithm. The effectiveness of the model is proved by a case study in Section V. Finally, we provide a summary in Section VI.

II. MODEL AND EVALUATION OF CPPS

A. Structural Analysis and Model of CPPS

According to the system operation characteristics and basic architecture theory of CPPS [15], we divide CPPS into the electrical physical layer and the power communication layer, as shown in the Fig. 1. We equate it as an undirected graph, which can be represented as (L_P, L_C) .

1) Power communication layer: This layer mainly includes the communication network, uplink channels/downlink channels, communication network, and transmitted information flow. The communication network is mainly composed of communication equipment and communication protocols whose main function is information transmission. The uplink/downlink channel is responsible for uploading data and issuing instructions. At present, the transmission technology commonly used in the backbone communication network is the passive optical ethernet network, which has large transmission capacity and strong anti-interference capability, completing data transmission in an end-to-end manner. The information transmission layer is represented by $L_C = (G_C, M_C)$, which incorporates the network structure G_C and operation status M_C . $G_C = (V_C, E_C)$, where $V_C = \{v_{c1}, v_{c2}, \dots, v_{cN_{VC}}\}$ is the information node set, which represents the router, dispatch center, and subnet in the communication network. $E_C = \{e_{c1}, e_{c2}, \dots, e_{cN_{EC}}\}$ is the communication link set. N_{VC} and N_{EC} are the number of communication nodes and communication links, respectively. $M_C = (S, R)$, where $S = (v_{k,s}, v_{k,d}, I_k)$ is the service information matrix, which is composed of the source node $v_{k,s}$, destination node $v_{k,d}$ of service k , and service importance I_k . The "1 + 1" protection mode is often used in path planning of key services, which plans a main routing and an alternative routing for each service [16]. R represents the routing strategy composed of the main routing strategy $R_M = \{r_{m,ij}^k\}$ and the alternate routing strategy $R_S = \{r_{s,ij}^k\}$. If the path of main routing of service k includes the link (i, j) , $r_{m,ij}^k = 1$. Otherwise, $r_{m,ij}^k = 0$. The value of $r_{s,ij}^k$ is taken in the same way.

2) Electrical physical layer: This layer mainly includes electric primary equipment, secondary equipment, and the transmitted energy flow. Intelligent terminals serve as the interface between the electrical physical layer and the power communication layer and are responsible for information collection and instructions transmission. The electrical physical layer is represented by $L_P = (G_P, M_P)$, including grid structure G_P and operating status M_P . $G_P = (V_P, E_P)$, where $V_P = \{v_{p1}, v_{p2}, \dots, v_{pN_{VP}}\}$ is the power node set representing generators, loads, and substations. $E_P = \{e_{p1}, e_{p2}, \dots, e_{pN_{EP}}\}$ is the power line set. N_{VP} and N_{EP} are the number of power nodes and power lines, respectively. The operating state variables of this layer include power P and power angle θ under DC power flow analysis, that is, $M_P = (P, \theta)$.

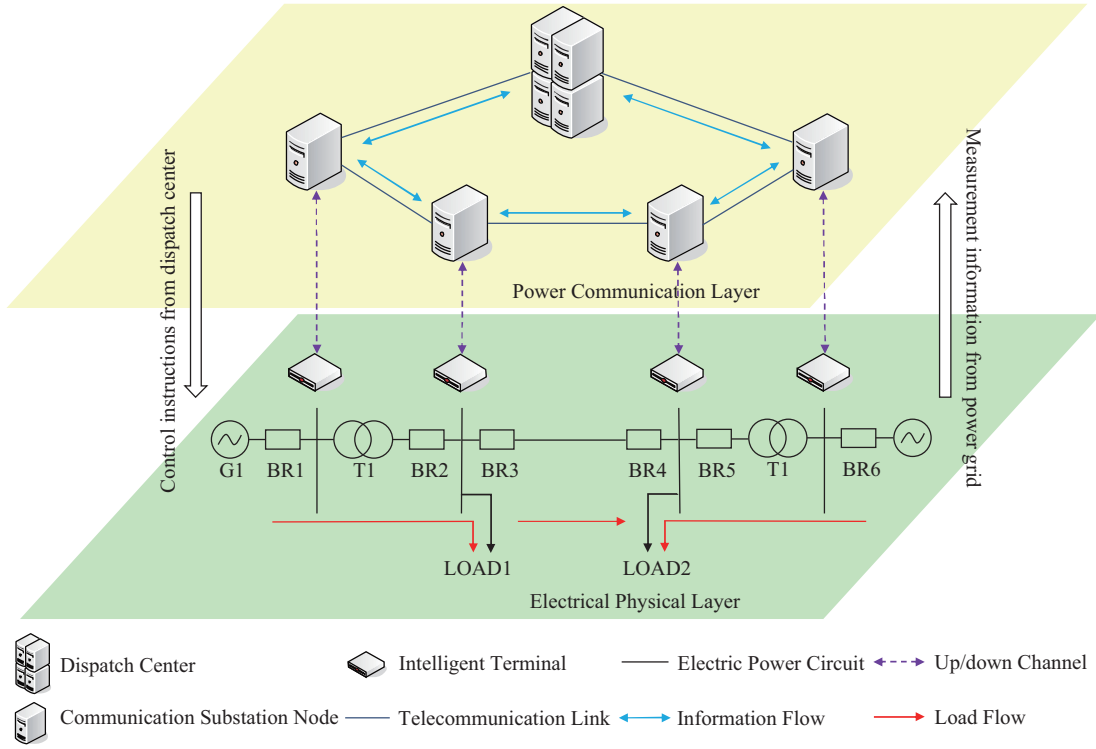


Fig. 1. Hierarchical structure of cyber-physical power system.

We use the associated matrix $\mathbf{G}$ [17] to describe the relationship between G_P and G_C :

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{C-P}^{(N_C \times N_C)} & \mathbf{G}_{C-E}^{(N_C \times N_E)} \\ \mathbf{G}_{C-P}^T & \mathbf{G}_{E-C}^{(N_E \times N_E)} \end{bmatrix} \quad (1)$$

where $\mathbf{G}_{C-P} = \{g_{C-P,ij}\}$ and $\mathbf{G}_{C-P}^T = \{g_{C-P,ij}^T\}$ are the cyber-physical associated submatrix and physical-cyber associated submatrix, respectively. The elements in G_{C-P} are defined as follows:

- 1) If there is no information transmission link between node i and node j , $g_{C-P,ij} = 0$.
- 2) If there is an information transmission link between node i and node j , $g_{C-P,ij} = 1$.

B. Cyber Attack and Evaluation

The existing network attack patterns can be divided into deliberate attack, random attack [18], and area attack [19], [20]. Deliberate attack is most effective among them. In this paper, deliberate attacks are divided into betweenness priority attacks based on network topology and key line attacks based on node functions. The effectiveness of the proposed routing optimization method is evaluated based on the attack results.

1) Betweenness Priority Attack Model

Although network attacks according to service distribution is the most destructive approach, it is difficult for the attacker to obtain accurate distribution information. Since communication optical cables are usually built into the overhead ground wires of high-voltage transmission lines, communication network topology is relatively easier to access. Therefore, the hacker may attack the communication network based on its topology.

The number of shortest paths for all node pairs passing through an edge is defined as its “betweenness.” The larger the betweenness of the line, the more services the line transmits and the more likely the line is to be attacked. Calculate the betweenness of each edge in E_C , sort them in descending betweenness order, and take the first n edges as betweenness priority attack set $A_{\text{between}} = \{e_{b,1}, \dots, e_{b,n} | e_{b,i} \in E_C\}$.

2) Key Link Attack Model

If the attacker invades the database and gets the information of the node functions and service distribution, the links connected to the dispatch center are very likely to be attacked because a large number of services are concentrated on them. Sort the links connected to the dispatch center in descending order according to the number of services transmitted on them and take the first n edges as the deliberately attack set $A_{\text{key}} = \{e_{o,1}, \dots, e_{o,n} | e_{o,i} \in E_{CC}\}$.

The primary task of the CPPS is to ensure reliable power supply. At the same time, the remaining grid structure after failure is related to the difficulty of system recovery. Therefore, we define topology loss of power system (TLPS) T_{loss} and load loss of power system (LLPS) P_{loss} to evaluate routing robustness from the perspectives of grid structure and operating state.

$$T_{\text{loss}}(\lambda) = \frac{G_E^\lambda - G_E^0}{G_E^0} \quad (2)$$

$$P_{\text{loss}}(\lambda) = \frac{P_D(G_E^\lambda) - P_D(G_E^0)}{P_D(G_E^0)} \quad (3)$$

where G_E^0 and G_E^λ are the number of initial transmission links and the remaining links after λ^{th} attack, respectively.

$P_D(G_E^0)$ and $P_D(G_E^\lambda)$ are the load before and after the attack, respectively.

III. ROUTING OPTIMIZATION MODEL UNDER THE RISK OF CYBER ATTACK

A. Importance Evaluation of Communication Service

The existing routing optimization mechanism defines service importance based on the service type rather than the interdependence between the communication and power system. However, because different power nodes have different impacts on the power system in practice, their corresponding communication services should also have different importance, and corresponding distinctions should be made when making routing strategies although they are in the same category from the perspective of communication.

We take the circuit breaker, an important physical device in the power system, as the cyber-physical interaction medium of the CPPS. The corresponding service is relay protection, whose failure causes the decision control layer to lose the monitoring and control function for the circuit breaker. The failure of the monitoring function does not change the topology and has an indirect effect on the power system. The breaker rejecting act only occurs during the adjustment of the system operation mode. To highlight the impact of cyber-attacks on system operation, the relay protection service considered in this paper refers to the breaker malfunction after the false data injection attack, which directly leads to changes in topology and power flow and has a direct effect on the power system.

The power network is composed of a large number of branches connecting various power sources and loads according to a certain electrical relationship. Its vulnerability largely depends on the electrical and structural characteristics of the energy transmission paths between the generation nodes and load nodes. Therefore, we evaluate the importance of transmission lines according to the power flow distribution and electrical parameters. Then, we propose the definition of node importance, which can reflect the vulnerability of the power network objectively and provide a basis for service importance assessment.

Define the transmission line importance $I_{\text{line},ij}$ as follows:

$$I_{\text{line},ij} = \frac{P_{Lij}}{P_{L\max}} \sum_{\substack{x \in G \\ y \in D}} W_{Gx} W_{Dy} |T_{ij}(x, y)| \quad (4)$$

in which

$$W_{Gx} = \frac{P_{Gx}}{\sum_{x=1}^{N_G} P_{Gx}} W_{Dy} = \frac{P_{Dy}}{\sum_{y=1}^{N_D} P_{Dy}} \quad (5)$$

where P_{Lij} and $P_{L\max}$ represent the transmission power and transmission capacity of line (i, j) , respectively. W_{Gx} represents the weight of power generation node x , which is determined by P_{Gx} , the ratio of power generation in node x to the total power generation; W_{Dy} represents the weight of load node y , which is determined by P_{Dy} , the ratio of load in node y to the total load. $T_{ij}(x, y)$ represents the power transmission distribution factor (PTDF) of the generation node x and the load node y on the branch (i, j) .

PTDF reflects the influence degree of the power transfer changes between generation nodes and load nodes on power flow of branches, which can be defined as follows:

$$T_{ij}(x, y) = \frac{\Delta P_{Lij}}{\Delta P_{Lxy}} = \frac{X_{i-x} - X_{j-y} - X_{i-x} + X_{j-y}}{x_{ij}} \quad (6)$$

where P_{Lxy} represents the power change between the generation node x and the load node y ; P_{Lij} represents the branch power change under the influence of P_{Lxy} ; X_{i-x} represents the element in the row i and column x of the node reactance matrix, and x_{ij} represents the reactance value of the branch (i, j) .

Line importance reflects the power flow distribution on the paths between generation nodes and load node and quantifies the contribution of each branch to energy transmission. Thus, the indicator considers the location of key generation and load nodes. The traditional definition of node degree is the number of directly connected nodes, in other words, the number of edges connected to the node. The node degree can reflect the node importance in the network and a positive correlation exists between them in complex network theory. We define the sum of line importance for the lines connected to the node as the electrical degree D_{Pi} . Since it is difficult to distinguish the important difference between nodes with the same degree, the mutual influence between adjacent nodes must be considered. The power node importance I_{Pi} is defined as follows:

$$D_{Pi} = \sum_{(i,j) \in E_P} I_{Pi} \quad (7)$$

$$I_{Pi} = \frac{D_{Pi}}{D_P^2} \sum_{(i,j) \in E_P} D_{Pj} \quad (8)$$

in which

$$\bar{D}_P = \frac{1}{N_{VP}} \sum_{i=1}^{N_{VP}} D_{Pi} \quad (9)$$

where $\bar{D}_P$ is the average electrical degree.

The relationship between the service importance from communication node and the power node importance can be defined as follows:

$$\mathbf{I}_C = (G_{C-P})^T \mathbf{I}_P G_{C-P} \quad (10)$$

where $\mathbf{I}_C$ is the service importance matrix. $\mathbf{I}_P$ is the power node importance matrix.

We utilized the Grainger & Stevenson IEEE 4 bus system to build a CPPS model to illustrate the effectiveness of the evaluation model. V_{c2} is the dispatch center at the cyber side. The node importance relationship in the physical side is $V_{e4} > V_{e3} > V_{e2} > V_{e1}$. Node V_{e4} has the greatest importance because it is the only power generation node and undertakes all loads. Although the load of node V_{e2} is less than that of node V_{e3} , the connected branch (V_{e2}, V_{e4}) carries the load of node V_{e1} and node V_{e2} . Thus, the importance of node V_{e2} is higher. The node V_{e1} has the lowest load power and the transmission power of the connected line, hence, its importance is the lowest. Traditional evaluation methods such as node degree and betweenness centrality do not consider the operation characteristics [21]. Therefore, the evaluation

results of the four nodes' importance are the same. In [22], the evaluation method only considers the structure characteristics and sorting result is $V_{e3} > V_{e4} > V_{e1} > V_{e2}$, which is not accurate. When the information nodes transmit relay protection services for the corresponding power nodes, the service number of the link (V_{e1}, V_{e2}) is greater than that of the link (V_{e2}, V_{e4}) . However, considering the coupling between the two networks, the interruption of the link (V_{c1}, V_{c2}) leads to the removal of the load nodes V_{c1} and V_{c3} , and the node V_{c2} maintains normal operation. The interruption of the link (V_{c2}, V_{c4}) causes the generation node V_{e4} to split and the whole system suffers a power loss. The communication technical indicators of the three relay protection services are not significantly different. However, considering that the main function of CPPS is to maintain power supply reliability, the services should be distinguished and protected according to the extent of their impact.

B. Operation Risk Assessment for Communication Networks

Attackers mostly disrupt the communication link based on its topology to interfere in service transmission and system operation. Therefore, we evaluated the operation risk of the communication network from the perspective of attack mode and service distribution. The cross-layer link importance $I_{\text{cross},ij}$ of the communication link (i, j) is determined by its betweenness $B_{\text{link},ij}$ and the transmitted service importance $I_{\text{service},ij}$, which is as follows:

$$I_{\text{cross},ij} = B_{\text{link},ij} I_{\text{service},ij} \quad (11)$$

in which

$$I_{\text{service},ij} = \sum_k^{N_K} I_k (r_{m,ij}^k + r_{s,ij}^k) \quad (12)$$

where I_k is the service importance, which is the power node importance in this paper; $r_{m,ij}^k$ and $r_{s,ij}^k$ are the 0–1 decision variable of the main routing and the alternative routing of the service k on the link m , respectively.

Equation (11) comprehensively considers the attack probability of the link and the corresponding loss. When the number of link betweenness and transmitted service importance are

both relatively large, the important services are concentrated on a small number of links. These links are more likely to be attacked and the loss can be relatively serious. The operation risk is at a high level. When the distribution of cross-layer importance is relatively balanced, for example, the former is large and the latter is small, the link is more likely to be interrupted; however, the attack just causes slight damage when the former is small but the latter is large, and the system operation risk is small because of the low attack probability on the critical paths. Then, we propose the comprehensive balanced degree I_{bal} by using information entropy to evaluate the system operation risks under betweenness priority attacks, which can be described as follows:

$$I_{\text{bal}} = \frac{-\sum_{(i,j) \in E_C} I'_{\text{cross},ij} \ln(I'_{\text{cross},ij})}{\ln(N_{EC})} \quad (13)$$

in which

$$I'_{\text{cross},ij} = \frac{I_{\text{cross},ij}}{\sum_{(i,j) \in E_C} I_{\text{cross},ij}}, I'_{\text{cross},ij} \in (0, 1] \quad (14)$$

The range of I_{vul} is $(0, 1]$. The larger the value, the more balanced is the service importance distribution and lower the system operation risk.

C. Routing Optimization Model for Reducing System Operation Risk in CPPS

From the above analysis, take F_{between} as the routing risk of communication network under betweenness priority attack, which is defined as follows:

$$F_{\text{between}} = 1 - I_{\text{bal}} \quad (15)$$

The operation risk of the power communication network under a key link attack is related to the service distribution on these links. The more important the services that the line transmits, the more damage the system suffers after the attack. Take F_{key} as the routing risk of communication network under key link attack, which is defined as:

$$F_{\text{key}} = \sqrt{\frac{1}{|E_{CC}|} \sum_{(i,j) \in E_{CC}} (I_{\text{serve},ij}^* - \bar{I}_{\text{serve}}^*)^2} \quad (16)$$

in which

$$I_{\text{serve},ij}^* = \frac{I_{\text{serve},ij}}{\max_{(i,j) \in E_{CC}} I_{\text{serve},ij}}, \quad \bar{I}_{\text{serve}}^* = \frac{1}{|E_{CC}|} \sum_{(i,j) \in E_{CC}} I_{\text{serve}}^* \quad (17)$$

where $\bar{I}_{\text{serve}}^*$ is the average service importance on link (i, j) .

If we only consider the attack risk in route optimization, the communication service bypasses the key lines and chooses a longer path to reduce the attack probability, which leads to a higher transmission delay. The longer transmission path also increases service interruption probability under stochastic failure of the link. An ideal allocation strategy is that the path length of the important service should be as short as possible to reduce the transmission delay and minimize the interruption risk. Low importance services should try their

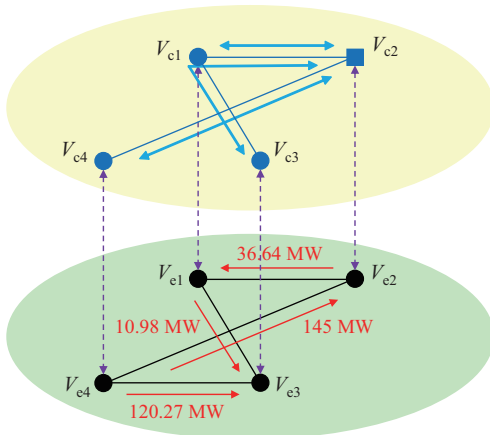


Fig. 2. Information flow and energy flow of a CPPS.

best to bypass the links with concentrated services or large betweenness and meet the delay constraint, to reduce service losses in case of cyber-attacks. The routing optimization model (ROM) that comprehensively considers cyber-physical interdependence and transmission risk is defined as:

$$\min \alpha_1 \sum_k \frac{I_k}{I_{k, \max} N_{MC}} \sum_{(i,j) \in E_C} d_{ij}(r_{m,ij}^k + \varepsilon r_{s,ij}^k) + \alpha_2 F_{\text{between}} + \alpha_3 F_{\text{key}} + \Lambda F_{\text{overlap}} \quad (18)$$

such that

$$\begin{cases} \sum_{(i,j) \in E_C} r_{m,ij}^k d_{ij} \leq T^k & \forall k \\ \sum_{(i,j) \in E_C} r_{s,ij}^k d_{ij} \leq T^k & \forall k \end{cases} \quad (19)$$

$$\sum_{k=1}^{N_K} (r_{m,ij}^k + r_{s,ij}^k) \leq b_{ij} \quad \forall (i,j) \in E_C \quad (20)$$

Equation (18) is the objective function. D_{ij} represents the communication delay of the link (i, j) . Coefficient ε is set to a relatively small number, so that the main path pays more attention to time delay and outage probability, to reduce the path switching probability. The alternate path focuses more on failure severity to prevent a lot of service losses after attack. $\alpha_1 + \alpha_2 + \alpha_3 = 1$ are coefficients of interruption risk, betweenness attack risk, and key line attack risk, which are determined according to risk aversion preference for route optimization. Λ is set to a relatively large value to ensure that the service can be transmitted by two independent paths. F_{overlap} is the penalty factor for the overlap between the main and alternate route paths, which is described as follows:

$$F_{\text{overlap}} = \sum_k \sum_{(i,j) \in E_C} J(r_{m,ij}^k, r_{s,ij}^k) \quad (21)$$

in which

$$J(r_{m,ij}^k, r_{s,ij}^k) = \begin{cases} 1 & r_{m,ij}^k = r_{s,ij}^k = 1 \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

where $J(r_{m,ij}^k, r_{s,ij}^k)$ is the overlap index for the main and alternate route paths.

Equations (19) and (20) are service delay constraints and bandwidth constraints, respectively. T^k is the maximum delay limit of service k . b_{ij} is the bandwidth of the link (i, j) .

IV. IMPROVED GENETIC ALGORITHM

For the non-linear integer programming problem of multi-objective routing planning, efficient optimization algorithms are needed to obtain the optimal solution. As a global optimization method, the genetic algorithm has strong parallelism and robustness [23]–[25]. Based on the traditional algorithm, we determine the probability of selection, crossover, and mutation according to the individual fitness in the evolution process, so that the genetic algorithm can adapt to environmental fluctuations and improve its search ability and convergence speed.

A. Chromosome Code and Fitness Function

We use real numbers for chromosome coding. Each chromosome represents a service path and each gene represents a path node. The gene sequence is the service transmission path. The multiple chromosomes representing different services form a routing optimization plan, which is regarded as the individual in the population.

In the genetic algorithm of routing planning, the objective function is transformed into the population fitness function, as shown in equation (21). When the algorithm result converges to a certain value, the individual with the highest fitness is the optimal solution. Restrictions screen the results of each generation to retain the satisfied results and regenerate the unsatisfied individuals, to ensure that each optimization result meets the requirements:

$$f_i^\lambda = 1 - \frac{C_i^\lambda - C_{\min}^\lambda}{C_{\max}^\lambda - C_{\min}^\lambda} \quad (23)$$

where I is the fitness value of individual i in the λ^{th} generation. C_i^λ is the cost function corresponding to the objective function in the routing model. $C_{\max}^\lambda$ and $C_{\min}^\lambda$ are the maximum and minimum costs of the λ^{th} generation, respectively.

B. Selection Operator Improvement

The classic genetic algorithm uses the roulette wheel method in individual selection. The higher the fitness value that the individual gets, the greater the probability it will have to be inherited by the next generation [26]. However, the high randomness of the selection result impacts the convergence speed. Therefore, we propose a selection method that can adjust the inheritance probability according to the fitness distribution. The selection operator η_i^λ generation of individual i in λ^{th} generation is described as follows:

$$\eta_i^\lambda = \frac{f_i^\lambda}{\sum_i^n f_i^\lambda} + k \left(\frac{f_{\text{ave}}^\lambda}{f_{\max}^\lambda} - \frac{f_{\text{ave}}^0}{f_{\max}^0} \right) \quad (24)$$

where f_{ave}^0 and $f_{\max}^0$ are the average and maximum value of individual fitness in the initial population, respectively. f_{ave}^λ and $f_{\max}^\lambda$ are the average and maximum value of fitness in the λ^{th} generation, respectively. k is the fitness distribution coefficient.

In the early stage of evolution, when the fitness distribution is relatively wide, the roulette wheel method can ensure population diversity. In the later stage of evolution, since the population fitness distribution is relatively concentrated, we increase the selection operator values to make individuals with high fitness more likely to be inherited in the next generation and improve the convergence speed of the algorithm.

C. Improvement of Crossover and Mutation Operators

Crossover and mutation are the key factors affecting the search of an optimal solution. The crossover operator determines the global search ability of the genetic algorithm and the mutation operator maintains the population diversity. Too large or too small a crossover rate, are more likely to destroy the optimal solution and decrease gene recombination, respectively. Too large or too small a mutation rate can close

the algorithm to random search and reduce the population diversity. Thus, we redefine an adaptive crossover operator η_c^λ and mutation operator η_m^λ to enhance the generalization ability, which can be defined as follows:

$$\eta_c^\lambda = \begin{cases} \eta_{cmax} & f_i^\lambda + f_j^\lambda / 2 \leq f_{ave}^\lambda \\ \eta_{cmin} + (\eta_{cmax} - \eta_{cmin})\beta_e^\lambda & f_i^\lambda + f_j^\lambda / 2 > f_{ave}^\lambda \end{cases} \quad (25)$$

$$\eta_m^\lambda = \begin{cases} \eta_{mmax} & f_i^\lambda \leq f_{ave}^\lambda \\ \eta_{mmin} + (\eta_{mmax} - \eta_{mmin})\beta_m^\lambda & f_i^\lambda > f_{ave}^\lambda \end{cases} \quad (26)$$

in which

$$\beta_e^\lambda = \frac{\arcsin\left(\frac{f_{max}^\lambda - (f_i^\lambda + f_j^\lambda)/2}{f_{max}^\lambda - f_{ave}^\lambda}\right)}{\pi/2} \quad (27)$$

$$\beta_m^\lambda = \frac{\arcsin\left(\frac{f_{max}^\lambda - f_i^\lambda}{f_{max}^\lambda - f_{ave}^\lambda}\right)}{\pi/2} \quad (28)$$

where f_i^λ and f_j^λ are the fitness values of the two individuals to be crossed. f_{max}^λ and f_{min}^λ are the maximum and minimum fitness of the individual in the λ^{th} generation, respectively. η_{cmax} and η_{cmin} are the maximum and minimum crossover rates, respectively. η_{mmax} and η_{mmin} are the maximum and minimum mutation rates, respectively.

It can be seen from equations (25) and (26) that if the individual fitness is close to the population average fitness, the individuals have a higher probability of crossover and mutation. In the early evolution stage, when the greatest number of individuals are close to the average fitness of population, new individuals can be generated with a high probability to ensure population diversity and avoid premature termination of the algorithm, and the algorithm convergence efficiency is improved. When the individual fitness value is close to the maximum fitness, the crossover and mutation probability is reduced to avoid algorithm falling into the local optimal solution. From the first generation, we repeat the above process until the fitness does not change significantly.

We perform crossover and mutation operations after getting the operators. The crossover operation adopts the staggered crossover method, which selects the chromosomes to cross in turn from the two individuals. We choose the same nodes from the two chromosomes as the exchange boundary. The mutation operation is to regenerate the path from any gene point in the chromosome. At the same time, it is necessary to judge and eliminate loops during the operation. The process is shown in Fig. 3.

V. CASE STUDIES

In this section, the CPPS model is constructed based on the IEEE30 bus system. To prove the effectiveness of the optimization method, the conventional short path routing model (SPRM) is tested as reference.

A. Description of Simulation Scenarios

Considering that most of the current communication lines (such as optical fibers) are laid along the transmission lines,

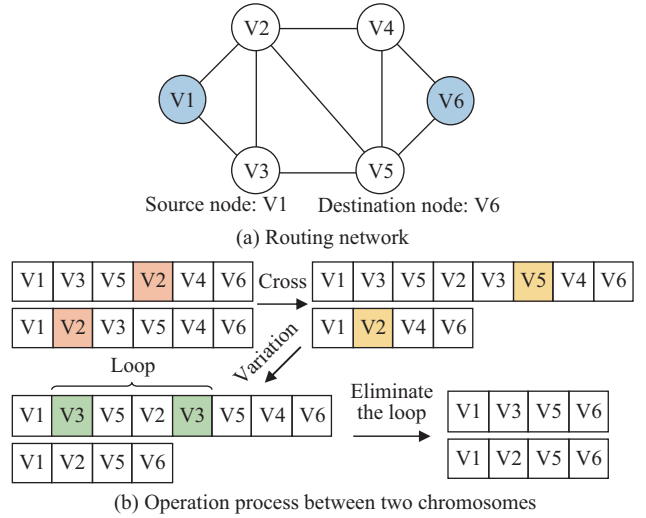


Fig. 3. The process of chromosome crossing, mutation, and loop elimination.

the topological structure of the transmission grid and the corresponding communication network is usually highly coupled. For example, the optical fiber is located inside the overhead ground wire, especially for new high-voltage transmission lines [27]. Therefore, in our case study, the nodes and branches in the communication network and the power network adopt a one-to-one coupling. The two networks have the same topological structure. Node 6 with the largest node degree is taken as the dispatch center on the cyber side. We assume that the communication services meet the delay requirements. The bandwidth is set to 15. In this paper, we focus on the influence of the communication routing methods on the power system operation. To reduce the influence of the power network structure and power-load distribution on the simulation results, we assume that the communication nodes are equipped with uninterruptible power supplies.

B. Parameter Setting and Algorithm Performance Comparison

We adopt the relay protection service as the communication side service in CPPS and “et” 1 + 1” route protection mechanism. When the main and alternate routings are interrupted, the communication node and the corresponding power node are inaccessible. We assume that the optimization target has the same weight for the three risks, that is, $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$. Since the values of the interruption risk and the whole risks are generally less than 1, we take ε as 0.1, to distinguish the functions between the main and alternate routing and take Λ as 1 to prevent them from overlapping. We use MATLAB to simulate the algorithm on the Intel(R) Core(TM) i7-7700HQ processor. The parameters of the simple genetic algorithm (SGA) and improved genetic algorithm (IGA) are shown in Table I. The population size is 200 and the evolution algebra is 1000. Due to the randomness of the genetic algorithm, we ran each algorithm 20 times and took the average of the results. The convergence performance is shown in Fig. 4.

The average operation time per generation for SGA and IGA are 1.9 seconds and 1.99 seconds, respectively. From Fig. 4, IGA has a better convergence performance than SGA. This is because the IGA adaptively adjusts selection, crossover, and

TABLE I
ALGORITHM PARAMETER SETTING

Algorithm	Operator	Value
SGA	Crossover Rate	0.8
	Mutation Rate	0.01
IGA	Maximum and Minimum Crossover Rate	0.8, 0.6
	Maximum and Minimum Mutation Rate	0.15, 0.05

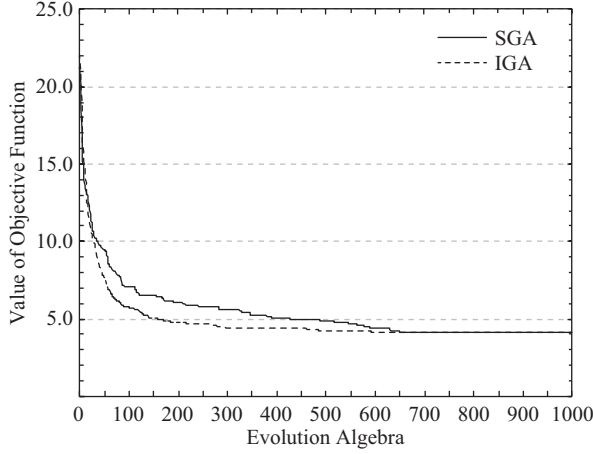


Fig. 4. Algorithm convergence performance comparison.

mutation rate according to the evolution stage, which enriches the population and improves the performance of the genetic algorithm in terms of convergence and stability.

C. Optimization Routes Based on Operation Risk

The transmission delay in the optical network is mainly composed of line transmission delay and node processing delay. The O/E/O conversion at the node needs much more time than for the line transmission [28], so we use routing hops to represent the transmission delay.

Figure 5 displays the comparison of the main and alternate routing delays between ROM and SPRM. The main route delays of the ROM and the SPRM are nearly equal. Some alternate route delays of ROM are significantly higher than that of SPRM. The optimal model reduces the service transmission risk under cyber-attacks at the expense of communication delay.

Figure 6 displays the correlation between the planning results and the risk distribution. The color from light to dark reflects the importance of the transmitted service from low to high. The red line represents the high-risk line. It can be seen from Fig. 6(a) that SPRM considers the proximity principle in route planning so that the important services are transmitted on the high betweenness links. It can be seen from Fig. 6(b) that ROM transfers some important services from line 4–6, 6–10, 4–12, and 6–28 to line 2–6, 6–9, and 9–10. ROM reduces the attack risk by distributing important information flows to reliable links to bypassing high-risk links.

The network vulnerabilities of ROM and SPRM under the betweenness priority attack and the key link attack are shown in Fig. 7 and Fig. 8, respectively. It can be seen from Fig. 7 that the vulnerability of SPRM under the betweenness priority attack is significantly higher than that of ROM. When the first 4 edges with the largest betweenness are attacked, the

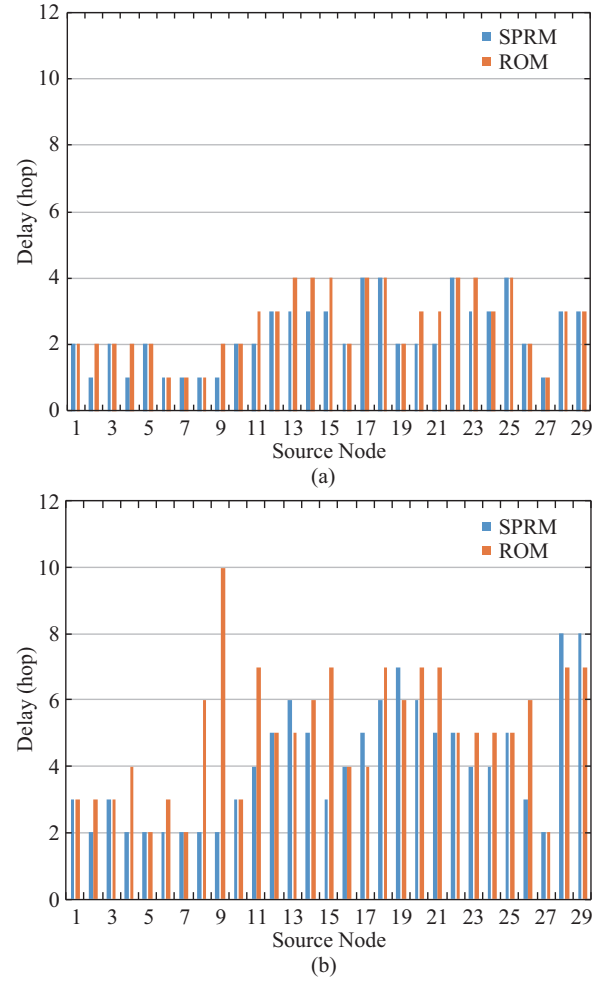


Fig. 5. Delay comparison between SPRM and ROM. (a) Delay comparison of main routing; (b) Delay comparison of alternate routing.

TLPS and the LLPS of SPRM rise rapidly to 60.98% and 60.56%, respectively, which is due to the proximity principle of SPRM making the three largest betweenness links—6–28, 6–10, and 4–12 undertake a large number of services. When these three links are interrupted, the services of nodes 12 to 26 in the bottom half of the network are largely cut off. However, the ROM allocates some important services to 6–9 and 9–10, which improves the information transmission reliability. The corresponding TLPS and LLPS are 31.7% and 33.7%, respectively. When the number of failed links reaches 16, the key links such as 27–28, 10–20, 22–24, 10–17, and 6–7 are interrupted and the relay protection service for power generation node cannot be transmitted to the dispatch center. The load nodes lose energy supply and the two overlap.

Through the simulation results, it is observed that TLPS and LLPS of CPPS are both 0 under $N - 1$ topology, because the main and alternate routing share the risk together. With the increasement of failed links, the curve is roughly a ladder-like distribution, which indicates that a link with a large betweenness transmits many services. Simultaneous failure of some key links results in simultaneous service disruption. Therefore, the large betweenness links are an important factor for system vulnerability, and the number of importance services on these

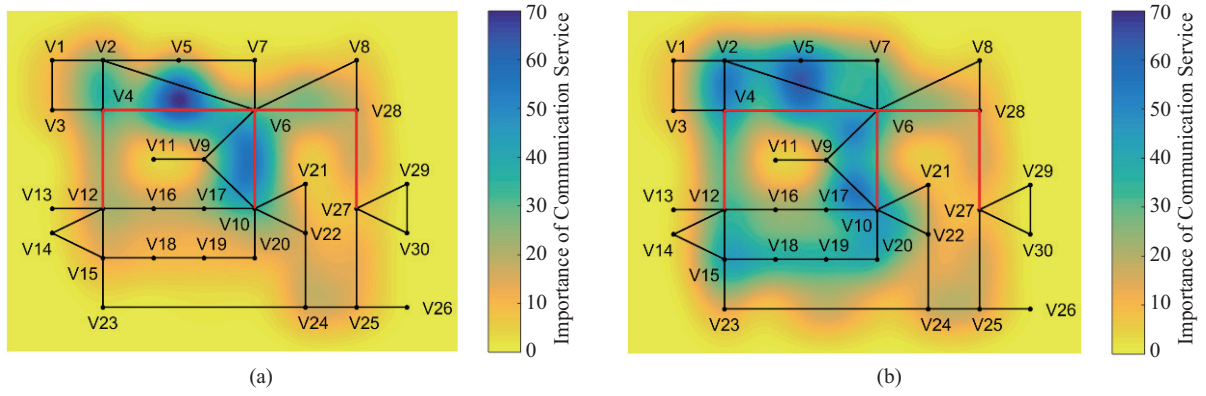


Fig. 6. Comparison of correlation degree between the routing path and the link risk. (a) Correlation degree under SPRM; (b) Correlation degree under ROM.

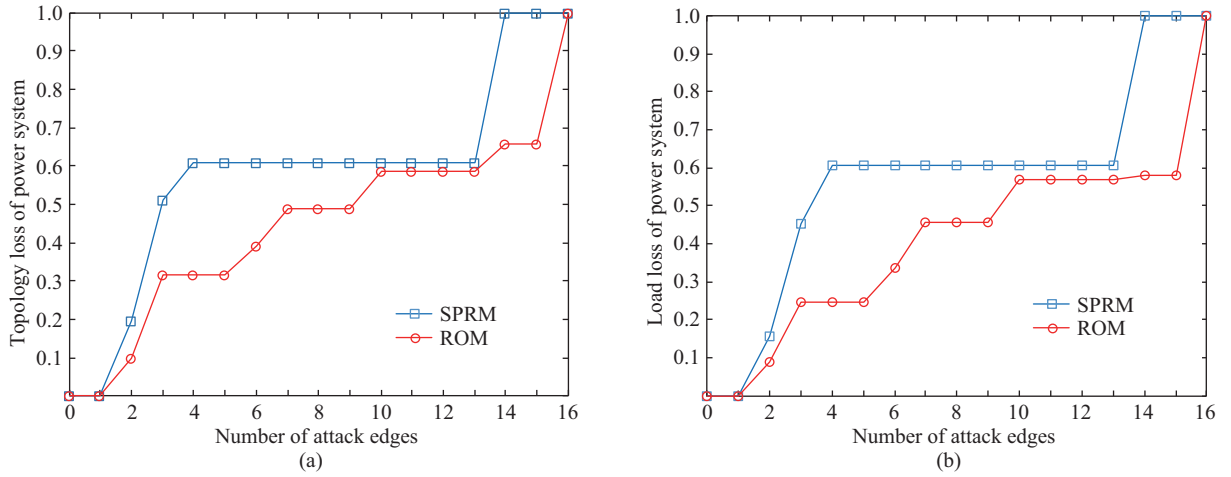


Fig. 7. Network vulnerability under betweenness priority attack. (a) TLPS under betweenness priority attack; (b) LLPS under betweenness priority attack.

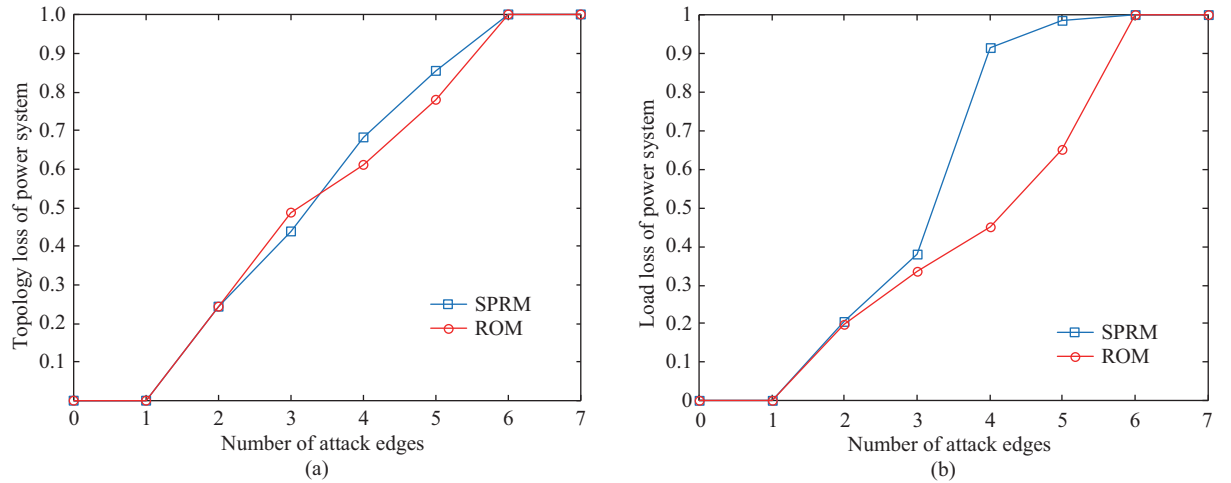


Fig. 8. Network vulnerability under key link attack. (a) TLPS under key link attack; (b) LLPS under key link attack.

links should be transferred appropriately, to enhance system robustness.

According to the simulation results under the key link attack shown in Fig. 8, ROM spreads the main and alternate routings of important services to the link adjacent to the dispatch center, but SPRM concentrates important services on the link 4–6, 6–10, 6–28, and 2–6. When the heavy loaded links are attacked,

the LLPS of SPRM reaches 91.61% while that of ROM is only 45.09%. The planning result of ROM is obviously more reasonable.

VI. CONCLUSION

To better support the safe operation of power systems, we discuss the routing optimization mechanism for power

communication from the interdependence of cyber-physical power systems, in this paper. We reduced operational risks from the following two perspectives: the relative importance of the communication services, which reflects the service impact on the reliable operation of power networks, and the operation risk of information systems, which reflects the impact of link interruption risks on service transmission. For the former, we defined service importance based on its role in power production and transmission. Taking the relay protection service as an example we defined its importance based on the role of its corresponding power node in the integrity and operation of the power grid. For the latter, we constructed the communication vulnerability index to describe service transmission risks based on service distribution and cyber-attacks. Based on the above analysis, we propose a routing optimization model according to the operating characteristics of the information system, utilizing an adaptive genetic algorithm. The case study shows that the system resilience of the proposed ROM is significantly better than that of SPRM under betweenness priority attack, key link attack, and the serious scenarios in $N - 2$ topology, and verifies the effectiveness of our approach.

However, the service categories are simplified in this paper, and only information transmission of relay protection is considered. The assessment of the communication service's impact on the power system only starts from the perspective of circuit breaker maloperation, which is a small probability event in practice. In future research work, we will optimize the routings of multiple communication services from the perspective of practical operation.

REFERENCES

- [1] S. J. Xin, Q. L. Guo, H. B. Sun, B. M. Zhang, J. H. Wang, and C. Chen, "Cyber-physical modeling and cyber-contingency assessment of hierarchical control systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2375–2385, Sep. 2015.
- [2] X. Y. Shi, Y. Li, Y. J. Cao, and Y. Tan, "Cyber-physical electrical energy systems: challenges and issues," *CSEE Journal of Power and Energy Systems*, vol. 1, no. 2, pp. 36–42, Jun. 2015.
- [3] J. W. Cao, Y. X. Wan, H. G. Hua, and G. Yang, "Performance modeling for data monitoring services in smart grid: A network calculus based approach," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 3, pp. 610–618, Sep. 2020.
- [4] 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.
- [5] S. Soltan, M. Yannakakis and G. Zussman, "REACT to cyber attacks on power grids," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 3, pp. 459–473, Sep. 2019.
- [6] D. Wu, F. Ma, M. Javadi, and J. N. Jiang, "Fast screening severe cyber attacks via transient energy-based impact analysis," *CSEE Journal of Power and Energy Systems*, vol. 2, no. 3, pp. 28–34, Sep. 2016.
- [7] K. H. Wu, J. W. Li, Y. Y. Zhu, S. W. Miao, S. X. Zhu, and C. J. Zhou, "Interactive visual analysis on the attack and defense drill of grid cyber-physical systems," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 1, pp. 45–56, Jan. 2021.
- [8] Q. T. Zeng, X. S. Qiu, S. Y. Guo, F. Qi, and L. M. Meng, "Risk balancing based routing mechanism for power communications service," *Journal of Electronics & Information Technology*, vol. 35, no. 6, pp. 1318–1324, Jun. 2014.
- [9] H. Y. Zhu, H. Zang, K. Y. Zhu, and B. Mukherjee, "A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks," *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, pp. 285–299, Apr. 2003.
- [10] H. W. Lee, K. Lee, and E. Modiano, "Maximizing reliability in WDM networks through lightpath routing," *IEEE/ACM Transactions on Networking*, vol. 22, no. 4, pp. 1052–1066, Aug. 2014.
- [11] M. Xia, M. Tornatore, C. U. Martel, and B. Mukherjee, "Risk-aware provisioning for optical WDM mesh networks," *IEEE/ACM Transactions on Networking*, vol. 19, no. 3, pp. 921–931, Jun. 2011.
- [12] W. Cai, H. Yang, F. Xiong, J. E. Li, J. Liu, Z. Y. Zhao, and K. P. Liu, "An optimized service routing allocation method for electric power communication network considering reliability," *Power System Technology*, vol. 37, no. 12, pp. 3541–3545, Dec. 2013.
- [13] B. Li, C. Lu, D. S. Jing, C. Y. Zhu, Y. Sun, and B. Qi, "An optimized routing algorithm with load and risk joint balance in electric communication network," *Proceedings of the CSEE*, vol. 39, no. 9, pp. 2713–2722, May 2019.
- [14] N. Z. Xing, S. Y. Xu, S. D. Zhang, and S. Y. Guo, "Load balancing-based routing optimization mechanism for power communication networks," *China Communications*, vol. 13, no. 8, pp. 169–176, Aug. 2016.
- [15] J. H. Zhao, F. S. Wen, Y. S. Xue, X. Li, and Z. Y. Dong, "Cyber physical power systems: architecture, implementation techniques and challenges," *Automation of Electric Power Systems*, vol. 34, no. 16, pp. 1–7, Aug. 2010.
- [16] T. Y. Chow, F. Chudak, and A. M. Ffrench, "Fast optical layer mesh protection using pre-cross-connected trails," *IEEE/ACM Transactions on Networking*, vol. 12, no. 3, pp. 539–548, Jun. 2004.
- [17] Y. Tang, X. Han, Y. J. Wu, Y. Ju, X. Zhou, and M. Ni, "Electric power system vulnerability assessment considering the influence of communication system," *Proceedings of the CSEE*, vol. 35, no. 23, pp. 6066–6074, Dec. 2015.
- [18] Y. Vardi and C. H. Zhang, "Measures of network vulnerability," *IEEE Signal Processing Letters*, vol. 14, no. 5, pp. 313–316, May 2007.
- [19] I. Mishkovski, M. Biey, and L. Kocarev, "Vulnerability of complex networks," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 1, pp. 341–349, Jan. 2011.
- [20] W. Peng, Z. M. Li, Y. Liux, and J. S. Su, "Assessing the vulnerability of network topologies under large-scale regional failures," *Journal of Communications and Networks*, vol. 14, no. 4, pp. 451–460, Aug. 2012.
- [21] T. Zhou, W. J. Bai, B. H. Wang, Z. J. Liu, and G. Yan, "A brief review of complex networks," *Physics*, vol. 34, pp. 31–36, Jan. 2005.
- [22] Y. D. Tan, X. R. Li, Y. Cai, and Y. D. Zhang, "Critical node identification for complex power grid based on electrical distance," *Proceedings of the CSEE*, vol. 34, no. 1, pp. 146–152, Jan. 2014.
- [23] L. M. Su, C. Z. Hou, Z. J. Dai, and Y. J. Zhang, "Application of Improved genetic algorithm in network fault diagnosis expert system," *Journal of Beijing Institute of Technology*, vol. 12, no. 3, pp. 225–229, Sep. 2003.
- [24] Y. H. Du, J. C. Fang, and C. X. Miao, "Frequency-domain system identification of an unmanned helicopter based on an adaptive genetic algorithm," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 2, pp. 870–881, Feb. 2014.
- [25] J. Mendes, R. Araújo, T. Matias, R. Seco, and C. Belchior, "Automatic extraction of the fuzzy control system by a hierarchical genetic algorithm," *Engineering Applications of Artificial Intelligence*, vol. 29, pp. 70–78, Mar. 2014.
- [26] M. Srinivas and L. M. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 24, no. 4, pp. 656–667, Apr. 1994.
- [27] K. Ooura, K. Kanamaru, R. Matsubara, and S. Ibuki, "Application of a power line maintenance information system using OPGW to the Nishi-Gunma UHV line," *IEEE Transactions on Power Delivery*, vol. 10, no. 1, pp. 511–517, Jan. 1995.
- [28] A. L. Chiu and E. H. Modiano, "Traffic grooming algorithms for reducing electronic multiplexing costs in WDM ring networks," *Journal of Lightwave Technology*, vol. 18, no. 1, pp. 2–12, Jan. 2000.



Baozhong Ti received the bachelor's degree from the Department of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China, in 2016, where he is currently pursuing the Ph.D. degree. His research interests include resilience assessment and failure recovery of cyber-physical power system.



Jianxiao Wang (S'14–M'19) received the double bachelor's degrees in Electrical Engineering, Economic Management and Ph.D. degree from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 2014 and 2019 respectively. He is currently an assistant professor in the Department of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China. He was awarded as the Outstanding Ph.D. Graduate of Tsinghua University, and Junior Fellowships for Advanced Innovation Think-Tank

Programming by China Association for Science and Technology. His research interests include cyber-physical system, energy integration planning, market mechanism design, energy economics and policy.



Ming Zhou (SM'21) received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from North China Electric Power University in 1989, 1992, and 2006, respectively. Since 1992, she has been with the Department of Electrical and Electronic Engineering at North China Electric Power University, where she is currently a professor. Her research interests include power system analysis and reliability, power system economics, HVDC transmission, electricity market and power quality analysis.



Gengyin Li received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from North China Electric Power University, in 1984, 1987 and 1996 respectively. Since 1987, he has been with the Department of Electrical and Electronic Engineering at North China Electric Power University, where he is currently a professor. His research interests include HVDC transmission, analysis and control of power system with renewable energy, power system economics.