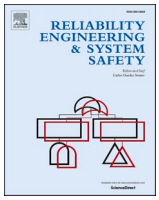
Reliability Engineering and System Safety 256 (2025) 110740

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress



Attack-defense game modeling framework from a vulnerability perspective to protect critical infrastructure systems

Yanfang Wu [a,b,*], Peng Guo [a], Ying Wang [a], Enrico Zio [b,c]

[a] School of Management, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China
[b] Energy Department, Politecnico di Milano, Via La Masa 34, Milano 20156, Italy
[c] MINES Paris-PSL University, CRC, Sophia Antipolis, France

ARTICLE INFO

Keywords:
Critical infrastructure system
Attack-defense game model
Vulnerability
Cascading failure
Risk attitudes

ABSTRACT

Critical infrastructure systems (CISs) are increasingly vulnerable to attacks due to their complex interdependencies. To enhance the resilience of CISs against deliberate attacks, we propose a novel Stackelberg attack-defense game model (ADGM) framework based on game theory and network science. In this framework, the functional vulnerability of CISs that accounts for cascading effects is innovatively considered in the construction of the ADGM. Critical target selection and resource allocation problems are integrated into the strategy model, while cumulative prospect theory (CPT) is applied to evaluate payoffs considering the risk attitudes of agents. The particle swarm optimization (PSO) algorithm is utilized to determine the game equilibrium. The results of applying the ADGM to a power grid highlight the necessity to study resource allocation for infrastructure protection from a vulnerability perspective. Identifying critical targets based on their importance, as determined by the vulnerability metric, provides the foundation for players to develop optimal resource allocation strategies. The analysis of optimal strategies under varying levels of resources for both players reveals the importance of striking an investment balance between enhancing component capacity and safeguarding critical components. The ADGM framework proposed in this paper provides valuable decision-making support for the protection of infrastructure systems.

1. Introduction

Critical infrastructure systems (CISs), such as power grids, water distribution systems and transportation networks, have become the backbone of society operations [1]. However, CISs have grown increasingly susceptible to disruptions due to the potential cascade effect of failures spreading through the complex interdependencies among infrastructures [2]. Recently, the significance and vulnerability of CISs have attracted the attention of terrorist organizations [3]. For instance, in October 2021, a cyberattack on Iran's fuel distribution system paralyzed gas stations across the country. Similarly, in December 2022, a shooting attack on two electrical distribution substations in Moore County, North Carolina, United States, caused power outages that affected up to 40,000 residential and business customers. Therefore, safeguarding critical infrastructures from deliberate attacks has become essential for societal stability and security.

Adversarial game theory has been extensively applied to the study of complex systems' protection under targeted attacks. It provides a mathematical framework to model and analyze strategic interactions between attackers and defenders. The attack-defense game model (ADGM) proposed in recent research has proven effective for exploring the protection of CISs [4]. In the context of ADGM applied to CISs, the strategies and payoffs of individual elements can be influenced by their own actions, those of their adversaries, and the structure of systems [5]. Strategy sets and payoff functions are defined from a system-level perspective. The equilibrium strategies derived from ADGM can provide valuable support for decision-making in protecting CISs.

Payoffs in ADGM are typically defined based on the system value, which can be evaluated by the overall performance of the system or as a cumulative sum of the values of its components. For instance, Zhou et al. [6] evaluated payoffs for players based on the operational cost of the pipeline network from a network perspective. Wu et al. [7] quantified payoffs using the performance response function (PRF) of the power network. Hausken [8] defined utilities based on the system value, calculated as the sum of all targets' values in complex interdependent systems. Zhang et al. [9] considered the additional effects of the simultaneous destruction of related targets on systems to measure the

* Corresponding author at: School of Management, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China.
E-mail address: wwuyanfang@mail.nwpu.edu.cn (Y. Wu).

https://doi.org/10.1016/j.ress.2024.110740
Received 30 May 2024; Received in revised form 17 November 2024; Accepted 3 December 2024
Available online 4 December 2024

Acronyms

CIS	critical infrastructure system
ADGM	attack-defense game model
CPT	cumulative prospect theory
PSO	particle swarm optimization
CSF	contest success function
CPV	cumulative prospect value
CNIV	critical nodes identified based on vulnerability metric
CNIN	critical nodes identified based on network characteristic

Symbols

$G(N, L)$	graph with vertex set N and edge set L
l_{ij}	power line from node i to node j
P	real power injection
θ	voltage phase angles
B	bus susceptance matrix
b_{ij}	susceptance of line l_{ij}
f_{ij}	power flow on line l_{ij}
C_{ij}	capacity of line l_{ij}
$\widetilde{L}_{ij}$	initial load of line l_{ij}
$\widetilde{d}_i$	real power demand
$\overline{d}_i$	required power demand
V	vulnerability of the system
I_i	importance of component i
CT	critical target set
R_a/ R_d	resource available to the attacker/defender
a/d	attack/defense resources of players on targets
p_i	probability of target i being successfully attacked
m	intensity of the contest over component
μ_i/γ_i	proportion of resources invested on each target of the attacker/defender
K	number of critical targets
U	utilities of both players
V_A/V_D	CPVs of attacker and defender
π^-/π^+	decision weight of the value for the loss/gain
w^-/w^+	weighting function for the loss/gain

payoffs in ADGM. These studies quantify payoffs from a system perspective, considering complex interdependencies. However, there is limited research has focused on characterizing strategies and payoffs from a vulnerability perspective, particularly considering the cascading failure effects on CISs.

Vulnerability has been defined and conceptualized in various ways by researchers across different fields. Bier et al. [10] described vulnerability as "any weakness in an asset's or infrastructure's design, implementation or operation that can be exploited by an adversary", as referenced in Risk Analysis and Management for Critical Asset Protection (RAMCAP™) Framework. Yaghlane et al. [11,12] characterized the vulnerability of a system component based on its probability of failing to survive an attack. Fang et al. [13] defined vulnerability as "a measure of system susceptibility to scenarios for more narrowly identifying weak points in the system within the context of a scenario". Bellè et al. [14] viewed vulnerability as the assessment of potential adverse outcomes that may arise within a system when subjected to external strain. Referring to Ref. [14], we define vulnerability as a drop in system functionality under attacks.

CISs operate as complex systems of systems, and their vulnerability is intricately linked to the phenomenon of cascading failures. Cascading failures occur when the failure of a single target leads to the dysfunction of other interconnected components. As these failures propagate, they can eventually cause the collapse of a substantial portion of nodes or even the entire system [15]. Shan [16] indicated that the defender's best response is not solely determined by direct attacks but also by the potential spread of damage from connected networks. Chaoqi et al. [17] proposed an ADGM for CISs that incorporates the cascading effect using the load-capacity model. Huang et al. [18] further analyzed the impact of cascading failures on the game equilibrium of ADGM, suggesting that enhancing the tolerance of components can reduce the differences in attack performance across various strategies. Recent studies on cascading failures in ADGM for CISs often focus on the topological characteristics of networks. However, given the service-oriented nature of CISs, it is essential to propose a novel ADGM framework from a functional vulnerability perspective, considering the effects of cascading failures.

The existing research on strategy models in ADGM for CISs mainly focuses on identifying critical components within CISs. Wu et al. [19] introduced a tri-level defender-attacker-operator game-theoretic model to prioritize components for allocating protective resources. Kuttler et al. [20] extended the tri-level protection-interdiction-restoration model from a single commodity to multiple commodities, aiming to identify the optimal strategies against intelligent attacks. In their models, component states are considered binary under the interaction between players: a component is operational if protected and failed if unprotected under attack [21]. While the binary attacker-defender model provides a foundation for the analysis of ADGM, its applicability in real-world scenarios is limited. In practice, the dynamic nature of component states depends on the degree of attack and defense exerted by players. The strength of attack and defense resource allocation on targets would influence the probability of a successful attack [22].

The research on the resource allocation strategies for CISs has garnered attention in recent years. In this context, the effectiveness of players' resources in influencing the state of components has been assessed by using the contest success function (CSF) [23]. According to the CSF, the investments and payoffs of players are defined based on unit investment cost, contest intensity for each component, and system functionality [24]. Hausken [25] proposed a novel strategic defense and attack model that integrates CSF to determine the optimal allocation of resources between defending and attacking multiple targets. Mo et al. [26] incorporated CSF into a three-stage attack+defend-defend-attack framework to optimize the allocation of attack-defense resources. The ADGM based on a resource allocation perspective can effectively characterize component states by incorporating CSF. However, most research focuses on distributing resources across all system targets, which is not practical for analyzing large-scale and complex CISs. Therefore, considering both target selection and resource allocation within ADGM is crucial for defenders to develop optimal protection strategies on large-scale CISs.

Recent studies on ADGM for the protection of CISs mostly assume that players are risk-neutral [27,28]. In these models, the attacker seeks to maximize disruption impacts on the system by finding the optimal attack strategy within resource constraints, whereas the defender aims to determine the optimal protection strategy to minimize system performance deterioration [29] or maximize system functionality maintenance [30,31]. However, in reality, individuals often depart from risk neutrality when making decisions under uncertainty, with risk-taking behaviors affected by subjective risk attitudes [32]. Zhang et al. [33] indicated that the risk preferences of attackers can influence the optimal allocation in strategic scenarios. Peng et al. [34] incorporated cumulative prospect theory (CPT) into ADGM on interdependent networks, where the objectives of players are to maximize their cumulative prospect values (CPVs). Lin et al. [35] applied CPV as the performance metric instead of vulnerability to study the defense-attack game between M defenders and N attackers. CPT has been widely used to analyze

attack-defense games by considering risk attitudes [36]. Therefore, this paper incorporates the risk attitudes of agents into the ADGM framework to evaluate player payoffs by using CPT.

Based on the above review, existing literature on ADGM for CISs has made significant contributions to the knowledge of CISs protection against intelligent attacks, but several critical limitations remain that need to be addressed: (i) Research on ADGM for CISs mainly focuses on the impact of target failures on the system, with limited consideration of cascading failure effects; (ii) Studies on resource allocation strategies across all components are generally effective for small-scale systems but fail to adequately address the resource limitations in large-scale CISs; (iii) Additionally, most current models overlook the varying risk attitudes of attackers and defenders, which are crucial for accurately predicting and optimizing attack-defense strategies for CISs.

To address these challenges, this paper introduces an innovative attack-defense game model framework that integrates the concept of functional vulnerability. The major contributions of this paper are as follows:

(1) A novel ADGM framework is proposed to analyze optimal defense strategies from a functional vulnerability perspective, taking into account the cascading failure effects.
(2) The critical target selection is incorporated into the resource allocation problem to construct a comprehensive strategy model, which enables the analysis of attack-defense strategies in large-scale CISs.
(3) The contest success function is integrated with cumulative prospect theory to evaluate all potential states of CISs and characterize the payoffs of agents by considering their diverse risk attitudes. The ADGM framework is subsequently applied to a power grid to provide valuable insights for system protection against intelligent attacks.

The remainder of this paper is organized as follows. Section 2 introduces the framework of the Stackelberg attack-defense game model for CISs. Section 3 proposes the vulnerability assessment model considering cascading failures. In Section 4, the Stackelberg attack-defense game model is formulated. Section 5 presents the solution of game equilibrium. Section 6 conducts a case study on the ADGM using the IEEE57-bus system. Finally, Section 7 concludes the paper with a summary of the research findings.

2. The framework of ADGM

In this section, the building blocks of the attack-defense game model framework for CISs are illustrated in Fig. 1. The vulnerability assessment model is proposed to define the strategies and payoffs for the ADGM, and a two-level PSO algorithm is utilized to solve the game equilibrium of the Stackelberg game model.

Step 1 formalizes the process of the functional vulnerability assessment model that considers cascading failures. In this framework, the vulnerability of CISs is viewed as a function of the system's state, external disruptions, and resulting impacts [37]. CISs operate as highly interconnected networks due to their structural interdependencies, where a failure in one component can quickly trigger a chain reaction of failures in connected components. This cascading effect can initiate a domino-like process, potentially leading to the catastrophic collapse of the entire system. Therefore, the impact of cascading failures is considered in this section to assess the system's overall vulnerability. Consequently, the vulnerability assessment model in this section involves four key steps: Initially, the CIS functional model is extracted by collecting initial data from the specific infrastructure system. Subsequently, failure scenarios are identified based on N-1 contingencies or attack-defense interactions. Once the failure scenarios are initialized, cascading failures are then simulated until the system reaches stability. Finally, the vulnerability of the CIS can be quantified using a functional vulnerability metric. This vulnerability assessment model serves as the foundation for constructing the attack-defense game model, which includes both the strategy model and payoff model.

Step 2 proposes the ADGM from a vulnerability perspective, utilizing a Stackelberg game model. In the ADGM, the defender is defined as the owner of CISs and acts as the leader by initially allocating resources to protect targets, aiming to maximize the defense capability of CISs. The attacker, characterized as an intelligent terrorist, allocates offensive resources after observing the defender's choices, intending to maximize the expected damage on CISs. The ADGM consists of two main modules: the strategy model and the payoff model. (1) In the strategy model, player strategies are formulated based on critical component selection

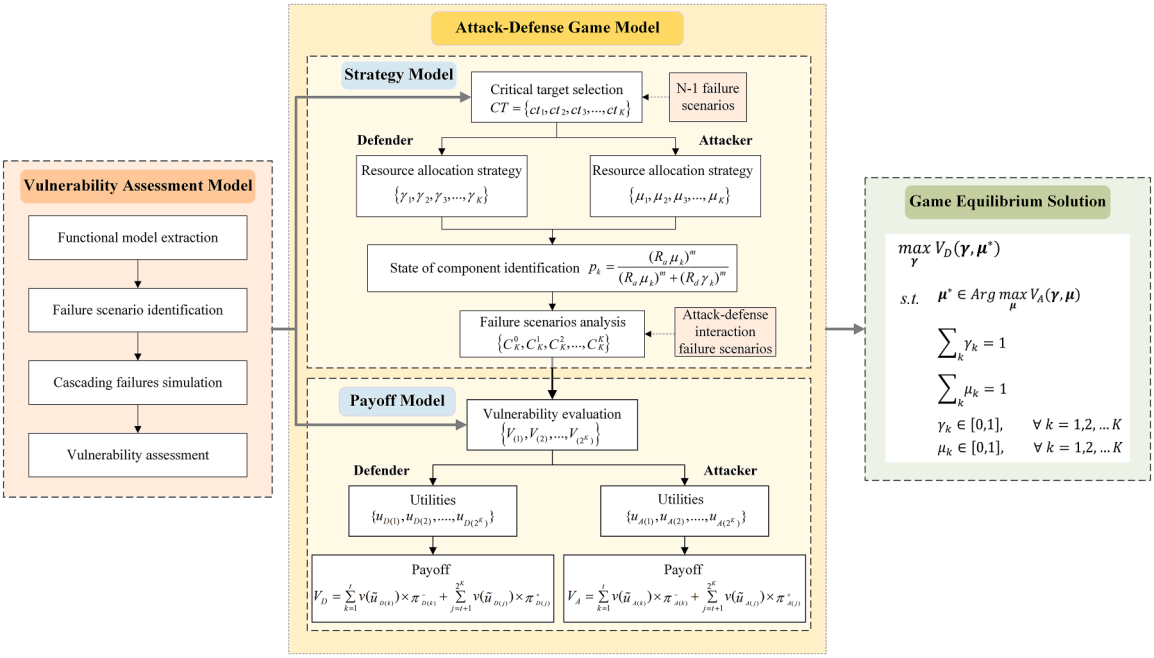


Fig. 1. The main building blocks of ADGM framework.

and resource allocation problems. The CSF is then applied to determine the target states resulting from attack-defense interactions, facilitating the identification of all possible failure scenarios. (2) In the payoff model, the vulnerabilities of CISs are first assessed across various failure scenarios. These vulnerability values are subsequently used to evaluate the utilities of players. Finally, the payoffs are represented by CPVs, which incorporate the risk attitudes of the agents.

Step 3 introduces a method for analyzing the game equilibrium. In this part, the equilibrium solution of the sequential two-player game is formulated as a bi-level optimization problem. A two-level Particle Swarm Optimization (PSO) algorithm is applied to analyze the optimal equilibrium strategies for both players.

3. Vulnerability assessment model

The power grid is a critical infrastructure system (CIS) that supports the operation of various other CISs, including transportation networks, water distribution systems, heating networks, etc. It is essential to both national economic stability and public safety [2]. Additionally, as the interdependencies among its components grow more complex, the power grid becomes increasingly vulnerable to disruptions and cascading failures [38]. These characteristics make it a representative case for demonstrating the applicability of methods used in critical infrastructure research [39,40]. Therefore, this paper focuses on analyzing the attack-defense game in the context of the power grid, aiming to identify the optimal defense resource allocation strategy to defend against intentional attacks.

The structure of the power grid is modeled as a graph $G(N,L)$, where N represents the set of nodes, including generator, substations and transformers, L represents the set of transmission lines. The DC load flow model is applied to calculate the flow through the lines and nodes of the power grid. Additionally, a load-capacity model is used to simulate the cascading failures resulting from line overloads due to flow redistribution [41]. In this research, we have extended the MATCASC tool to assess the functional vulnerability of power grids, building on its proven effectiveness in simulating cascading failures within these systems [42]. Given the critical role of power grids in delivering electricity services, this paper defines their functional vulnerability as a decline in power service levels, reflecting the impact of potential disruptions on the system [43]. The vulnerability assessment architecture is shown in Fig. 2.

3.1. Functional model extraction

To assess the vulnerability of the power grid, the following initial data are required: (*i*) the power grid structure including electrical properties, and (*ii*) a tolerance parameter for transmission lines. These data are used to construct the power grid functional model and initialize its state.

3.1.1. DC load flow equation

DC power flow equations provide a linear relationship between power flowing through the lines and power input into nodes [44]. The DC load flow equation is given by:

$$P = B\theta \tag{1}$$

where P is the set of real power injections; B is the bus susceptance matrix, $B_{ii} = \sum_{j=1}^{d} b_{ij}$ and $B_{ij} = -b_{ij}$, b_{ij} is the susceptance of line l_{ij} and d is the degree of node i; θ is the n-th dimensional vector that contains the voltage angles at each node. Given the bus susceptance matrix B, the voltage angles at each node can be calculated by using:

$$\theta = B^{-1}P \tag{2}$$

The power flow f_{ij} through the transmission line l_{ij} from node i to node j can then be assessed as follows:

$$f_{ij} = b_{ij}\theta_{ij} \tag{3}$$

where θ_{ij} is the voltage phase difference between node i and node j.

3.1.2. Line capacity estimation

The load-capacity model is applied to simulate the failure of power lines due to overloads. The maximum capacity of a line l_{ij} is defined as the maximum power flow that the line can support. It is proportional to the initial load of the node [45]:

$$C_{ij} = \alpha \cdot \widetilde{L}_{ij} \tag{4}$$

where C_{ij} represents the maximum capacity of the line l_{ij}; $\widetilde{L}_{ij}$ is the initial load represented by the flow $\widetilde{f}_{ij}$ on the line l_{ij} evaluated following the DC load flow equation in the initial stable state of the system; α is the tolerance parameter of the transmission line. The line would trip once its load exceeds a certain threshold level under the protection of a circuit breaker. The tolerance parameter α in the load-capacity model plays a crucial role in influencing the system's vulnerability by affecting the extent of cascading failures. Specifically, α impacts the cascading failures within the system by adjusting the load tolerance level of power lines, which in turn affects the overall vulnerability of the power grid.

3.2. Failure scenario identification

Failure scenario identification refers to identifying situations where components may fail under specific environmental conditions. In modeling the ADGM framework, two key types of failure scenarios are considered. First, N-1 failure scenarios are applied to simulate the impact of single-component failures on the performance of power grids, helping to assess component importance and identify critical targets. Second, attack-defense interaction failure scenarios are considered to evaluate the system's response capabilities under various attack-defense interactions, which is crucial for calculating the payoffs for players.

3.3. Cascading failures simulation

The DC load flow equations and the load-capacity model of lines are combined to simulate the cascading failure process. ① Once a failure scenario is identified, the nodes that have been successfully attacked become dysfunctional, causing the failure of the lines connected to these nodes. ② The failed lines are removed from the graph, altering the topology of the power grid. ③ Subsequently, the power flow in the update network is redistributed using the DC load flow model, as described by Eq. (1)-Eq. (3). ④ Overloaded lines are then identified using the load-capacity model and deactivated, the process returns to step ②. Steps ②, ③, and ④ are iterated until no lines are overloaded, indicating that the system has reached a stable state.

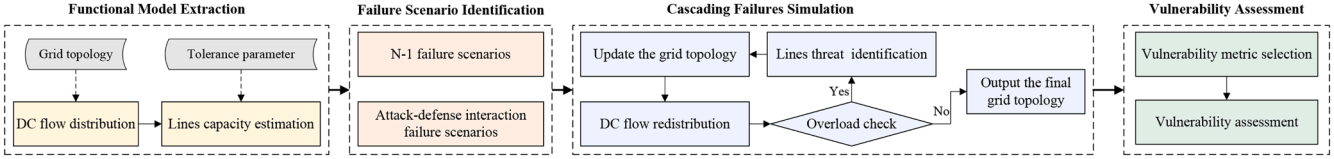


Fig. 2. Vulnerability assessment model of power grids.

3.4. Vulnerability assessment

Once the system reaches stability, the total service level can be calculated by summing the individual demands of all components, denoted as $\sum_{i\in N} d_i$. The initial required power demand is derived from the input data of electric characteristics. Subsequently, the functional vulnerability of the system can be assessed based on the reduction level of system function. The vulnerability metric is formulated as follows:

$$V = 1 - \frac{\sum_{i\in N} \widetilde{d}_i}{\sum_{i\in N} \overline{d}_i} \tag{5}$$

where V is the vulnerability of the system; $\widetilde{d}_i$ and $\overline{d}_i$ represent the real power demand and the required power demand of component i in the system, respectively. The value of $\widetilde{d}_i$ is defined as $min\ (\widehat{d}_i, \overline{d}_i)$, where $\widehat{d}_i$ represents the demand of each node i under the steady state of the power grid. The demand $\widehat{d}_i$ can be calculated as:

$$\widehat{d}_i = \bar{s}_i + \sum_{j=1}^{d} f_{ji} - \sum_{j=1}^{d} f_{ij} \tag{6}$$

where $\bar{s}_i$ represents the supply power of component i under the initial steady state of the power grid.

4. The attack-defense game model

The reliability of the power grid is crucial for national economic stability and public safety. However, due to its highly interconnected structure, the power grid is particularly vulnerable to targeted attacks. To mitigate these risks, this section introduces a Stackelberg attack-defense game model aimed at protecting the power grid. This game model framework comprises both the strategy model and the payoff model for the players. In the game, the defender, representing the power grid's security department, assumes the role of the leader, committing to defense strategies to protect critical power nodes. Conversely, the attacker, representing the terrorists, plays the follower role and formulates optimal attack strategies after observing the defender's actions [46]. Both players have complete information about the power grid and each other, with strategies and payoffs defined based on the vulnerabilities of the power grid.

4.1. Strategy model

Given the finite resources available to the players, attackers are likely to focus on critical components to maximize the expected damage to the CIS. Consequently, defenders need to prioritize protecting these critical nodes by allocating resources accordingly. In this section, we incorporate the selection of critical targets into the resource allocation problem to model the strategy sets for both players.

4.1.1. Targeted target selection

Considering the scale of the system, agents will allocate resources to a limited number K of critical targets for attacking or protecting. In this part, the importance of components is first evaluated based on their impact on system vulnerability. These evaluations are then used to determine the critical target set $CT = \{ct_1, ct_2, \ldots ct_K\}$ [47]. The vulnerability V_i is assessed according to the N-1 scenarios, where the failure of each component i out of N system components is considered. The set of component importances can be denoted as $I = \{I_1, I_2, \ldots I_n\}$, where $I_i = V_i / \sum_i V_i$, for all $i \in N$.

4.1.2. Resource allocation on targets

Both players allocate their resources to each target once the critical target set is confirmed. The vectors $a = \{a_1, a_2, \ldots, a_K\}$ and $d = \{d_1, d_2, \ldots, d_K\}$ represent the allocation of resources across the targets by the attacker and defender, respectively. To determine the status of components, a ratio-based contest success function is applied to evaluate the probability p_k of target k being successfully attacked [22,24]. The probability p_k depends on the resource allocation strategies chosen by the defender and attacker:

$$p_k = \frac{{a_k}^m}{{a_k}^m + {d_k}^m} = \frac{(R_a \mu_k)^m}{(R_a \mu_k)^m + (R_d \gamma_k)^m} \tag{7}$$

where $m \geq 0$ scales the intensity of the contest over component k; R_a and R_d represent the resource available to the attacker and defender, respectively; The variables $\mu = \{\mu_1, \mu_2, \ldots \mu_K\}$ and $\gamma = \{\gamma_1, \gamma_2, \ldots \gamma_K\}$ represent the resource allocation strategies for the attacker and defender, where μ_k and γ_k denote the resource allocation ratios for target k. Table 1 displays the 2^K possible failure scenarios of the game, including the associated probabilities and system vulnerabilities. In Table 1, $i,j,k \in \{1, 2, \ldots K\}$.

4.2. Payoff model

In the interaction between the attacker and the defender, the attacker aims to maximize the expected impact on the system, whereas the defender endeavors to protect the system against the attack. The utilities for both players across various failure scenarios are defined in terms of system vulnerability: the attacker's utility is represented by the vulnerability of the system, and the defender's utility is represented by the robustness of the system. The utilities for both players in each scenario can be assessed using formulas (8) and (9) below:

$$\mathrm{U_A} = V \tag{8}$$

$$\mathrm{U_D} = 1 - V \tag{9}$$

In accordance with the 2^K possible outcomes of the game, the defender's utilities are represented as $u_{D(1)}, u_{D(2),}, \ldots, u_{D(2^K)}$ with probabilities $\mathscr{P}_{D(1)}, \mathscr{P}_{D(2)}, \ldots, \mathscr{P}_{D(2^K)}$. Similarly, the attacker's utilities are denoted as $u_{A(1)}, u_{A(2),}, \ldots, u_{A(2^K)}$ with corresponding probabilities $\mathscr{P}_{A(1)}, \mathscr{P}_{A(2)}, \ldots, \mathscr{P}_{A(2^K)}$. The 2^K utilities of the defender can be ranked as $\widetilde{u}_{D(1)} \leq \ldots \leq \widetilde{u}_{D(t)} \leq 0 \leq \widetilde{u}_{D(t+1)} \leq \ldots \leq \widetilde{u}_{D(2^K)}$, with the associated probabilities $\widetilde{\mathscr{P}}_{D(1)}, \widetilde{\mathscr{P}}_{D(2)}, \ldots, \widetilde{\mathscr{P}}_{D(2^K)}$. Analogously, the ranked utilities of the attacker can be expressed as $\widetilde{u}_{A(1)} \leq \ldots \leq \widetilde{u}_{A(t)} \leq 0 \leq \widetilde{u}_{A(t+1)} \leq \ldots \leq \widetilde{u}_{A(2^K)}$, with the corresponding probabilities denoted as $\widetilde{\mathscr{P}}_{A(1)}, \widetilde{\mathscr{P}}_{A(2)}, \ldots, \widetilde{\mathscr{P}}_{A(2^K)}$.

Following Ref. [35], CPVs incorporate the real-world risk preferences of players, specifically risk-seeking behavior for losses and risk aversion for gains. The values are calculated as weighted sums based on separate decision weight functions and utility functions for losses and

Table 1
Different possible failure scenarios.

Number of failed nodes	Number of cases	Probability ($\mathscr{P}$)	Vulnerability (V)
0	C_K^0	$\prod_{k=1}^{K}(1-p_k)$	0
1	C_K^1	$p_i \cdot \prod_{k=1,k\neq i}^{K}(1-p_k)$	V_i
2	C_K^2	$p_i \cdot p_j \cdot \prod_{\substack{k=1 \\ k\neq i,j}}^{K}(1-p_k)$	$V_{ij},\ i<j$
…	…	…	…
K	C_K^K	$\prod_{k=1}^{K} p_k$	$V_{12\ldots K}$

gains. In this section, the CPVs for the defender and the attacker are calculated using Eq.(10) and Eq.(11), respectively:

$$V_D = \sum_{\tau=1}^{t} \pi_{D(\tau)}^{-} \times v(\widetilde{u}_{D(\tau)}) + \sum_{\tau=t+1}^{2^K} \pi_{D(\tau)}^{+} \times v(\widetilde{u}_{D(\tau)}) \tag{10}$$

$$V_A = \sum_{\tau=1}^{t} \pi_{A(\tau)}^{-} \times v(\widetilde{u}_{A(\tau)}) + \sum_{\tau=t+1}^{2^K} \pi_{A(\tau)}^{+} \times v(\widetilde{u}_{A(\tau)}) \tag{11}$$

where $\pi_{D(\tau)}^{-}/\pi_{A(\tau)}^{-}$ represents the decision weight for the value of losses and $\pi_{D(\tau)}^{+}/\pi_{A(\tau)}^{+}$ denotes the decision weight for the value of gains. The utility function $v(u)$ can be represented by Eq.(12):

$$v(u) = \begin{cases} u^g & u \geq 0 \\ -\lambda(-u)^l & u < 0 \end{cases} \tag{12}$$

where g and l are exponent parameters representing the risk aversion over gains and risk seeking over losses, respectively, both of them belong to the interval [0,1]; λ is the loss-aversion factor and $\lambda > 1$. Further explanations of these parameters can be found in reference [35].

The decision weight $\pi_{(\tau)}^{-}$ can be calculated as:

$$\pi_{(\tau)}^{-} = w^{-}\left(\sum_{h=1}^{\tau} \widetilde{\mathscr{P}}_{(h)}\right) - w^{-}\left(\sum_{h=1}^{\tau-1} \widetilde{\mathscr{P}}_{(h)}\right) \tag{13}$$

for $\tau \geq 2$. When $\tau = 1$, $\pi_{(1)}^{-} = w^{-}(\widetilde{\mathscr{P}}_{(1)})$, $\pi_{(\tau)}^{+}$ is written as:

$$\pi_{(\tau)}^{+} = w^{+}\left(\sum_{h=\tau}^{2^K} \widetilde{\mathscr{P}}_{(h)}\right) - w^{+}\left(\sum_{h=\tau+1}^{2^K} \widetilde{\mathscr{P}}_{(h)}\right) \tag{14}$$

for $\tau \leq 2^K - 1$. When $\tau = 2^K$, $\pi_{(2^K)}^{+} = w^{+}\left(\widetilde{\mathscr{P}}_{(2^K)}\right)$.

The weighting function for losses $w^{-}(\mathscr{P})$ and gains $w^{+}(\mathscr{P})$ are obtained as:

$$w^{-}(\mathscr{P}) = \frac{\mathscr{P}^{\delta}}{[\mathscr{P}^{\delta} + (1-\mathscr{P})^{\delta}]^{1/\delta}} \tag{15}$$

$$w^{+}(\mathscr{P}) = \frac{\mathscr{P}^{\chi}}{[\mathscr{P}^{\chi} + (1-\mathscr{P})^{\chi}]^{1/\chi}} \tag{16}$$

where δ and χ are weighting parameters.

5. Game equilibrium solution

In the Stackelberg game model, the defender acts as the leader who chooses the defense strategy first. The attacker, as the follower, commits to a strategy after observing the defender's actions. Both players choose their resource allocation strategies to maximize their CPVs as shown in Eq.(10) and Eq.(11). Given the resource strategies for players $\mu = \{\mu_1, \mu_2, \ldots \mu_K\}$ and $\gamma = \{\gamma_1, \gamma_2, \ldots \gamma_K\}$, the optimal strategy model for the defender and attacker can be formulated as a two-phase optimization model as follows:

$$\max_{\gamma} V_D(\gamma, \mu^*) \tag{17}$$

$$s.t. \quad \mu^* \in \arg\max_{\mu} V_A(\gamma, \mu) \tag{18}$$

$$\sum_{k=1}^{K} \gamma_k = 1 \tag{19}$$

$$\sum_{k=1}^{K} \mu_k = 1 \tag{20}$$

$$\gamma_k \in [0,1], \ \forall\, k = 1, 2, \ldots K \tag{21}$$

$$\mu_k \in [0,1], \ \forall\, k = 1, 2, \ldots K \tag{22}$$

Eq. (17) represents the objective function for maximizing the expected payoff for the defender. Eq. (18) is utilized to determine the optimal strategies for the attacker based on the specific strategy adopted by the defender. Eqs. (19)–(22) are the constraints on the resource allocation ratios.

It is important to note that the model is an NP-Hard problem, and it can be solved using methods employed for bi-level programming problems. In this paper, the two-level PSO algorithm [48] is utilized to analyze the optimal attack-defense strategies. The equilibrium solution-solving process is illustrated in Fig. 3, and the corresponding equilibrium solution is calculated and demonstrated in Algorithm 1.

6. Case study

This section presents a numerical experiment conducted on the power system using the IEEE57-bus test system [49]. The node data for the IEEE57-bus test system in the initial steady state is shown in Appendix-Table A. It is assumed that 10 % of the nodes are selected as targets. The model parameters are set according to [35], as detailed in Table 2.

6.1. Targeted node selection

In this section, the characteristics of the top 5 critical nodes, identified based on the vulnerability metric with α=1.5, are presented in Table 3.

It can be observed that when cascading failure is considered, identifying important nodes based solely on network topology and initial electrical significance is insufficient. For instance, although the importance of node 14, as determined by degree, betweenness, and electrical power, is lower, its failure has a more significant impact on the power grid compared to node 15. This highlights the necessity to assess the importance of nodes based on the vulnerability metric. The vulnerability $V_{(i)}$ for the critical target set CT={8,14,15,2,16} under different tolerance values α are shown in Fig. 4.

As shown in Fig. 4, the vulnerabilities caused by node failures decrease at different rates as the tolerance parameter α increases, leading to differences in the importance of the same node under different α values. Whereas the vulnerability of the system tends to decrease with increasing α, the relationship between $V_{(i)}$ and α is not strictly negatively correlated. This is attributed to the fact that the vulnerability of power grids can be influenced not only by the structure characteristics but also by the DC load flow model of the power grid. Therefore, it is essential to identify critical nodes based on different α values.

6.2. Game equilibrium analysis

This section explores the influence of two different critical target selection methods on the game equilibrium strategies and expected payoffs: critical node selection based on node importance measured by static network characteristic (CNIN) and dynamic vulnerability metric (CNIV). For CNIN, node importance is evaluated by combining the initial topology and energy flow metrics, denoted as $I_{N(i)}$, where $I_{N(i)} = (I_{D(i)} + I_{B(i)} + I_{E(i)})/3$. For CNIV, the dynamic vulnerability metric is applied to evaluate node importance, denoted as $I_{V(i)}$. In this section, it is assumed that the total resources available to both the attacker and the defender are equal. The expected payoffs for the attacker and defender under different critical target identification methods are shown in Fig. 5. The critical target set identified based on CNIN is CT={1,8,9,13,15}, whereas the critical targets identified using CNIV vary for different values of α, as shown in Table 4.

As shown in Fig. 5, the expected payoffs of the attacker based on

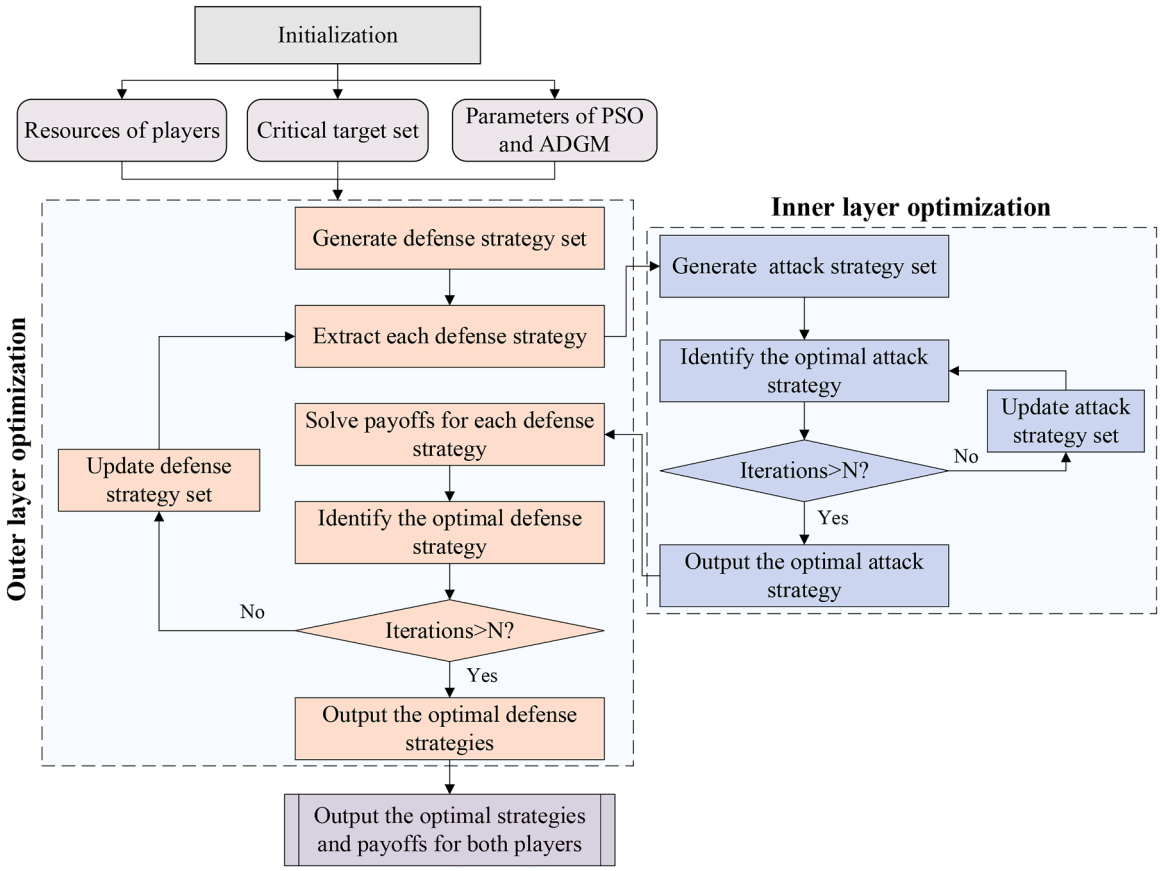


Fig. 3. Equilibrium solution-solving process.

CNIV are generally higher than those based on CNIN, indicating that resource allocation based on CNIV can achieve higher cumulative payoffs for the attacker. Therefore, to counteract the attacker's optimal strategy and maximize their cumulative payoffs, the defender must prioritize protecting these critical nodes. The equilibrium strategies for both players under different values of α are represented in Table 4.

Comparing Fig. 5 and Table 4, we observe that when the resources of the defender are equal to those of the attacker, the equilibrium payoff for the attacker does not strictly decrease with an increase in the tolerance parameter α, nor does the defender's payoff strictly increase. This is due to the fact that both critical targets and utility values for the players can be influenced by functional vulnerabilities, which vary with different α values. To maximize the expected payoffs, infrastructure managers should comprehensively design and allocate investments to enhance both component load capacity and resistance to attacks.

6.3. The influence of different resource investments on game equilibrium

Understanding the relationship between game equilibrium payoffs and resource investments is critical for analyzing the dynamic interactions between the attacker and defender on CISs. By identifying the critical target sets based on CNIV with varying tolerance parameter α, this section explores how the game equilibrium shifts as a function of the defender-attacker resource ratio, denoted as $\rho = R_d/R_a$. The expected equilibrium payoffs for both players across different values of ρ for a given α are shown in Fig. 6.

Fig. 6 illustrates that, for a given α, the attacker's expected payoffs generally decrease as ρ increase, whereas the defender's expected payoffs increase correspondingly. This trend suggests that as the defender's resources increase relative to the attacker's, the defender gains a strategic advantage, leading to higher payoffs, whereas the attacker's benefits diminish. The inverse correlation between the payoffs of the attacker and defender highlights the adversarial nature of their interaction.

Comparing different values of tolerance parameter α, it is evident that lower α values are associated with higher initial payoffs for the attacker, although this advantage wanes as ρ increases. Conversely, higher α values tend to favor the defender, providing consistent advantages as ρ increases. Specifically, lower α values aggravate the vulnerability of CISs by amplifying the effects of cascading failures, which initially benefit the attacker. In contrast, higher α values can mitigate the propagation of failures, providing a consistent defensive advantage.

Fig. 6(b) illustrates that in scenarios where α>1.5 (expect for α=1.7), the defender can achieve a high payoff with relatively minimal investment. In these scenarios, the parameter ρ has a limited impact on equilibrium payoffs. Therefore, for α>1.5, setting α=1.6 is recommended when designing the power grid. In contrast, for $\alpha \leq 1.5$, the defender's payoffs increase significantly as ρ increases, and this improvement becomes more pronounced as α increases. Thus, the relationships between equilibrium strategies and ρ for α=1.4 and α=1.5 are analyzed in detail, with results presented in Figs. 7 and 8.

As illustrated in Figs. 7 and 8, the critical target sets are identified as CT={8,3,4,15,2} for α=1.4 and CT={8,14,15,2,16} for α=1.5. In both cases, when ρ<1, the attacker has a resource advantage and focuses on high-value target node 8. The defender, constrained by limited resources, prioritizes protecting this node. When ρ>1, the defender's resources become more abundant, forcing the attacker to allocate resources across multiple targets to counter the comprehensive defense. In this scenario, the defense strategy continues to focus on the critical node, whereas the attack strategy becomes more dispersed. These results highlight the importance of designing an appropriate tolerance parameter α and implementing effective resource allocation strategies to protect CISs against intelligent attacks.

Algorithm 1
Algorithm pseudocode to determine the equilibrium solution of ADGM.

Input:
Algorithm parameters:
Population size N, Number of iterations G, Dimensionality of the search space D, Inertia weight w,
Acceleration coefficients c_1 and c_2, Upper and lower bounds for position and velocity
Fitness functions:
Outer layer ($f1$): $-V_D$ (V_D see formula (10))
Inner layer ($f2$): $-V_A$ (V_A see formula (11))
Fitness function parameters: K, m, δ, χ, λ, l, g

Output:
The optimal strategy and payoff for the defender g^1_{best} and $f1\left[g^1_{best}\right]$
The optimal strategy and payoff for the attacker g^2_{best} and $f2\left[g^2_{best}\right]$

Begin:
Step 1. Outer Layer: Initialize particle positions (P1) and velocities (V1)
Generate the initial particle swarm (defense strategy set) and deliver it to the inner layer
Step 2. Inner Layer PSO: (Output the optimal attack strategy set against the defense strategy set)
for each particle i *in* P1
Initialize particle positions (P2) and velocities (V2)
Generate the initial particle swarm (attack strategy set)
for each particle j in P2
Evaluate the fitness function value ($f2$) of particle j
Assign p^2_j as p^2_{best} and $f2\left[p^2_j\right]$ as $f2\left[p^2_{best}\right]$
end for
$f2\left[g^2_{best}\right] = min\left\{f2\left[p^2_{best}\right]\right\}$, $g^2_{best} = argmin\left\{f2\left[p^2_{best}\right]\right\}$
while not stop
for each particle j in P2
Update the position and velocity of particle j
Evaluate the fitness function $f2\left[p^2_j\right]$ of particle j
If $f2\left[p^2_j\right] < f2\left[p^2_{best}\right]$, $p^2_{best} = p^2_j$
If $f2\left[p^2_{best}\right] < f2\left[g^2_{best}\right]$, $g^2_{best} = p^2_{best}$
end for
end while
print g^2_{best} and $f2\left[g^2_{best}\right]$
end for
Step 3. Outer Layer: Evaluate and identify the solution with best fitness
for each particle i *in* P1
Evaluate the fitness function value ($f1$) of particle i
Assign p^1_i as p^1_{best} and $f\left[p^1_i\right]$ as $f\left[p^1_{best}\right]$
end for
$f1\left[g^1_{best}\right] = min\left\{f1\left[p^1_{best}\right]\right\}$, $g^1_{best} = argmin\left\{f1\left[p^1_{best}\right]\right\}$
Step 4. Outer Layer: Update the best know position of particles and the swarm
while not stop
for each particle i in P1
Update the position and velocity of particle i
Return to Step 2
Evaluate the fitness function $f1\left[p^1_i\right]$ of particle i
If $f1\left[p^1_i\right] < f1\left[p^1_{best}\right]$, $p^1_{best} = p^1_i$
If $f1\left[p^1_{best}\right] < f1\left[g^1_{best}\right]$, $g^1_{best} = p^1_{best}$
end for
end while
print g^1_{best} and $f1\left[g^1_{best}\right]$, g^2_{best} and $f2\left[g^2_{best}\right]$
End

7. Conclusions

CISs are becoming increasingly vulnerable to deliberate attacks due to their critical role in urban operations. Although extensive research has explored attack-defense game models to protect CISs from potential intelligent attacks, several limitations remain unaddressed. The current research on ADGM for CISs mainly focuses on the identification of critical components or resource allocation across components, with relatively little attention given to the cascading effects of failures on CISs' performance and the impact of varying risk attitudes on strategic interactions.

To address these limitations, this paper introduces a novel ADGM framework aimed at enhancing the protection of CISs. Our work is innovative in that integrates a functional vulnerability perspective that considers cascading failure effects into the strategy and payoff models. Additionally, the proposed framework introduces a new strategy model that combines critical target selection with resource allocation, as well as an improved payoff model that utilizes CPT to account for players' risk attitudes. Furthermore, a two-level PSO algorithm is employed to solve the equilibrium of the Stackelberg attack-defense game. A case study involving the IEEE57 power grid is conducted to validate the feasibility and effectiveness of the proposed model. The simulation results demonstrate that:

(1) Identifying critical targets from a functional vulnerability perspective is crucial for the strategy model. Relying on the static network characteristics (CNIN) for optimal resource allocation tends to underestimate the attacker's expected payoffs and overestimate those of the defender. This discrepancy highlights the importance of prioritizing targets identified based on the dynamic vulnerability metric (CNIV).
(2) Cascading failures significantly affect the selection of critical targets and the expected payoffs of both players by influencing the functional vulnerability of CISs. When the attack and defense resource investments are equal, the defender's equilibrium payoff does not increase strictly with a higher tolerance parameter α. Therefore, it is crucial for the defender to determine an appropriate α value to effectively mitigate cascading failures.
(3) For given tolerance parameter α values, the attacker's expected payoffs generally decrease as the defender-attacker resource ratio ρ increases, whereas the defender's payoffs increase accordingly. However, when α is sufficiently large, the impact of ρ on these payoffs diminishes.
(4) It is crucial for the defender to balance investments between enhancing component capacity to mitigate cascading failures and strengthening defenses for critical components to reduce the impact of attacks.

This study applies the ADGM framework to the power grid to validate its applicability and analyzes the local cascading impacts on optimal resource allocation strategies within a single infrastructure system. However, with the development of Internet of Things (IoT) technology, infrastructure systems are becoming increasingly interdependent, such as power-gas and power-water networks. In future work, it would be valuable to extend the ADGM framework to other

Table 2
Model parameters settings.

Target number (K)	Contest success function (p_k)	Weighting function ($w(\mathscr{P})$)		Utility function ($v(u)$)		
		Losses	Gains	Losses	Losses	Gain
$K = 5$	$m = 1$	δ=0.69	χ=0.61	λ=2.25	l = 0.92	g = 0.89

Table 3
Characteristics of the top 5 critical nodes.

Number	ID of Node	I_V	I_D	I_B	I_E
1	8	0.072	0.019	0.030	0.156
2	14	0.061	0.019	0.008	0.029
3	15	0.050	0.032	0.032	0.080
4	2	0.050	0.013	0.002	0.049
5	16	0.043	0.013	0.001	0.015

Notations: $I_{V(i)} = \frac{V_i}{\sum_{j=1}^{N} V_j}$, $I_{D(i)} = \frac{D_i}{\sum_{j=1}^{N} D_j}$, $I_{B(i)} = \frac{B_i}{\sum_{j=1}^{N} B_j}$, $I_{E(i)} = \frac{E_i}{\sum_{j=1}^{N} E_j}$. D, B, and E represent the degree, betweenness, and initial electric power of nodes, respectively.

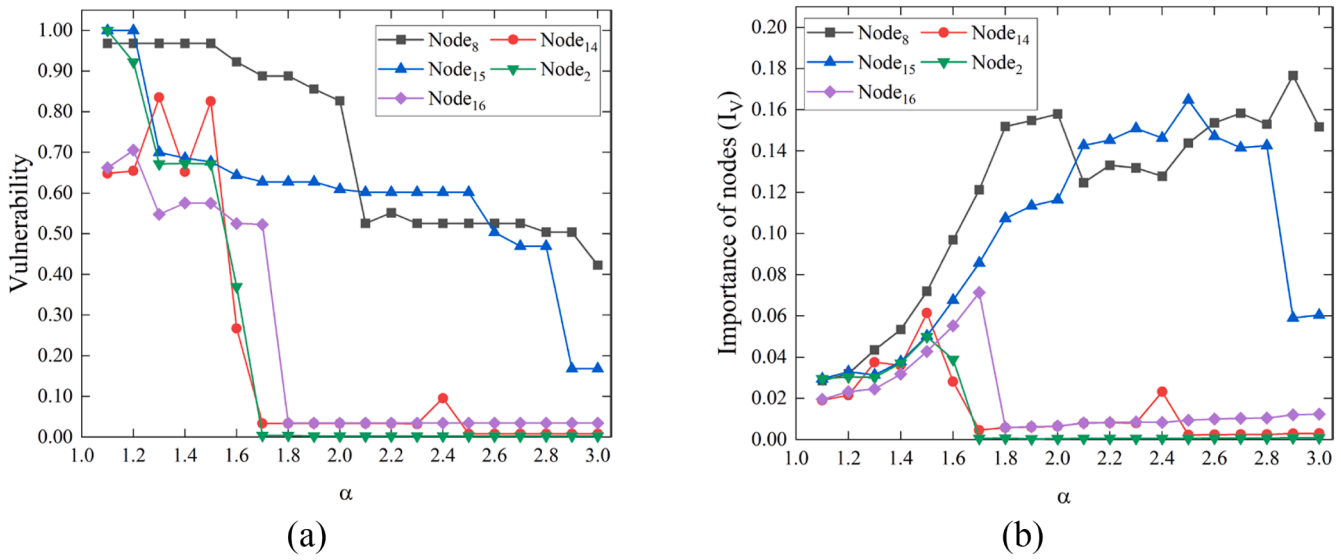


Fig. 4. The vulnerability of the system (a) and importance of nodes (b) according to N-1 scenarios under different values of the tolerance parameter α.

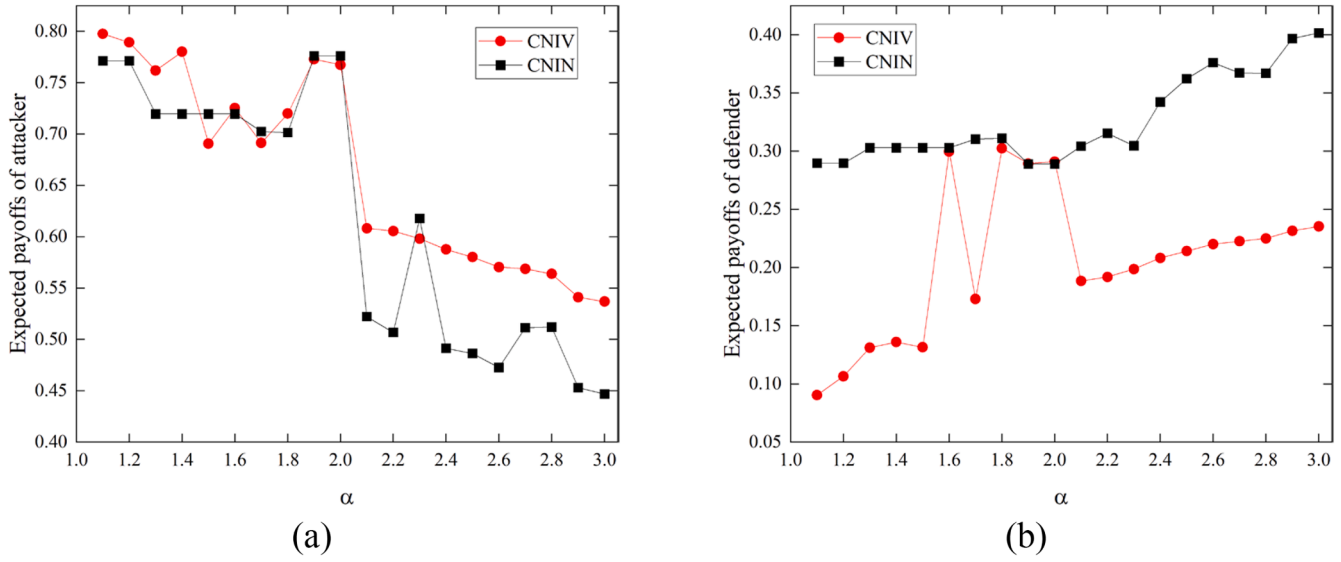


Fig. 5. Expected payoffs of the attacker (a) and defender (b) based on CNIV and CNIN.

Table 4
The Equilibrium strategies and payoffs of both players under different $\boldsymbol{\alpha}$.

α	Optimal Strategy			Expected Payoffs	
	Critical Targets	Resource allocation ratio		Defender	Attacker
		Defender	Attacker		
1.1	2,15,17,8,3	0.4,0.024,0.355,0.191,0.031	0,0.296,0.037,0.42,0.247	0.088	0.832
1.2	15,8,3,2,45	0.121,0.466,0.095,0.133,0.185	0.12,0.212,0.272,0.25,0.145	0.105	0.747
1.3	8,3,14,7,17	0.259,0.033,0.035,0.323,0.35	0.251,0.273,0.213,0.006,0.257	0.117	0.786
1.4	8,3,4,15,2	0.511,0.015,0.009,0.077,0.389	0.38,0.323,0.169,0.125,0.002	0.156	0.786
1.5	8,14,15,2,16	0.461,0.132,0.077,0.13,0.201	0.261,0.153,0.277,0.181,0.129	0.143	0.699
1.6	8,15,3,1,16	0.304,0.21,0.189,0,0.298	0.26,0.26,0.248,0.006,0.227	0.299	0.727
1.7	8,15,1,17,16	0.28,0.157,0.078,0.261,0.224	0.21,0.2,0.228,0.188,0.174	0.181	0.652
1.8	8,15,1,9,17	0.22,0.276,0,0.363,0.141	0.21,0.333,0.089,0.036,0.331	0.300	0.724
1.9	8,1,15,9,17	0.348,0,0.21,0.357,0.085	0.036,0.445,0.363,0,0.156	0.289	0.776
2	8,1,15,9,17	0.22,0,0.226,0.415,0.14	0.307,0.377,0.306,0,0.01	0.289	0.773

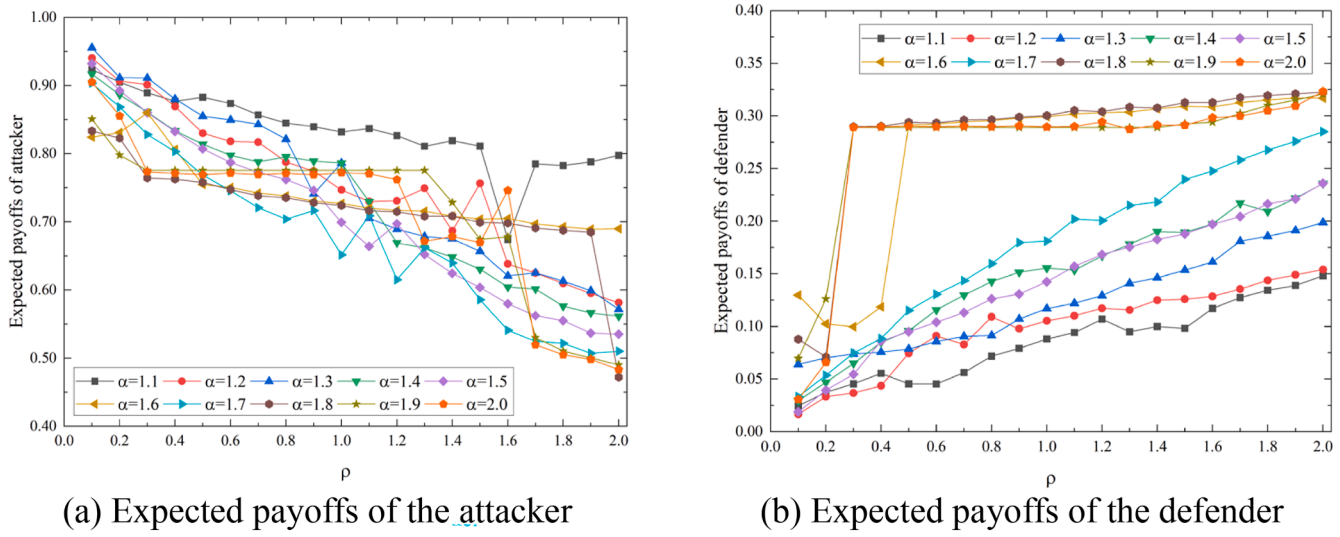


Fig. 6. Expected payoffs of players under different resource investment scenarios.

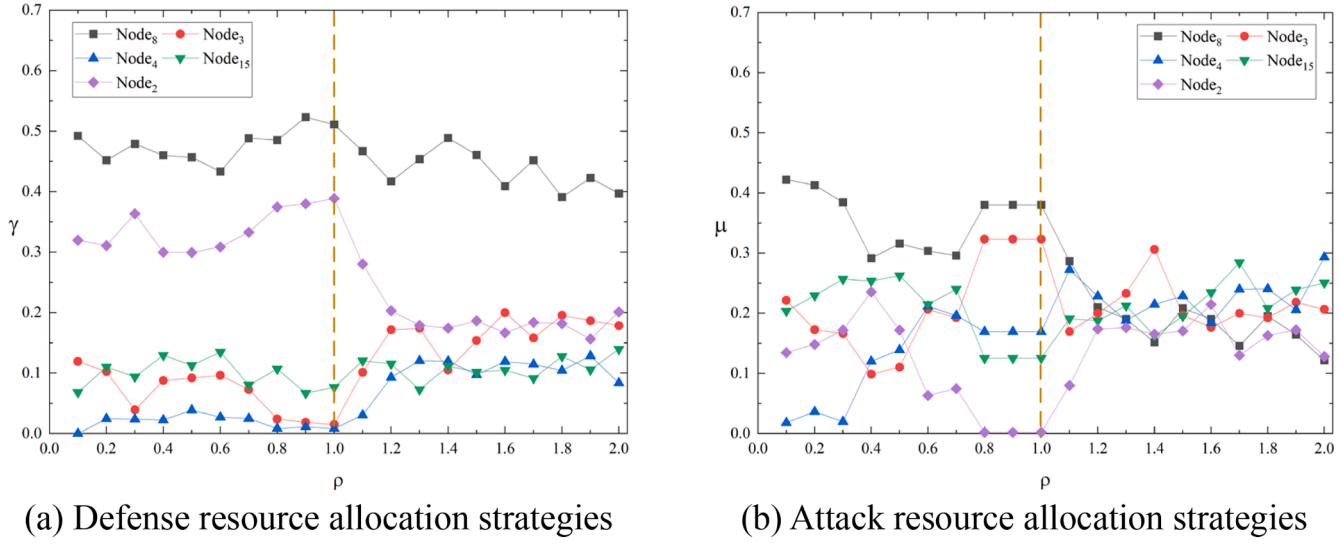


Fig. 7. Equilibrium strategies under different resource investment scenarios when α=1.4.

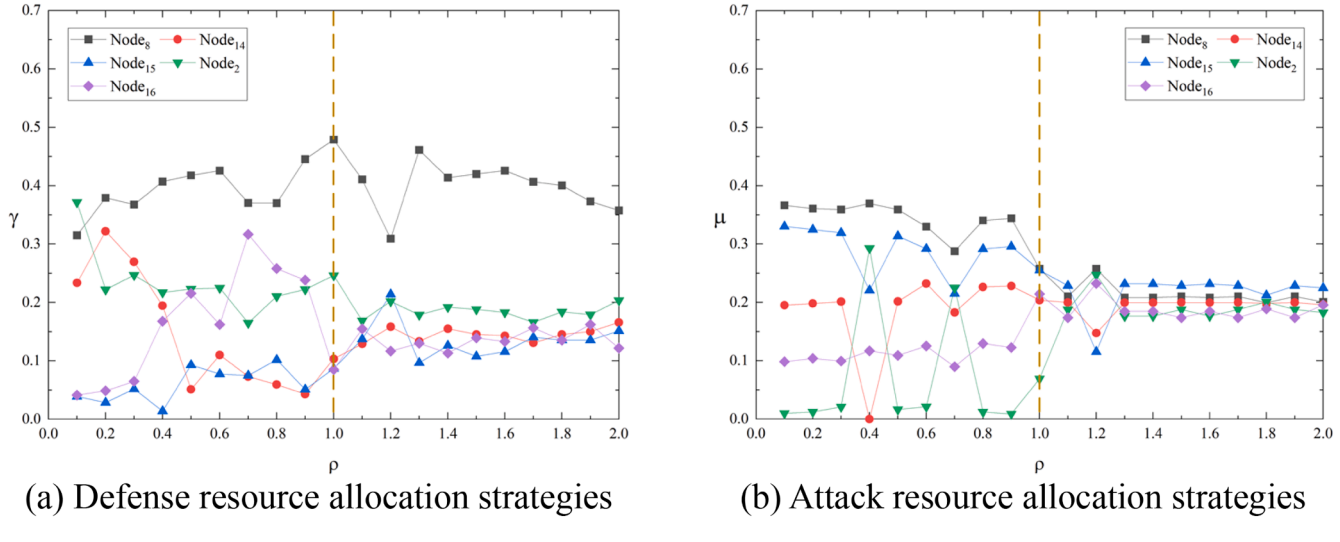


Fig. 8. Equilibrium strategies under different resource investment scenarios when α=1.5.

interdependent critical infrastructure systems, including natural gas and water distribution networks, to explore optimal defense strategies across interdependent CISs.

CRediT authorship contribution statement

Yanfang Wu: Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Peng Guo:** Supervision, Funding acquisition, Conceptualization. **Ying Wang:** Writing – review & editing, Software, Formal analysis. **Enrico Zio:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge the support of the National Natural Science Foundation of China (Grant No. 72171195).

Appendix

Table A

Table A
Node data in IEEE57-bus test system (Initial Steady State) [Units: MW].

Node	Supply ($\bar{s}_i$)	Demand ($\bar{d}_i$)	Node	Supply ($\bar{s}_i$)	Demand ($\bar{d}_i$)	Node	Supply ($\bar{s}_i$)	Demand ($\bar{d}_i$)
1	450.8	55	20	0	2.3	39	0	0
2	0	3	21	0	0	40	0	0
3	40	41	22	0	0	41	0	6.3
4	0	0	23	0	6.3	42	0	7.1
5	0	13	24	0	0	43	0	2
6	0	75	25	0	6.3	44	0	12
7	0	0	26	0	0	45	0	0
8	450	150	27	0	9.3	46	0	0
9	0	121	28	0	4.6	47	0	29.7
10	0	5	29	0	17	48	0	0
11	0	0	30	0	3.6	49	0	18
12	310	377	31	0	5.8	50	0	21
13	0	18	32	0	1.6	51	0	18
14	0	10.5	33	0	3.8	52	0	4.9
15	0	22	34	0	0	53	0	20
16	0	43	35	0	6	54	0	4.1
17	0	42	36	0	0	55	0	6.8
18	0	27.2	37	0	0	56	0	7.6
19	0	3.3	38	0	14	57	0	6.7

Data availability

Data will be made available on request.

References

[1] Tang D, Fang YP, Zio E. Vulnerability analysis of demand-response with renewable energy integration in smart grids to cyber attacks and online detection methods. Reliab Eng Syst Safe 2023;235:109212.
[2] Tu H, Gu F, Zhang X, Xia Y. Robustness analysis of power system under sequential attacks with incomplete information. Reliab Eng Syst Saf 2023;232:109048.
[3] Fang Y, Sansavini G. Optimizing power system investments and resilience against attacks. Reliab Eng Syst Saf 2017;159:161–73.
[4] Hausken K. Fifty years of operations research in defense. Eur J Oper Res 2024;318:355–68.
[5] Hausken K, Levitin G. Review of systems defense and attack models. Int J Perform Eng 2012;8:13.
[6] Zhou J, Zhu J, Liang G, Ma J, He J, Du P, et al. Three-layer and robust planning models to evaluate the strategies of defense layer, attack layer, and operation layer for optimal protection in natural gas pipeline network. Reliab Eng Syst Saf 2024;249:110196.
[7] Wu Y, Chen Z, Dang J, Chen Y, Zhao X, Zha L. Allocation of defensive and restorative resources in electric power system against consecutive multi-target attacks. Reliab Eng Syst Saf 2022;219:108199.
[8] Hausken K. Defence and attack of complex interdependent systems. J Oper Res Soc 2019;70:364–76.
[9] Zhang X, Ding S, Ge B, Xia B, Pedrycz W. Resource allocation among multiple targets for a defender-attacker game with false targets consideration. Reliab Eng Syst Safe 2021;211:107617.
[10] Bier VM, Cox LA, Azaiez MN. Why both game theory and reliability theory are important in defending infrastructure against intelligent attacks. In: Bier VMM, Azaiez MN, editors. Game theoretic risk analysis of security threats. Boston, MA: Springer US; 2009. p. 1–11.
[11] Yaghlane AB, Azaiez MN, Mrad M. System survivability in the context of interdiction networks. Reliab Eng Syst Saf 2019;185:362–71.
[12] Yaghlane AB, Azaiez MN. Systems under attack-survivability rather than reliability: concept, results, and applications. Eur J Oper Res 2017;258:1156–64.
[13] Fang YP, Sansavini G, Zio E. An optimization-based framework for the identification of vulnerabilities in electric power grids exposed to natural hazards. Risk Anal 2019;39:1949–69.
[14] Bellè A, Zeng Z, Duval C, Sango M, Barros A. Modeling and vulnerability analysis of interdependent railway and power networks: application to British test systems. Reliab Eng Syst Saf 2022;217:108091.
[15] Shen Y, Yang H, Ren G, Ran B. Model cascading overload failure and dynamic vulnerability analysis of facility network of metro station. Reliab Eng Syst Saf 2024;242:109711.
[16] Shan XG, Zhuang J. A game-theoretic approach to modeling attacks and defenses of smart grids at three levels. Reliab Eng Syst Saf 2020;195:106683.
[17] Chaoqi F, Yangjun G, Jilong Z, Yun S, Pengtao Z, Tao W. Attack-defense game for critical infrastructure considering the cascade effect. Reliab Eng Syst Saf 2021;216:107958.
[18] Huang Y, Wu J, Tse CK, Zheng Z. Sequential attacker-defender game on complex networks considering the cascading failure process. Ieee Trans Comput Soc Syst 2021.
[19] Wu Y, Chen Z, Gong H, Feng Q, Chen Y, Tang H. Defender–attacker–operator: tri-level game-theoretic interdiction analysis of urban water distribution networks. Reliab Eng Syst Saf 2021;214:107703.
[20] Kuttler E, Ghorbani-Renani N, Barker K, González AD, Johansson J. Protection-interdiction-restoration for resilient multi-commodity networks. Reliab Eng Syst Saf 2024;242:109745.

[21] Ghorbani-Renani N, González AD, Barker K, Morshedlou N. Protection-interdiction-restoration: tri-level optimization for enhancing interdependent network resilience. Reliab Eng Syst Saf 2020;199:106907.
[22] Guan P, Zhuang J. Modeling resources allocation in attacker-defender games with "Warm Up" CSF. Risk Anal 2016;36:776–91.
[23] Hausken K, He F. On the effectiveness of security countermeasures for critical infrastructures. Risk Anal 2016;36:711–26.
[24] Hausken K. Strategic defense and attack for reliability systems. Reliab Eng Syst Saf 2008;93:1740–50.
[25] Hausken K. Defense and attack for interdependent systems. Eur J Oper Res 2017;256:582–91.
[26] Mo H, Xiao X, Sansavini G, Dong D. Optimal defense resource allocation against cyber-attacks in distributed generation systems. Proc Inst Mech Eng O J Risk Reliab 2023.
[27] Zhang J, Wang Y, Zhuang J. Modeling multi-target defender-attacker games with quantal response attack strategies. Reliab Eng Syst Saf 2021;205:107165.
[28] Hou H, Wu W, Zhang Z, Wei R, Wang L, He H, et al. A tri-level typhoon-DAD robust optimization framework to enhance distribution network resilience. Reliab Eng Syst Saf 2024:110004.
[29] Li YP, Tan SY, Deng Y, Wu J. Attacker-defender game from a network science perspective. Chaos 2018;28:051102.
[30] Sun J, Wang S, Zhang J, Dong Q. Attack-defense game in interdependent networks: a functional perspective. J Infrastruct Syst 2023;29:04023020.
[31] Li Y, Qiao S, Deng Y, Wu J. Stackelberg game in critical infrastructures from a network science perspective. Phys A Stat Mech Appl 2019;521:705–14.
[32] Hunt K, Agarwal P, Zhuang J. On the adoption of new technology to enhance counterterrorism measures: an attacker–defender game with risk preferences. Reliab Eng Syst Saf 2022;218:108151.
[33] Zhang J, Zhuang J, Jose VRR. The role of risk preferences in a multi-target defender-attacker resource allocation game. Reliab Eng Syst Saf 2018;169:95–104.
[34] Peng R, Wu D, Sun M, Wu S. An attack-defense game on interdependent networks. J Oper Res Soc 2021;72:2331–41.
[35] Lin C, Xiao H, Peng R, Xiang Y. Optimal defense-attack strategies between M defenders and N attackers: a method based on cumulative prospect theory. Reliab Eng Syst Saf 2021;210:107510.
[36] Wu D, Yan X, Peng R, Wu S. Risk-attitude-based defense strategy considering proactive strike, preventive strike and imperfect false targets. Reliab Eng Syst Saf 2020;196:106778.
[37] Liu K, Wang M, Zhu W, Wu J, Yan X. Vulnerability analysis of an urban gas pipeline network considering pipeline-road dependency. Int J Crit Infrastruct Prot 2018;23:79–89.
[38] Shuvro RA, Das P, Jyoti JS, Abreu JM, Hayat MM. Data-integrity aware stochastic model for cascading failures in power grids. IEEE T Power Syst 2023;38:142–54.
[39] Han L, Zhao X, Chen Z, Gong H, Hou B. Assessing resilience of urban lifeline networks to intentional attacks. Reliab Eng Syst Saf 2021;207:107346.
[40] Caetano HO, LD N, Fogliatto MSS, Maciel CD. Resilience assessment of critical infrastructures using dynamic Bayesian networks and evidence propagation. Reliab Eng Syst Saf 2023:109691.
[41] Koc Y, Verma T, Araujo NAM, Warnier M. MATCASC: a tool to analyse cascading line outages in power grids. In: Proceedings of the 2013 IEEE international workshop on inteligent energy systems (IWIES). IEEE; 2013. p. 143–8.
[42] Wang X, Xue F, Lu S, Jiang L, Bompard E, Masera M, et al. Coordinated cyber-physical attack on power grids based on malicious power dispatch. Int J Electr Power 2024;155:109678.
[43] Zhang H, Ouyang M, Wu S, Hong L. Simplified operation models of integrated power and gas systems for vulnerability analysis. Phys A 2019;531:121428.
[44] Dobson I, Department E, Carreras BA, Lynch VE, Newman DE. An initial model for complex dynamics in electric power system blackouts. Hawaii Int Conf Syst Sci 2001:9.
[45] Wu Y, Guo P, Wang Y, Du M, Wang X, Zhang D. Vulnerability analysis of interdependent infrastructures considering the sensitivity of components to different risks. In: Proceedings of the 2023 IEEE international conference on systems, man, and cybernetics (SMC); 2023. p. 2403–8.
[46] Liu N, Liu S, Chai QW, Zheng WM. A method for analyzing Stackelberg attack-defense game model in 5G by tCPSO. Expert Syst Appl 2023;228:120386.
[47] Fang C, Dong P, Fang YP, Zio E. Vulnerability analysis of critical infrastructure under disruptions: an application to China Railway High-speed. Proc Inst Mech Eng O J Risk Reliab 2020;234:235–45.
[48] Chen Z, Zhang J, Du WB, Lordan O, Tang J. Optimal allocation of node capacity in cascade-robustness networks. PLoS One 2015;10:e0141360.
[49] Zhai C, Zhang H, Xiao G, Pan T. A model predictive approach to protect power systems against cascading blackouts. Int J Electr Power 2019.