Original Article

Institution of MECHANICAL ENGINEERS
Journal of RISK AND RELIABILITY

Proc IMechE Part O:
J Risk and Reliability
1–15

DOI: 10.1177/1748006X251336007
journals.sagepub.com/home/pio
Sage

Attack-defense game of interdependent infrastructure systems considering cascading failures

Yanfang Wu[1,2], Peng Guo[1], Ying Wang[1] and Enrico Zio[2,3]

Abstract

Interdependent infrastructure networks (IINs) are increasingly vulnerable to potential threats from terrorist activities, which can severely disrupt their performance. The dynamic interactions between intelligent attackers and defenders are crucial in determining the resilience of IINs. Based on game theory and complex network theory, this paper proposes a Stackelberg attack-defense game model considering cascading failures. The proposed two-player game model prioritizes the actions of the defender, with the attacker adopting the role of a follower who formulates a response to the defender's moves. The strategies and payoffs are defined based on the vulnerability of IINs under disruptions, accounting for cascading failures both within individual networks and between heterogeneous networks. An interdependent power and gas network is applied to explore equilibrium strategies and expected payoffs for both the attacker and defender. Simulation results reveal the importance of considering cascading effects from a network perspective when evaluating the performance of IINs. The findings demonstrate that narrowing the importance gap between nodes is an effective strategy for enhancing system resilience and mitigating the impact of attacks. The equilibrium strategies derived from this model offer valuable insights for improving the resilience of IINs against disruptive events.

Keywords

Interdependent infrastructure networks, attack-defense game, vulnerability distribution function, cascading failure, Strong Stackelberg Equilibrium

Date received: 16 April 2024; accepted: 27 March 2025

Introduction

Critical infrastructure systems are increasingly interconnected to form networks of interdependent infrastructure networks (IINs), such as Power-Transportation-Communication networks and Power-Water-Gas networks.[1] The complex interdependencies among infrastructures render IINs highly vulnerable to disruptions, as a cascading failure can propagate through interconnections once initiated by the failure of a component within the networks.[2] Recently, the growing significance and vulnerability of IINs have made them prime targets for malevolent attacks.[3] For instance, a terrorist attack involving the explosion of a natural gas pipeline in Syria in August 2020 resulted in a widespread power outage across the country. Similarly, a cyberattack on Colonial Pipeline Company in May 2021 led to the shutdown of its 5550-mile gasoline pipeline, causing an "energy shortage" in many eastern states in the U.S. Ensuring the security of IINs in the face of intelligent adversarial entities has become a pressing imperative.[4] However, the large scale of IINs, the complex interdependencies among components, the presence of intelligent adversaries, and limited protection resources pose a challenge to the security of IINs.

The attack-defense game model has been widely used to simulate confrontations between intelligent adversaries, serving as a suitable framework for analyzing strategic interactions in various domains.[5,6] Within this modeling paradigm, the attacker aims to maximize the damage of targeted systems, whereas the defender focuses on deploying optimal defense strategies to

[1]School of Management, Northwestern Polytechnical University, Xi'an, Shaanxi, China
[2]Energy Department, Politecnico di Milano, Milano, Italy
[3]MINES Paris-PSL University, CRC, Sophia Antipolis, France

Corresponding author:
Yanfang Wu, School of Management, Northwestern Polytechnical University, 1 Dongxiang Road, Chang'an District, Xi'an, Shaanxi 710129, China.
Email: wwuyanfang@mail.nwpu.edu.cn

counter attacks.[7] Recently, attack-defense game modeling has been extensively applied to the protection of critical infrastructures.[8,9] Zhang et al.[10] proposed a sequential defender-attacker game model that considers the risk preferences of players to address resource allocation problems across multiple targets. Hausken[11] developed a simultaneous game model for interdependent systems, accounting for the interdependence probabilities between two targets. Zhang et al.[12] improved the payoff function of the game model by incorporating the additional effects of the simultaneous destruction of related targets on systems. The attack-defense game model has been used to analyze the allocation of defense resources to mitigate attacks on infrastructure systems. However, most of these studies evaluate system performance by accumulating the values of individual targets, which overlooks the fact that interdependent infrastructure systems operate as networks of networks.[13] Therefore, it is essential to assess the performance of IINs from a network-wide perspective.

In recent research, complex network theory has been widely utilized to assess the performance of infrastructure systems characterized by complex interdependencies.[14,15] Ouyang[16] and Fang et al.[17] used a network flow-based approach to evaluate the resilience of infrastructure systems, which served as the foundation for the development of attacker-defender (AD) and defender-attacker-defender (DAD) models. Li et al.[18,19] developed attack-defense game models by defining strategies and payoffs based on the topological characteristics of the network. These studies focus on the selection of targeted components by the attacker and defender, aiming to maximize their expected payoff from a network perspective under resource constraints.[20,21] However, within these networks, the failure of one or a few nodes can trigger a cascade of node collapses due to the interdependencies among components in IINs, a phenomenon known as cascading failure.[22] To address this challenge, various methods have been proposed to simulate cascading failure mechanisms within complex networks, including flow-based models, load-capacity models, and percolation theory.[23–25]

There have been relatively few studies on attack-defense game models for infrastructure networks that consider potential cascading failure processes within IINs. Chaoqi et al.[26] considered the cascading failure process within an individual network and its impact on the payoffs of players in the attacker-defender game model. Huang et al.[27] analyzed how cascading failures impact the optimal strategies of both the defender and attacker. However, the focus of these studies has primarily been on individual networks. Understanding and modeling the propagation of failures across heterogeneous infrastructure systems is crucial for evaluating the performance of IINs under intelligent attack and defense actions.[28,29]

The identification of critical nodes is a fundamental consideration for both attackers and defenders when formulating their strategies model.[30] Recent research has applied component importance metrics to identify critical nodes that may be targeted for deliberate attacks and protection. Wu et al.[31] selected nodes with the highest degree and betweenness in a network to simulate terrorist attacks. Beyza et al.[32] proposed two different deliberate attack strategies, including the disruption of facilities with the most links and the most crucial facilities in terms of flow. Wang et al.[33] introduced a novel method, TOPSIS, to assess node importance by combining various metrics, including degree centrality (DC), closeness centrality (CC), betweenness centrality (BC), and eigenvector centrality (EC). However, the critical components identified in these studies are based solely on the topological characteristics of the system, overlooking the cascading failure effects triggered by the failure of critical components. Therefore, importance measures should consider the contribution of components to overall system performance.[34] In this context, the vulnerability distribution function is utilized in our research to analyze the importance of different components.[35]

To address the above challenges, this paper proposes a Stackelberg attack-defense game model that considers cascading failure effects both within and between networks, aiming to analyze the optimal strategy for the defender in response to intelligent attacks on IINs from a network perspective. The model allows for the assessment of the performance of IINs under the dynamic interactions between intelligent attackers and defenders. The main contributions of this work are summarized as follows:

(1) A Stackelberg attack-defense game model framework for interdependent infrastructure networks is proposed, with a focus on the selection of defensive strategies under resource constraints.
(2) The effects of cascading failures, both within individual infrastructure networks and across heterogeneous infrastructure networks, are incorporated into the payoff functions for both players, from a network perspective.
(3) The vulnerability distribution function according to N-1 scenarios is utilized to assess the importance of nodes within IINs. The resulting importance index serves as the referential property for cost evaluation in the attack-defense game model.
(4) A case study of an interdependent power-gas system is used to analyze the selection of equilibrium strategies for both the attacker and defender under varying resource constraints.

The rest of the paper is organized as follows. Section "Interdependent infrastructure networks" introduces the interdependent infrastructure networks model, including cascading failure mechanism design and critical node identification. Section "Attack-defense game model on IINs based on Stackelberg game" presents

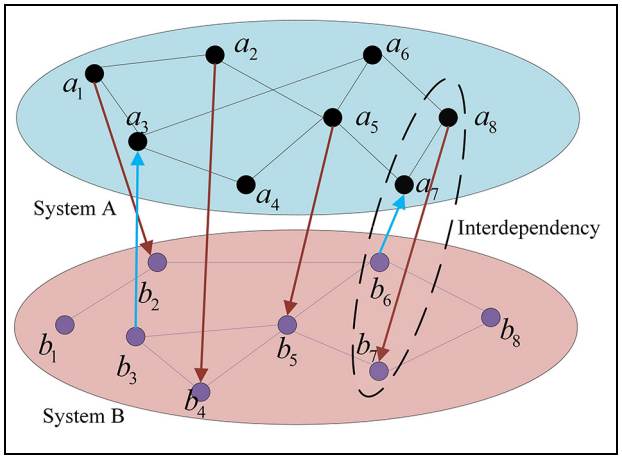

Figure 1. Interdependent infrastructure networks.

the Stackelberg attack-defense game model and equilibrium solution method. Section "Case study" demonstrates the application of the game on IINs through a case study of an interdependent power-gas network system. Finally, conclusions are shown in Section "Conclusions".

Interdependent infrastructure networks

In our research, complex network theory is used to model the topology of interdependent infrastructure systems. First, individual infrastructure systems are modeled as undirected networks, based on considerations of bidirectional flow transmission designs in infrastructures.[36,37] Second, interdependent infrastructure systems are modeled as directed multilayer networks, considering the directional physical interdependencies between infrastructures.[38,39] In the application of network theory to infrastructure systems, components are represented as nodes, and interdependencies between them are represented as links.[40] Drawing from the formal definitions of multilayer networks,[41] the set of IINs is denoted as $\mathcal{M}=(\mathcal{G},\ \mathcal{C})$ where $\mathcal{G}=\{G_k;k\in\{1,2,\ldots,M\}\}$ represents a collection of individual undirected infrastructure networks $G_k=(X_k,E_k)$, $\mathcal{C}=\{E_{kl}\subseteq X_k\times X_l;k,l\in\{1,2,\ldots,M\},k\neq l\}$ denotes the set of directed physical interdependencies among components across different networks G_k and G_l. Considering the backup of infrastructure and pre-protection, the dependency strength between components in G_l and G_k is represented as $P_{kl}=P(G_l|G_k)$ and $0\leqslant P_{kl}\leqslant 1$. $P_{kl}=0$ indicates that networks G_k and G_l are independent, otherwise, there is an interaction between them. The magnitude of P_{kl} indicates varying levels of dependency strength, which further influences the intensity of cascading failure propagation between networks. The set of nodes in network $G_k=(X_k,E_k)$ is denoted by $X_k=\{x_1^k,x_2^k,\ldots,x_N^k\}$ and the set of edges is defined by E_k. A representative layout of IINs is visually shown in Figure 1.

Network cascading failure model

In IINs, a cascading failure can be triggered both within and between networks. This paper adopts the load-capacity model to simulate the failure propagation process within independent networks.[42] The cascading failure across interconnected networks is analyzed using the Monte Carlo simulation method.[43,44]

(1) Cascading failures within independent networks

In the load-capacity model, the initial load $\widehat{l}_i^k$ of node x_i in network G_k refers to the workload undertaken by the node in the absence of disruptions. The load can be quantified either by the functional flow (e.g. power flow at a substation of a power network) or the structural load (e.g. degree of nodes in a network). The capacity of a node C_i^k represents the maximum load that the node can handle and is proportional to its initial load $\widehat{l}_i^k$, as follows:

$$C_i^k=(1+\delta_k)\widehat{l}_i^k \tag{1}$$

where tunable variable δ_k represents the tolerance parameter of a node, and nodes within the same network are assumed to have the same level of δ_k.

The attack on nodes is simulated as the removal of targeted nodes from networks. The load of the removed node will be redistributed to its neighboring nodes. The load of node x_i^k, redistributed to its neighboring node x_j^k at time t, can be calculated as:

$$\Delta l_{ij}^k(t)=l_i^k(t)\cdot\left(\frac{l_j^k(t)}{\sum_{m\in\Gamma_i(t)}l_m^k(t)}\right) \tag{2}$$

where $\Gamma_i(t)$ is the set of neighboring nodes of x_i^k at time t. The load of node x_j^k at time $t+1$ can be obtained by the following equation:

$$l_j^k(t+1)=l_j^k(t)+\sum\nolimits_{i\in\Psi_j(t)}\Delta l_{ij}^k(t) \tag{3}$$

where $\Psi_j(t)$ is the set of neighboring failed nodes of node x_j^k at the end of time t. Subsequently, a comparison is made between the load and capacity of node x_j^k: node x_j^k fails if $l_j^k(t+1)>C_j^k$, and its load is then redistributed according to equations (2) and (3). This process continues iteratively until all nodes no longer experience overload, that is, network stability is achieved.

(2) Cascading failures across interconnected networks

Initially, the dependency strength P_{kl} is defined. If $P_{kl}>0$, the failure of network G_k can propagate to its interdependent network G_l. Consequently, once network G_k stabilizes, the set of failed nodes $F_k=\{x_1^k,x_2^k,\ldots x_m^k\}$ in G_k can be obtained following (1). Subsequently, the dependent nodes set

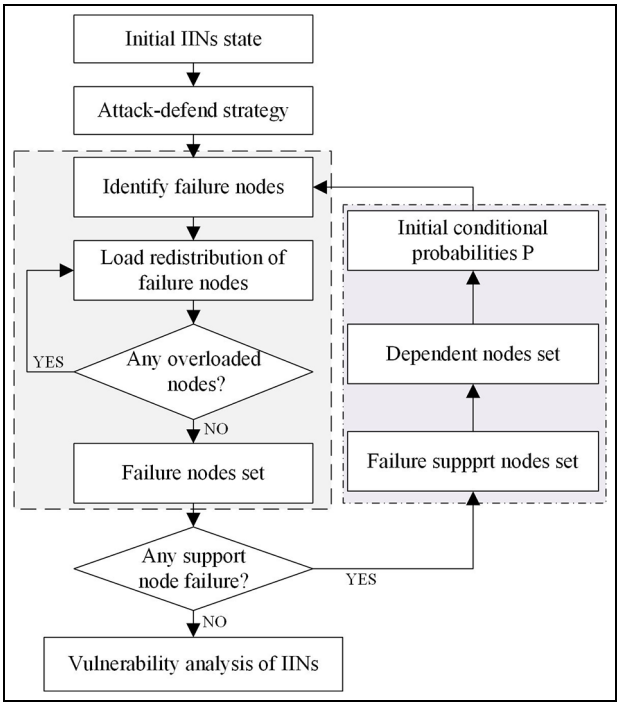

Figure 2. Propagation process of cascading failures in IINs.

$D_l = \{x_1^l, x_2^l, ...x_n^l\}$ in network G_l can be identified based on the dependency relationship E_{kl}. Considering the dependency strength, the Monte Carlo method is applied to simulate the probability of failure cascades for inter-network propagation. After obtaining $D_l = \{x_1^l, x_2^l, ...x_n^l\}$, a set of random numbers $R = \{r_1, r_2, ...r_n\}$ is generated using Monte Carlo simulation, where $r_i \in [0, 1]$. Node x_i^l fails if $r_i \leqslant P_{kl}$. The set of failure nodes in network G_l caused by cascading failures between IINs can be then identified as $F_l = \{x_1^l, x_2^l, ...x_u^l\}$. Subsequently, the cascading failure process within network G_l is triggered and can be simulated using the load-capacity model according to step (1). The cascading failure process stops when the IINs reach stability, with no overloaded nodes or the collapse of the entire system.

The propagation of cascading failures in IINs is illustrated in Figure 2. The gray dashed box shows the cascading failure process within a single network, modeled using the load-capacity model. The failure propagation between networks is shown in the purple dotted line box.

Critical node identification

For an intelligent attacker, the priority is to target critical nodes whose failure could lead to a massive collapse of IINs. Conversely, the defender aims to protect these critical nodes from attacks. Therefore, the identification of critical nodes is crucial for both attackers and defenders.

Recent research has considered vulnerability as a metric to represent the consequence of node failures on IINs. This metric quantifies the degree of functional decline when IINs encounter disruptions.[45] This paper applies the vulnerability metric to evaluate the performance of IINs under disruption. Connectivity loss is applied to assess the vulnerability of individual infrastructure networks. The vulnerability of IINs when node x_i fails can be expressed as the weighted sum of the vulnerability of each independent network:

$$V_i = \sum_{k \in K} w_k v_{k(i)} = \sum_{k \in K} w_k \left(\frac{\Gamma(G_k) - \Gamma(\hat{G}_{k(i)})}{\Gamma(G_k)} \right) = \sum_{k \in K} w_k \left(1 - \frac{n_{k(i)}}{N_k} \right) \tag{4}$$

where w_k is the vulnerability weighting factor for infrastructure network G_k and $\sum_{k \in K} w_k = 1$. $v_{k(i)}$ is the vulnerability of network G_k when node x_i fails. The function $\Gamma(G)$ represents the measure of network connectivity and is used to evaluate the vulnerability of networks. Specifically, $n_{k(i)}$ is the number of nodes in the giant component when node x_i fails, whereas N_k is the total number of nodes within network G_k.

Disruption scenarios are defined by the number of simultaneous failures of multiple components. Scenarios in which k out of N system components fail are referred to as N-k scenarios, where k represents the number of targeted components. To identify critical components, N-1 scenarios are considered in this paper. This approach effectively captures the entire spectrum of vulnerabilities of IINs by using the distributional metric. The disruption space $\Omega_{N-1} = (\omega_1, \omega_2, \ldots, , \omega_n)$ encompasses all distinct disruption scenarios, where ω_i represents the failure of node x_i. The cardinality of Ω_{N-1} is equal to N. The vulnerability in these disruption scenarios is treated as a discrete random variable, with each scenario ω_i corresponding to an event in the disruption space Ω_{N-1}. The vulnerability vector under distinct disruption scenarios is denoted as $V = (V_1, V_2, \ldots, , V_n)$.

The cumulative distribution function (CDF) $F(v)$ of vulnerability can be obtained based on the vulnerability vector as follows[35]:

$$F(v) = \frac{\sum_{i \in N} H(V_i \leqslant v)}{N} \tag{5}$$

where $H(V_i \leqslant v) = 1$ if $V_i \leqslant v$ and 0 otherwise. The CDF is referred to as the vulnerability distribution function. It quantifies the vulnerability of the system under N-1 scenarios, with its value ranges from 0 to 1 and exhibits a non-decreasing trend. The vulnerability mass function can be defined as:

$$f(v) = \frac{\sum_{i \in N} H(V_i = v)}{N} \tag{6}$$

where $H(V_i = v) = 1$ if $V_i = v$, and 0 otherwise. For N-1 scenarios, the vulnerability of IINs varies between a maximum value (representing the worst-case scenario) and a

minimum value, depending on which component within the network is disrupted. The distribution of vulnerability can represent the variability of V, and the heterogeneity of the consequences can be used to identify critical nodes.

The vulnerability vector serves as a basis for assessing the importance of nodes within IINs. A higher vulnerability value for a node indicates a greater potential impact on the function of the whole system. To ensure a fair and accurate evaluation of the significance of each node, a normalization process is applied to the vulnerability values. The importance of each node x_i can be obtained as follows:

$$I_i = \frac{V_i}{\sum_{i \in N} V_i} \tag{7}$$

Attack-defense game model on IINs based on Stackelberg game

This paper presents a sequential adversarial game model for IINs, involving the intelligent attacker and defender. Within the framework of the Stackelberg game,[46] the defender assumes the role of the leader, choosing strategies based on historical information, whereas the attacker acts as the follower, making decisions by observing the strategy adopted by the defender. In the event of an attack on a node, the node would become dysfunctional unless the defender intervenes to protect it. It is assumed that both players have complete information about IINs and are knowledgeable about all potential strategies that their opponents may employ.

Attack-defense strategy model

The attack-defense strategy model for IINs can be formulated as node selection and resource allocation problems. Both the attacker and defender aim to optimize their expected payoffs by selecting targeted nodes to attack and defend, respectively, within the constraints of available resources.

The offensive and defensive costs associated with components depend on their referential property.[19] Denote $c_i^A(c_i^D)$ as the cost of node x_i in attack (defense):

$$c_i^A = r_i^{q_A}, c_i^D = r_i^{q_D} \tag{8}$$

where $r_i \geqslant 0$ represents the referential property of node x_i. The importance index I_i is adopted to evaluate r_i. $q_A(q_D)$ is the attack (defense)-cost-sensitive parameter. When $q_A = q_D = 0$, the costs for each node are equal. For both the attacker and defender, the resources they can invest in IINs are:

$$C_A = \alpha \sum_{i \in N} c_i^A = \alpha \sum_{i \in N} I_i^{q_A} \tag{9}$$

$$C_D = \beta \sum_{i \in N} c_i^D = \beta \sum_{i \in N} I_i^{q_D} \tag{10}$$

where $\alpha \in [0, 1]$ and $\beta \in [0, 1]$ are the cost-constraint parameters for the attacker and the defender, respectively.

The strategies of the attacker and defender are represented by the attack vector $\vec{a} = (a_1, a_2, \ldots, a_n) \in S_A$ and defense vector $\vec{d} = (d_1, d_2, \ldots, d_n) \in S_D$, respectively. $S_A(S_D)$ denote the strategy set for the attacker (defender) and $a_i(d_i) \in \{0, 1\}$. When $a_i(d_i) = 1$ the node x_i is targeted by the respective player, otherwise $a_i(d_i) = 0$. The resource constraint is shown as follows:

$$\sum_{i \in N} d_i c_i^D \leqslant C_D, \sum_{i \in N} a_i c_i^A \leqslant C_A \tag{11}$$

Given the definition of the strategies of the attacker and defender, the strategy space for each player expands significantly with the network size. Considering that decision-makers are bounded rational,[47] this paper introduces two typical strategies for both players. The strategies include random node selection strategies and targeted node selection strategies based on node importance. Specifically, we consider the following: random node attack strategy (RAS) and targeted node attack strategy (TAS) are included in the attack strategy set S_A, whereas random node defense strategy (RDS) and targeted node defense strategy (TDS) are part of the defend strategy set S_D. To determine the targeted node selection, nodes are first arranged in descending order of node cost obtained from equation (8); then, nodes are progressively incorporated into the targeted component following this ordered sequence until the cumulative cost surpasses the resource constraint. If node x_i is targeted by players, $a_i(d_i) = 1$, otherwise $a_i(d_i) = 0$. The same procedure applies to random strategies, where involve a random sequence of node importance.

Pay-off model

Considering the strategic involvement of the players, the state of a targeted node is strictly dependent on the interplay between the attacker and defender:

$$s_i = a_i - d_i \tag{12}$$

where $s_i = 1$ indicates the failure of node x_i; otherwise, it denotes the operational state of the node. An attacked node remains functional if it is protected by the defender, which is represented as $a_i = 1$ and $d_i = 1$. In the absence of defense, the attacked node would be removed from IINs, when $a_i = 1$ and $d_i = 0$.

Suppose $U^A(U^D) : S_D \times S_A \to \mathbb{R}$ is the payoff function of the attacker (defender). $u^A_{\vec{d},\vec{a}}$ and $u^D_{\vec{d},\vec{a}}$ are the payoff of attacker and defender when the defender

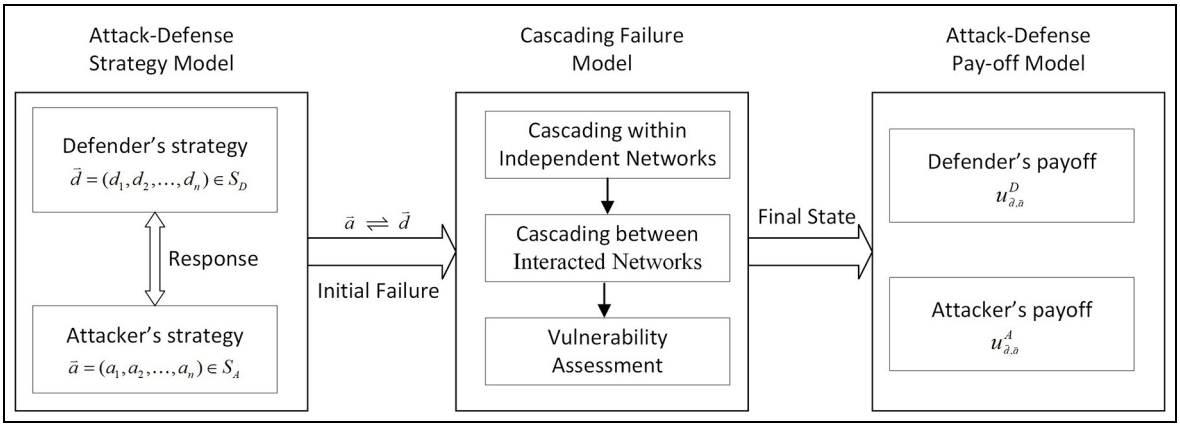

Figure 3. Flowchart of the attack–defense game on interdependent networks.

commits to strategy $\vec{d}$ and the attacker chooses strategy $\vec{a}$, respectively. The payoff function can be quantified using the vulnerability metric[19]:

$$u^A_{\vec{d},\vec{a}} = \sum_{k \in K} w^k \left(\frac{\Gamma(G^k) - \Gamma(\hat{G}^k)}{\Gamma(G^k)}\right) \tag{13}$$

$$u^D_{\vec{d},\vec{a}} = \sum_{k \in K} w^k \frac{\Gamma(\hat{G}^k)}{\Gamma(G^k)} \exp\left(\frac{\Gamma(\hat{G}^k) - \Gamma(\tilde{G}^k)}{2\Gamma(G^k)}\right) \tag{14}$$

where $\hat{G}^k(\hat{N}^k, \hat{E}^k)$ represents the remaining network resulting from the interaction between the attacker and defender. $\tilde{G}^k(\tilde{N}^k, \tilde{E}^k)$ represents the remaining network with no defense. Equation (13) represents the payoff function of the attacker, which is depicted by the vulnerability of IINs. Equation (14) represents the payoff function of the defender, which is determined by two factors. The first factor denotes the system performance under the interaction between the attacker and defender. The second factor shows the effectiveness of defense, which means that the defender will receive a higher payoff if more attacked nodes are defended.

Equilibrium solution

In the Stackelberg game framework, the concept of Strong Stackelberg Equilibrium (SSE) is applied to analyze the equilibrium strategies.[19] The defender takes on the role of the leader, choosing either a pure strategy or a mixed strategy (M-S) to optimize their expected payoff, where p_i represents the probability of the defender selecting strategy $i \in S_D$. Meanwhile, the attacker acts as the follower, formulating the optimal strategy by observing the actions taken by the defender. The flowchart of the sequence game is shown in Figure 3.

Once the payoff matrix is obtained, the SSE can be assessed using the Linear Programming method.[48] The attack-defense game model is formulated as follows:

$$max \sum_{i \in S_D} \sum_{j \in S_A} p_i y_j u^D_{ij} \tag{15}$$

Subject to

$$max \sum_{i \in S_D} \sum_{j \in S_A} p_i y_j u^A_{ij} \tag{16}$$

$$\sum_{i \in S_D} p_i = 1 \tag{17}$$

$$\sum_{j \in S_A} y_j = 1 \tag{18}$$

$$0 \leqslant p_i \leqslant 1 \forall i \in S_D \tag{19}$$

$$y_j \in \{0, 1\} \forall j \in S_A \tag{20}$$

Equations (15) and (16) represent the objective functions for the defender and the attacker, respectively, aimed at maximizing their expected utility. Constraints (17) and (18) ensure that each agent makes a decision based on the available strategy set. Constraints (19) and (20) define the variables: p_i represents the probability of the defender selecting strategy $i \in S_D$. y_j is a binary variable indicating whether the attacker selects strategy $j \in S_A$ in response to the strategy chosen by the defender.

Case study

Modeling interdependent infrastructure networks

This section presents a numerical experiment conducted on an interdependent power and gas system. The power system is based on the IEEE57 bus test system,[49] whereas the gas system is adopted from the GAS-48 system.[50] To model the power network (PN), physical components such as generation units, transmission substations, and distribution substations are represented as

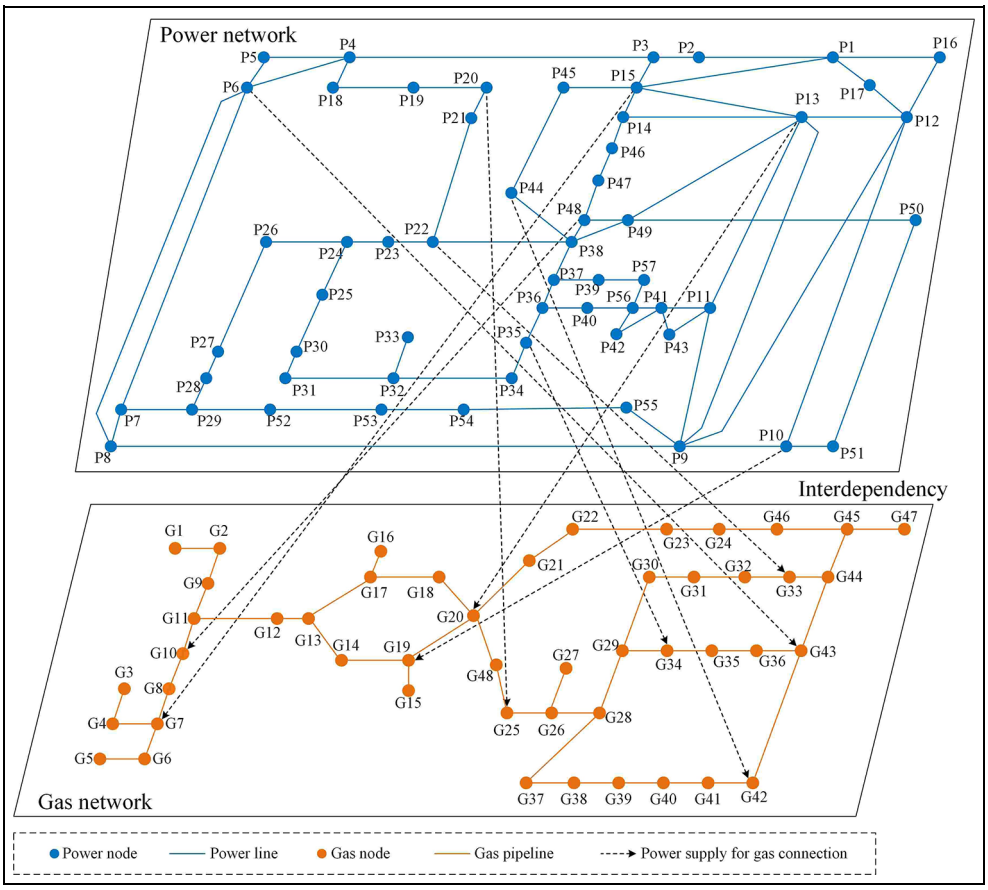

Figure 4. The layout of the interdependent power and gas systems.

power nodes, with power transmission lines modeled as power lines. For the gas network (GN), components like gas wells, transmission stations, and demand points are represented as gas nodes, with pipelines depicted as links. For the interdependency between PN and GN, the dependency of GN on PN is significant due to the essential role of electricity in the operation of critical components in GN, such as electric-driven gas sources and electricity-driven gas compressors. As for the reverse dependency, while PN does rely on GN for gas-fired generators that are driven by gas as fuel, the overall dependency of PN on GN is lower in regions where natural gas is not the primary source of power generation. Therefore, our research focuses on the dependency of GN on PN to analyze the impact of attacks on PN and the resulting failures in GN. Considering the dependency of GN on PN, a total of nine nodes are chosen from PN to provide support for nine nodes of GN. The layout of the IINs is illustrated in Figure 4. In this context, the set of attacked nodes is denoted as $N_A \subseteq N^{PN}$ and the set of defended nodes is represented by $N_D \subseteq N^{PN}$.

Critical node analysis

For ease of analysis, we assume that the capacity tolerance parameters are the same for both PN and GN, denoted as $\delta_{PN} = \delta_{GN}$, the vulnerability weighting factors are $w_{PN} = w_{GN} = 0.5$, and the degree of a node is adopted as the initial node load. The vulnerability distribution of the IINs subject to N-1 scenarios under different load tolerance values $\delta_{PN} = \delta_{GN}$, given $P_{kl} = 0.7$, is illustrated in Figure 5.

It can be observed that when $\delta_{PN} = \delta_{GN} = 0.3$, over 50% of nodes in PN are susceptible to causing a complete collapse of the IINs in the event of an attack. However, as the node capacity increases beyond $\delta_{PN} = \delta_{GN} = 0.5$, the failure of a single node tends to result in only a minor disruption in the IINs.

Figure 6 illustrates a consistent pattern in the vulnerability distribution of IINs across various values of P_{kl}. This pattern indicates that the failure of critical nodes can lead to a large vulnerability. Moreover, the vulnerability due to the failure of critical components increases with P_{kl}.

Based on the observations from Figures 5 and 6, the tolerance parameter plays a crucial role in the cascading failure process, significantly influencing the identification of critical nodes. In contrast, the influence of the dependency strength on node identification is relatively minor. In the following, we assume $\delta_{PN} = \delta_{GN} = 0.5$ and $P_{kl} = 0.7$ to assess the node importance values and the payoffs for both players. The importance values of the ten most critical nodes are shown in Table 1.

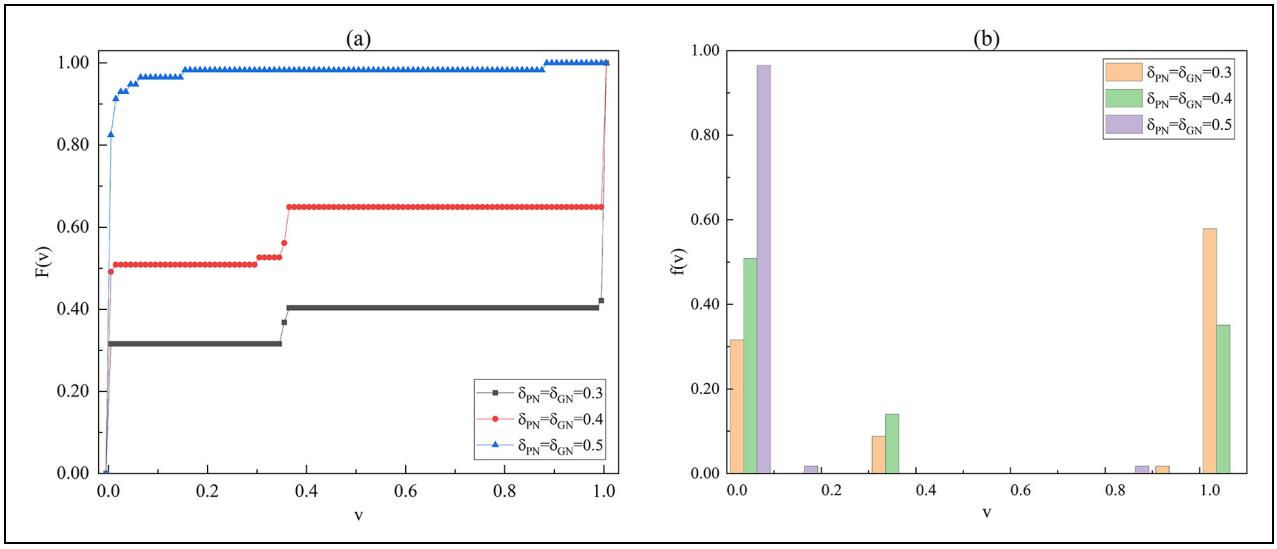

Figure 5. (a) The vulnerability distribution function $F(v)$ and (b) vulnerability mass function $f(v)$ of the IINs subject to $N-1$ scenarios under different load tolerance values $\delta_{PN} = \delta_{GN}$ given $P_{kl} = 0.7$.

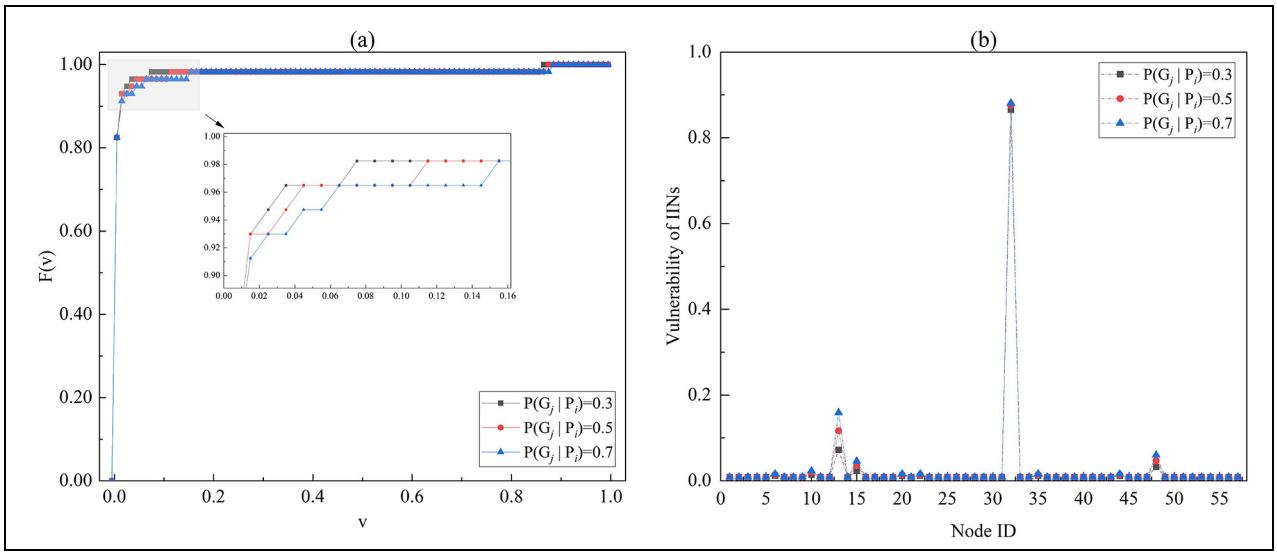

Figure 6. (a) Vulnerability distribution function $F(v)$ of the IINs and (b) Vulnerability of the IINs subject to N-1 scenarios under different P_{kl} values, given load tolerance $\delta_{PN} = \delta_{GN} = 0.5$.

Table 1. The characteristics of nodes

No.	Node ID	I_i	Degree	Betweenness
1	P32	0.5302	3	198.31
2	P13	0.0956	6	864.94
3	P48	0.0362	3	111.53
4	P15	0.0274	5	409.18
5	P10	0.0142	3	72.00
6	P44	0.0097	2	117.63
7	P22	0.0097	3	588.36
8	P20	0.0097	2	111.00
9	P6	0.0096	4	223.90
10	P35	0.0096	2	298.18

It is observed that the vulnerability resulting from a node failure is not always directly related to the nodes with high degree or betweenness centrality. For example, node P6, with a degree of 4 and betweenness of 223.90, but its importance value is only 0.0096. Conversely, node P32 with a degree of 3 and betweenness of 198.31 would trigger a substantial collapse in the IINs if it were to fail. This observation highlights the limitation of relying solely on network topology measures such as degree or betweenness to identify critical components. Evaluating the consequences of node failures is essential for both attackers and defenders.

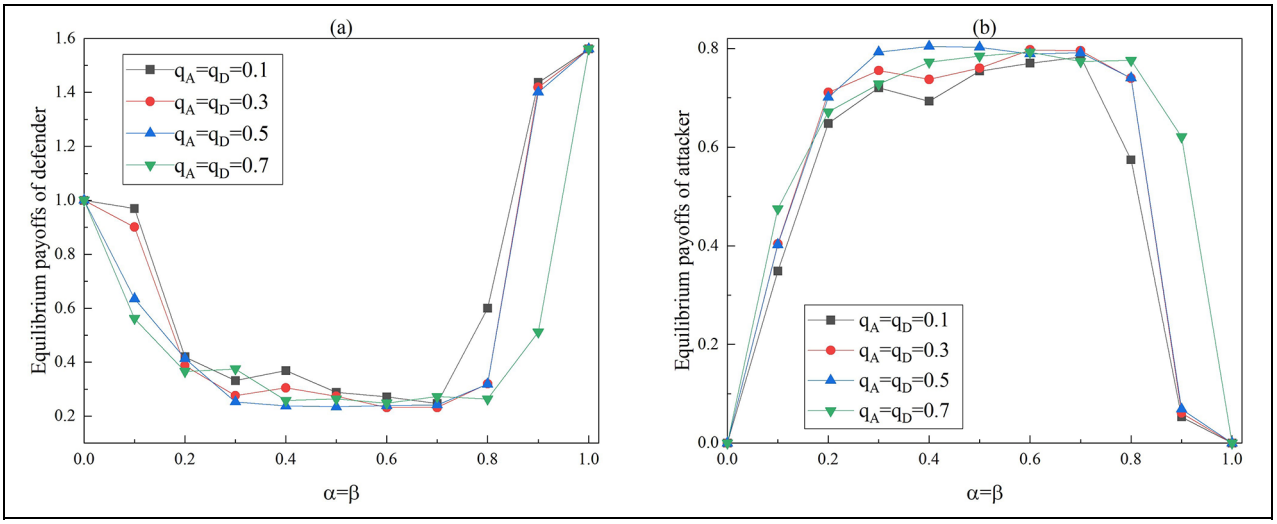

Figure 7. Equilibrium payoffs of the defender (a) and the attacker (b) versus cost-constraint parameters $\alpha(\alpha = \beta)$, with different values of $q_A = q_D$.

Equilibrium strategies based on Stackelberg game

Within the experimental setup described above, we first assume that $q_A = q_D$ and $\alpha = \beta$. A set of 1000 independent simulations is performed to assess the average payoff for both the attacker and defender. After obtaining the payoff matrix, equilibrium strategies are determined based on equations (15) to (20). The equilibrium payoffs for both players as a function of $\alpha(\alpha = \beta)$ and $q_A = q_D$ are shown in Figure 7.

Figure 7 shows that the trends in the expected equilibrium payoffs for the defender and attacker remain relatively consistent across different values of $q_A = q_D$. When $\alpha \leqslant 0.7$, the expected payoffs of the defender exhibit a declining fluctuation with the increase of $\alpha(\alpha = \beta)$, whereas the expected payoffs of the attacker show an increasing trend. This can be explained by the fact that as the attacker's resource available increases, he/she has the capacity to adapt the strategy in response to the defender's choices, thereby amplifying the effectiveness of the attacks. When $\alpha(\alpha = \beta) > 0.7$, the defender gains the ability to cover most of the attacked nodes regardless of the attacker's strategy. This leads to a rapid increase in the payoff of the defender, whereas the attacker's payoff decreases gradually and eventually reaches 0 when $\alpha(\alpha = \beta) = 1$.

Figure 7 illustrates that when $\alpha(\alpha = \beta) = [0.2{\sim}0.7]$, the expected equilibrium payoffs for both the defender and attacker remain relatively stable across varying values of $q_A = q_D$, implying that the impact of this parameter on payoffs is small in this scenario. However, when $\alpha(\alpha = \beta) = [0.1, 0.8, 0.9]$, with the increase of $q_A = q_D$, the expected payoff of the defender declines whereas the attacker's payoff increases. This is attributed to the fact that heightened cost sensitivity exacerbates cost differences between nodes: more crucial nodes demand a larger share of the defense investment. Consequently, when $q_A = q_D$ is large, the number of nodes that can be effectively defended decreases under the same resource constraint. It can be concluded that reducing the disparity in node importance levels can significantly enhance the defender's payoff. For this reason, we analyze the equilibrium strategy under varying resource constraints for the attacker and defender, given $q_A = q_D = 0.1$. The equilibrium strategies are shown in Figure 8.

Based on Figure 8, the findings can be summarized as follows.

(1) When the cost-constraint parameter $\alpha = 0, \beta = 0$, neither the defender nor the attacker has sufficient resources to cover any nodes for defense or attack. In this situation, both players are unable to implement any strategies. When $\alpha = 0, \beta > 0$, the defender is not required to make any protective decisions as the attacker is inactive. However, when $\alpha > 0, \beta = 0$, the attacker chooses TAS to optimize the payoffs, whereas the defender lacks the necessary resources to take any action.

(2) The diagonal line in the Figure illustrates a scenario where the amounts of offensive and defensive resources are equal. As resource increases, the defender moves from M-S to TDS to optimize the expected payoffs, leading the attacker to shift from TAS to RAS as a countermeasure. These results can be explained by Figures 8 and 9. When $\alpha \leqslant 0.15$, $u_{11}^D > u_{21}^D, u_{12}^D > u_{22}^D$, the choice of TDS by the defender prompts the attacker to choose RDS. In this context, u_{12}^D is not the optimal payoff for the defender, therefore, a mixed strategy is preferred by the defender to maximize the expected payoff. In response, the attacker adheres to TAS to counter the defender's decision. However, when $\alpha > 0.15$, the TDS becomes the dominant strategy for the defender no matter what strategy is chosen by the attacker. In response, the attacker would choose RAS because $u_{12}^A > u_{11}^A$.

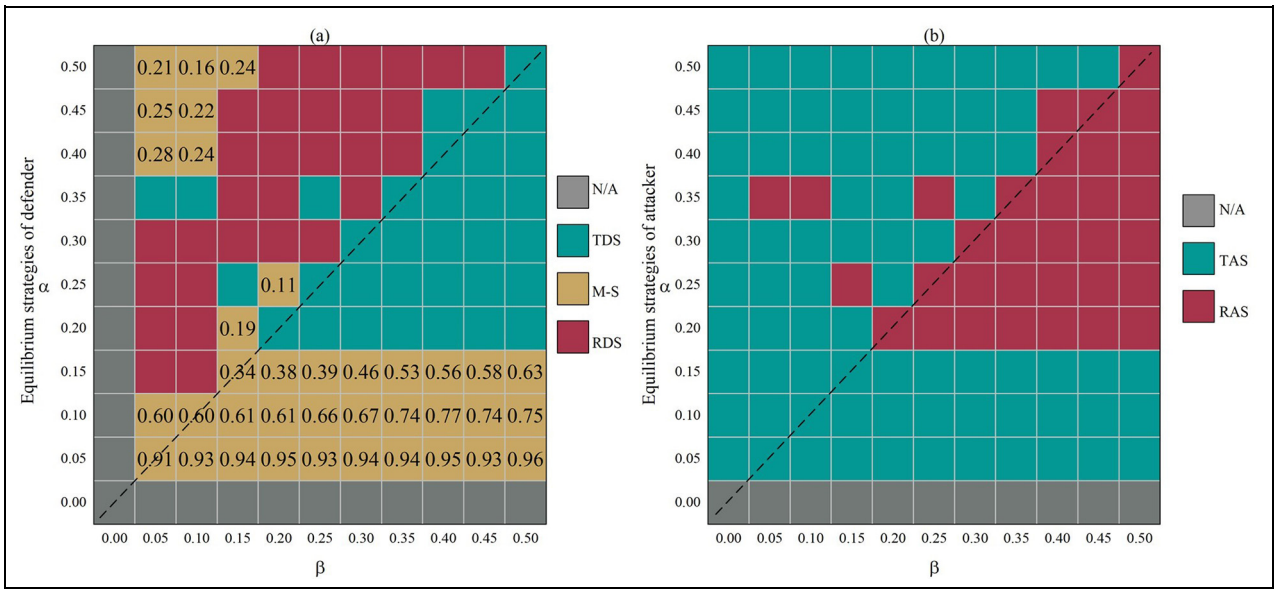

Figure 8. Equilibrium strategies of the defender (a) and the attacker (b) when α and β are different, with $q_A = q_D = 0.1$. The numbers in the blocks represent the probabilities of the TDS adopted by the defender in a mix-strategy (M-S).

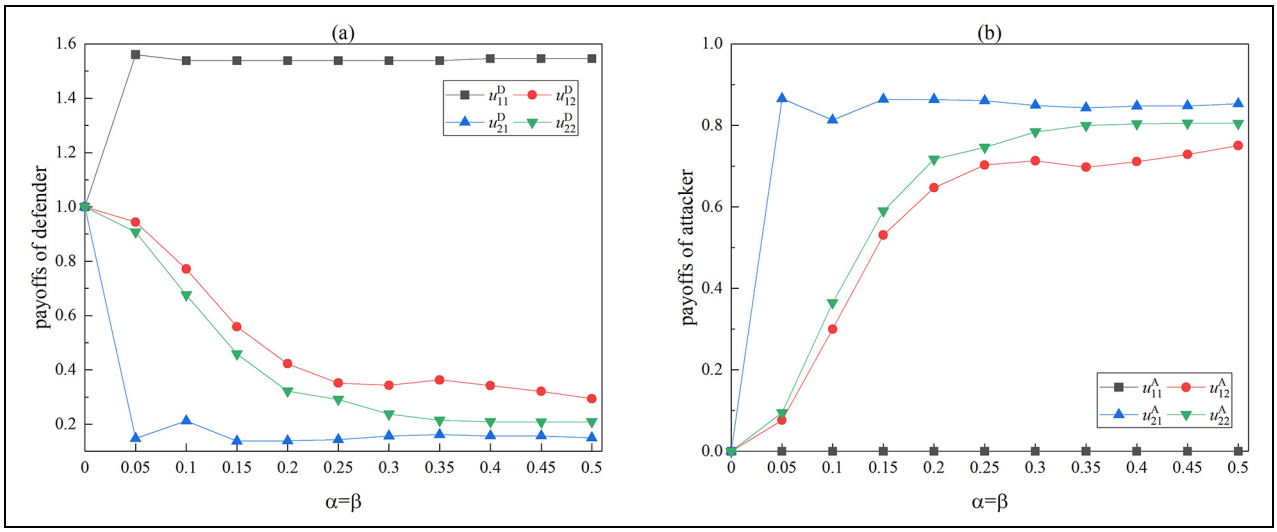

Figure 9. Payoffs of the defender (a) and the attacker (b) in each strategy profile with different cost-constraint parameters $\alpha = \beta$ of the defender (a) and the attacker (b).

(3) In the scenario when $\alpha < \beta$, the defender commits to mixed strategies in which the probabilities of the TDS decrease as the available attack resources increase when $\alpha \leqslant 0.15$. When $\alpha > 0.15$, the defender would choose TDS to protect critical nodes effectively. This can be explained by Figure 10.

(4) When $\alpha > \beta$, the defender's strategy choice becomes highly dependent on the available resource, as shown in Figure 11. Given the attack resource constraint $\alpha = 0.5$, when the defense resource is relatively low, $\beta \leqslant 0.15$, the optimal choice for the defender is M-S, whereas the attacker responds with TAS to optimize payoffs. As the defender's available resources increase within $0.15 < \beta \leqslant 0.5$, the attacker's best strategy remains TAS because $u_{11}^A > u_{12}^A, u_{21}^A > u_{22}^A$. In response to this strategic move by the attacker, the defender strategically adopts RDS.

(5) When the defender adopts the M-S, the attacker's expected payoffs remain the same, regardless of whether they choose RAS or TAS. However, by opting TAS, the defender can achieve the best-expected payoff, which coincides with the fundamental concept of the Strong Stackelberg Equilibrium (SSE). Consequently, when the defender commits to M-S, the attacker's strategic choice would lean toward TAS. Additionally, when the defender chooses TDS, the attacker would commit to RAS to optimize the expected payoffs, as the value of the payoff would be 0 when choosing TAS.

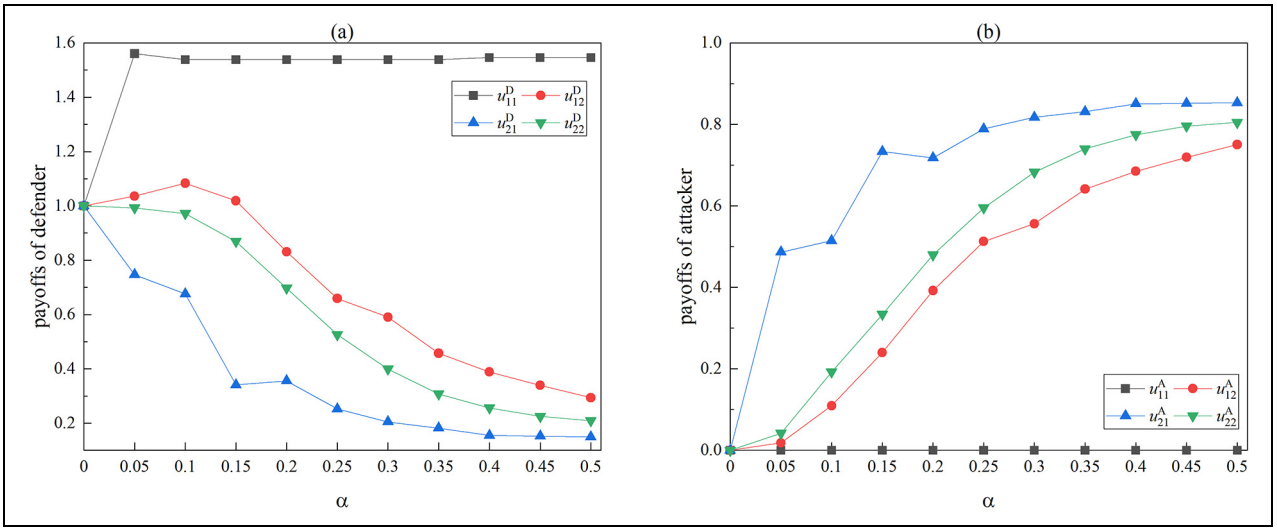

Figure 10. Payoffs of the defender (a) and the attacker (b) in each strategy profile with different cost-constraint parameters α when given $\beta = 0.5$.

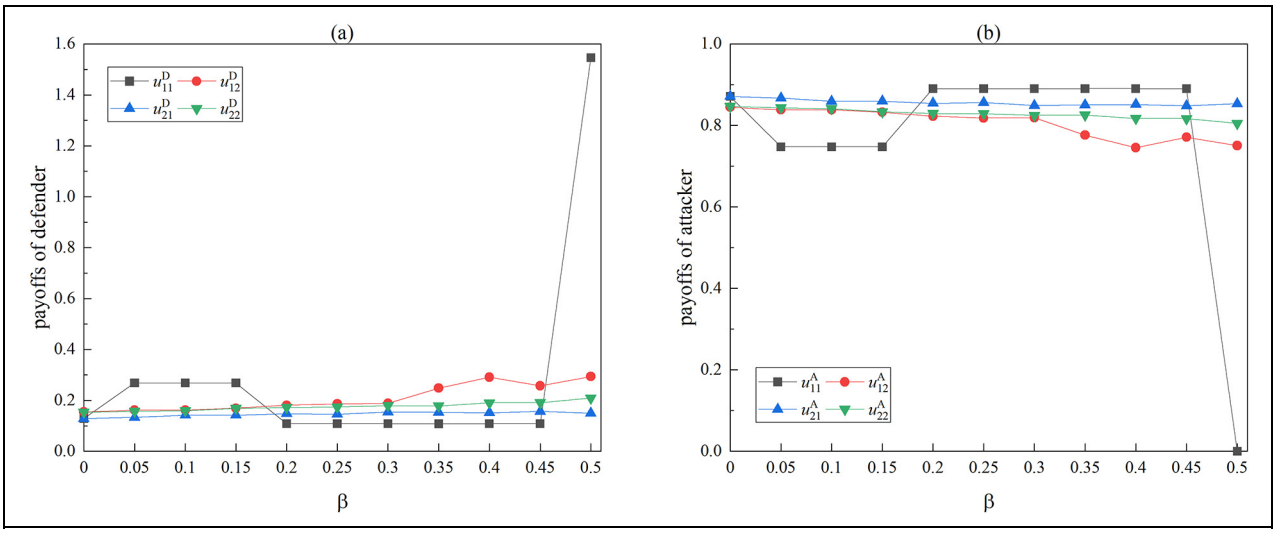

Figure 11. Payoffs of the defender (a) and the attacker (b) in each strategy profile with different cost-constraint parameter β when given $\alpha = 0.5$.

An interesting observation from Figure 11 is that given $\alpha = 0.5$, the defender's payoff may not exhibit an upward trend with the increase of defense resources. As shown in the figure, $u_{11}^{D}(0.2) < u_{11}^{D}(0.15)$, in this scenario, even though the defense resource increases, the attacker's payoff increases $u_{11}^{A}(0.2) > u_{11}^{A}(0.15)$. This implies that increasing defense resources does not necessarily lead to a proportional increase in payoffs for the defender, particularly within a specific range of attack resources. This can be attributed to the fact that cascading failures tend to amplify the impact of attacks when only a few nodes fail. The process of cascading failure can be mitigated when multiple nodes fail simultaneously. This propagation process of failure is shown in Figure 12.

Based on the analysis of the equilibrium strategies for both players, the equilibrium expected payoffs under different resource constraints for the attacker and defender are illustrated in Figure 13. When $\alpha > \beta$, the expected payoffs of the defender are relatively low, whereas the attacker's expected payoffs are high. In this scenario, the attacker does not need to invest substantial resources to achieve high payoffs. When $\alpha \leqslant \beta$, both the attacker and defender experience varying

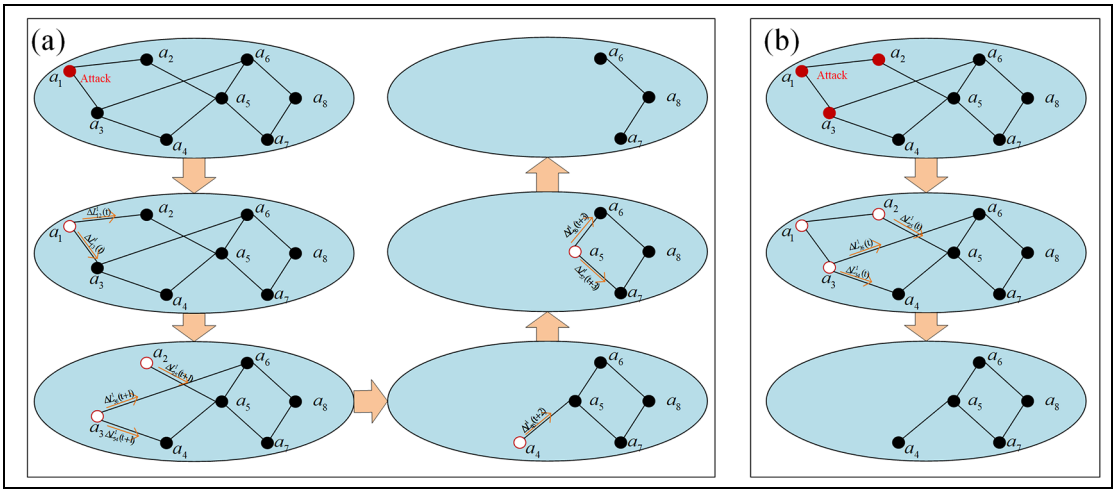

Figure 12. Cascading failure process of the network (with Load-capacity model): (a) there is one node attacked and (b) there are three nodes attacked simultaneously.

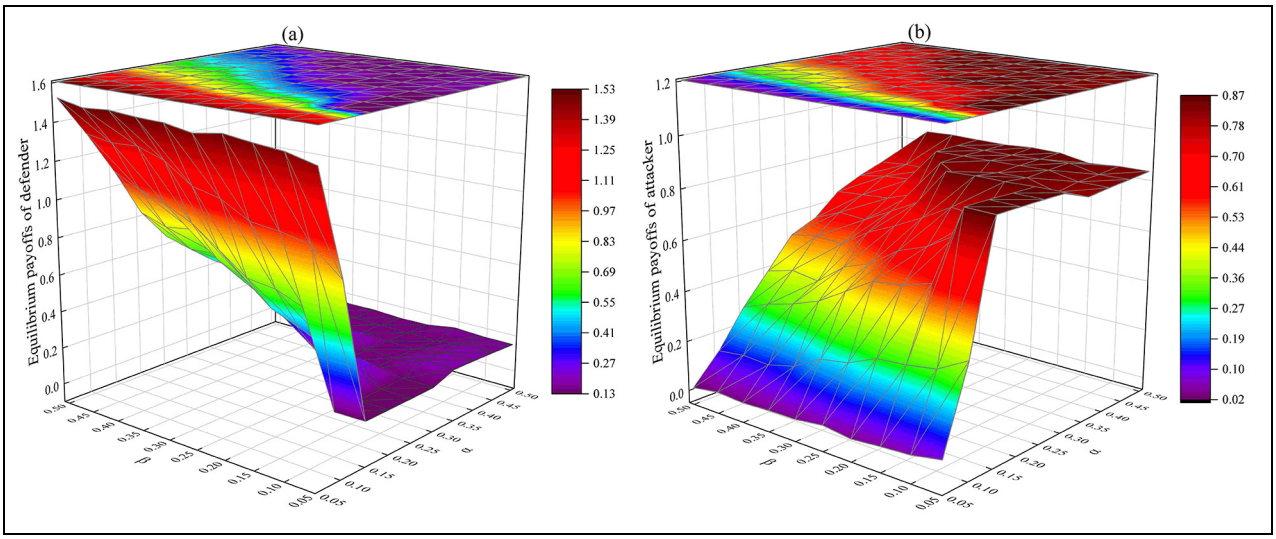

Figure 13. Equilibrium expected payoffs of the defender (a) and the attacker (b) when α and β are different with $q_A = q_D = 0.1$.

degrees of improvement in their expected payoffs as they invest more resources when given a certain opponent's resource levels.

Conclusions

The attack-defense game model provides an effective framework for analyzing the security of interdependent infrastructure networks (IINs) in the context of interactions between intelligent attackers and defenders. This paper proposes a Stackelberg attack-defense game model from a network science perspective, with a focus on considering cascading failures within IINs. In this study, we adopt the vulnerability distribution function to identify critical components, which serves as the foundation for the attack-defense game model. The load-capacity model and Monte Carlo simulation method are applied to simulate cascading failure processes both within individual networks and between heterogeneous networks. The equilibrium of the Stackelberg attack-defense model is analyzed by using linear programming methods, within the context of an interdependent power and gas network.

Meaningful conclusions confirm that: (1) Considering cascading failures is essential in evaluating the importance of components and the performance of complex networks. There is no positive correlation between group disruption or cooperative protection and the payoffs of both players when cascading failures are considered. Specifically, for the attacker, targeting critical nodes is more effective than simply investing additional resources in attacks throughout the entire network. For the defender, promptly isolating compromised nodes is crucial to prevent cascading failures. (2) In cases where the cost-sensitivity parameter is small, the defender's expected payoffs can be maximized while minimizing the

attacker's payoffs. This underscores the importance of reducing the importance gap between nodes for the defender when designing interdependent infrastructure systems. (3) The cost-constraint parameter plays a crucial role in determining the expected payoffs. While increasing resource allocation can enhance the resilience of IINs, it is crucial to recognize that the relationship between resource allocation and system resilience is not linear. Therefore, the defender must carefully balance resource investment costs with expected payoffs when making strategic decisions.

This study applies the attack-defense game model to unidirectional interdependent power and gas networks to validate its applicability and analyze the cascading impacts from PN to GN. Recently, to improve energy utilization, bidirectional interdependencies between infrastructures have been increasing significantly. Therefore, extending our game model to bidirectional IINs would provide a more comprehensive understanding of network dynamics. Additionally, within the attack-defense game model, future investigations should explore a broader range of strategy models based on the dynamic interactions between attackers and defenders, enabling a more effective exploration of potential strategic outcomes.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors would like to acknowledge the support of the National Natural Science Foundation of China (Grant No. 72171195).

ORCID iDs

Yanfang Wu https://orcid.org/0009-0009-6602-958X
Enrico Zio https://orcid.org/0000-0002-7108-637X

References

1. Karakoc DB, Barker K and Gonzalez AD. Analyzing the tradeoff between vulnerability and recoverability investments for interdependent infrastructure networks. *Socio Econ Plan Sci* 2023; 87: 101508.
2. Zio E, Golea LR and Sansavini G. Optimizing protections against cascades in network systems: a modified binary differential evolution algorithm. *Reliab Eng Syst Saf* 2012; 103: 72–83.
3. Fang Y and Sansavini G. Optimizing power system investments and resilience against attacks. *Reliab Eng Syst Saf* 2017; 159: 161–173.
4. Hausken K. Defense and attack of complex and dependent systems. *Reliab Eng Syst Saf* 2010;95:29–42.
5. Hausken K and Levitin G. Review of systems defense and attack models. *Int J Perform Eng* 2012; 8: 13.
6. Mo H, Xiao X, Sansavini G, et al. Optimal defense resource allocation against cyber-attacks in distributed generation systems. *Proc IMechE, Part O: J Risk and Reliability* 2024; 238(6): 1302–1329.
7. Nochenson A and Heimann CFL. Simulation and game-theoretic analysis of an attacker-defender game. In: J Grossklags and J Walrand (eds) *Decision and game theory for security*. Berlin, Heidelberg: Springer, 2012, pp.138–151.
8. Hausken K. Defence and attack of complex interdependent systems. *J Oper Res Soc* 2019; 70: 364–376.
9. Fang Y-P, Sansavini G and Zio E. An optimization-based framework for the identification of vulnerabilities in electric power grids exposed to natural hazards. *Risk Anal* 2019; 39: 1949–1969.
10. Zhang J, Zhuang J and Jose VRR. The role of risk preferences in a multi-target defender-attacker resource allocation game. *Reliab Eng Syst Saf* 2018; 169: 95–104.
11. Hausken K. Defense and attack for interdependent systems. *Eur J Oper Res* 2017; 256: 582–591.
12. Zhang X, Ding S, Ge B, et al. Resource allocation among multiple targets for a defender-attacker game with false targets consideration. *Reliab Eng Syst Saf* 2021; 211: 107617.
13. Havlin S, Kenett DY, Bashan A, et al. Vulnerability of network of networks. *Eur Phys J Spec Top* 2014; 223: 2087–106.
14. Han F and Zio E. A multi-perspective framework of analysis of critical infrastructures with respect to supply service, controllability and topology. *Int J Crit Infrastruct Prot* 2019; 24: 1–13.
15. Ouyang M. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliab Eng Syst Saf* 2014; 121: 43–60.
16. Ouyang M. A mathematical framework to optimize resilience of interdependent critical infrastructure systems under spatially localized attacks. *Eur J Oper Res* 2017; 262: 1072–1084.
17. Fang Y and Zio E. Game-theoretic decision making for the resilience of interdependent infrastructures exposed to disruptions. In: Gritzalis D, Theocharidou M and Stergiopoulos G (eds) *Critical infrastructure security and resilience*. Cham: Springer International Publishing, 2019, pp.97–114.
18. Li Y-P, Tan S-Y, Deng Y, et al. Attacker-defender game from a network science perspective. *Chaos* 2018; 28: 051102.
19. Li Y-P, Qiao S, Deng Y, et al. Stackelberg game in critical infrastructures from a network science perspective. *Phys A Stat Mech Appl* 2019; 521: 705–714.
20. Wu Y, Chen Z, Gong H, et al. Defender–attacker–operator: tri-level game-theoretic interdiction analysis of urban water distribution networks. *Reliab Eng Syst Saf* 2021; 214: 107703.
21. Ghorbani-Renani N, González AD, Barker K, et al. Protection-interdiction-restoration: tri-level optimization for enhancing interdependent network resilience. *Reliab Eng Syst Saf* 2020; 199: 106907.
22. Fang Y-P and Zio E. An adaptive robust framework for the optimization of the resilience of interdependent infrastructures under natural hazards. *Eur J Oper Res* 2019; 276: 1119–1136.

23. Motter AE and Lai Y-C. Cascade-based attacks on complex networks. *Phys Rev E* 2002; 66: 065102.
24. Bellè A, Zeng Z, Duval C, et al. Modeling and vulnerability analysis of interdependent railway and power networks: application to British test systems. *Reliab Eng Syst Saf* 2022; 217: 108091.
25. Turalska M and Swami A. Greedy control of cascading failures in interdependent networks. *Sci Rep* 2021; 11: 3276.
26. Chaoqi F, Yangjun G, Jilong Z, et al. Attack-defense game for critical infrastructure considering the cascade effect. *Reliab Eng Syst Saf* 2021; 216: 107958.
27. Huang Y, Wu J, Tse CK, et al. Sequential attacker-defender game on complex networks considering the cascading failure process. *IEEE Trans Comput Soc Syst* 2021; 9(2): 518–529.
28. Zio E and Sansavini G. Modeling interdependent network systems for identifying cascade-safe operating margins. *IEEE Trans Reliab* 2011; 60: 94–101.
29. Zio E. Challenges in the vulnerability and risk analysis of critical infrastructures. *Reliab Eng Syst Saf* 2016; 152: 137–150.
30. Zio E, Piccinelli R and Sansavini G. A framework for ranking the attack susceptibility of components of critical infrastructures. *Chem Eng Trans* 2012; 26: 309–314.
31. Wu B, Tang A and Wu J. Modeling cascading failures in interdependent infrastructures under terrorist attacks. *Reliab Eng Syst Saf* 2016; 147: 1–8.
32. Beyza J, Ruiz-Paredes HF, Garcia-Paricio E, et al. Assessing the criticality of interdependent power and gas systems using complex networks and load flow techniques. *Phys A Stat Mech Appl* 2020; 540: 123169.
33. Wang S, Sun J, Zhang J, et al. Attack-defense game analysis of critical infrastructure network based on Cournot model with fixed operating nodes. *Int J Crit Infrastruct Prot* 2023; 40: 100583.
34. Fang C, Dong P, Fang Y-P, et al. Vulnerability analysis of critical infrastructure under disruptions: an application to china railway high-speed. *Proc IMechE, Part O: J Risk and Reliability* 2020; 234: 235–245.
35. Galvan G and Agarwal J. Assessing the vulnerability of infrastructure networks based on distribution measures. *Reliab Eng Syst Saf* 2020; 196: 106743.
36. Dileep G. A survey on smart grid technologies and applications. *Renew Energy* 2020; 146: 2589–2625.
37. Shin J, Werner Y and Kazempour J. Modeling gas flow directions as state variables: does it provide more flexibility to power systems? *Electr Power Syst Res* 2022; 212: 108502.
38. Bao Z, Jiang Z and Wu L. Evaluation of bi-directional cascading failure propagation in integrated electricity-natural gas system. *Int J Electr Power Energy Syst* 2020; 121: 106045.
39. Zhang H, Ouyang M, Wu S, et al. Simplified operation models of integrated power and gas systems for vulnerability analysis. *Phys A Stat Mech Appl* 2019; 531: 121428.
40. Johansson J and Hassel H. An approach for modelling interdependent infrastructures in the context of vulnerability analysis. *Reliab Eng Syst Saf* 2010; 95: 1335–1344.
41. Boccaletti S, Bianconi G, Criado R, et al. The structure and dynamics of multilayer networks. *Phys Rep* 2014; 544: 1–122.
42. Zhang Y, Yang N and Lall U. Modeling and simulation of the vulnerability of interdependent power-water infrastructure networks to cascading failures. *J Syst Sci Syst Eng* 2016; 25: 102–118.
43. Di Maio F, Pettorossi C and Zio E. Entropy-driven Monte Carlo simulation method for approximating the survival signature of complex infrastructures. *Reliab Eng Syst Saf* 2023; 231: 108982.
44. Zio E. Stochastic simulation of critical infrastructures for electric power transmission. In: Gheorghe A and Muresan L (eds) *Energy security*. Dordrecht: Springer Netherlands, 2011, pp.109–124.
45. Johansson J, Hassel H and Cedergren A. Vulnerability analysis of interdependent critical infrastructures: case study of the Swedish railway system. *Int J Crit Infrastruct* 2011; 7: 289.
46. Hunt K and Zhuang J. A review of attacker-defender games: current state and paths forward. *Eur J Oper Res* 2024; 313: 401–417.
47. Sun J, Wang S, Zhang J, et al. Attack-defense game in interdependent networks: a functional perspective. *J Infrastruct Syst* 2023; 29: 04023020.
48. Conitzer V and Sandholm T. Computing the optimal strategy to commit to. In: *Proceedings of the 7th ACM conference on electronic commerce*, Ann Arbor, MI, 2006, pp.82–90. New York: ACM.
49. Zhai C, Zhang H, Xiao G, et al. A model predictive approach to protect power systems against cascading blackouts. *Int J Electr Power Energy Syst* 2019; 113: 310–321.
50. Chen S, Wei Z, Sun G, et al. Optimal power and gas flow with a limited number of control actions. *IEEE Trans Smart Grid* 2018; 9: 5371–5380.

Appendix

Acronyms

IINs	Interdependent infrastructure networks
AD	Attacker-Defender
DAD	Defender-Attacker-Defender
CDF	Cumulative distribution function
RAS	Random attack strategy
TAS	Targeted attack strategy
RDS	Random defense strategy
TDS	Targeted defense strategy
M-S	Mixed strategy
SSE	Strong Stackelberg Equilibrium
GN	Gas network
PN	Power network

Symbols

$\mathcal{M}$	Interdependent infrastructure networks
$\mathcal{G}$	Set of individual infrastructure networks
G	Infrastructure network of $\mathcal{G}$
$\mathcal{C}$	Set of directed physical interdependencies among components across different networks
P_{kl}	Dependency strength of network G_l on network G_k

X_k	Set of nodes in network G_k
E_k	Set of edges in network G_k
x_i	Node index
C_i^k	Capacity of node x_i in network G_k
$\widehat{l}_i^k$	Initial load of node x_i in network G_k
δ_k	Tolerance parameter of nodes in network G_k
$l_j^k(t)$	Load of node x_i in network G_k at time t
Γ_i	Set of neighboring nodes of x_i
Ψ_j	Set of neighboring failed nodes of node x_j
V	Vulnerability of IINs
w_k	Vulnerability weighting factor of network G_k
Ω_{N-1}	Set of disruption space
I_i	Importance index of node x_i
$c_i^A(c_i^D)$	Cost of node x_i for the attacker (defender)
r_i	Referential property of node x_i
$C_A(C_D)$	Investment resource of the attacker (defender)
$\alpha(\beta)$	Cost-constraint parameter of the attacker (defender)
$S_A(S_D)$	Strategy set of the attacker (defender)
$q_A(q_D)$	Cost-sensitive parameter of the attacker (defender)
s_i	State of node x_i
$U^A(U^D)$	Payoff of the attacker (defender)
p_i	Probability of the defender selecting strategy i