Exploring reliable strategies for defending power systems against targeted attacks

by Guo Chen, Zhao Yang Dong, David J. Hill, and Yu Sheng Xue

Copyright © 2011 IEEE.

This is an author-prepared version of the article, reprinted from IEEE Transactions on Power Systems, Vol. 26, Issue 3, p. 1000-1009.

http://dx.doi.org/10.1109/TPWRS.2010.2078524

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of University of Newcastle's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.
Exploring Reliable Strategies for Defending Power Systems under Terrorism Threat

Guo Chen, Student Member, IEEE, Zhao Yang Dong, Senior Member, IEEE
David J. Hill, Fellow, IEEE and Yu Sheng Xue, Member, IEEE

Abstract—Recently, game theory has been used to design optimized strategies for defending an electric power system against terrorist attacks. In this paper, we extend the current static model to a more generalized framework which includes several interaction models between defenders and attackers. A new criterion of reliable strategies for defending power systems has been derived. In addition, two effective allocation algorithms have been developed to seek reliable strategies for two types of defense tasks. The new criterion and algorithms are complementary to current security criteria and can provide useful information to assist decision-makers (governments), for protecting their power systems under possible terrorism threat. Numerical simulation examples using the proposed methods are given as well.

Index Terms—Game theory, Power system security, Risk management, Reliable strategies, Terrorism threat.

I. INTRODUCTION

Since September 11, 2001 and the frequent suicide bombing attacks in some countries, terrorism has become a major threat for national security. The US government [1] has spent over $150 billion on homeland security and appropriates about $15 billion on protecting the country’s critical infrastructure every year. Critical infrastructure is a term widely used by many governments to describe assets that are essential for the functioning of a society and its economy. So far, many countries have launched Critical Infrastructure Protection (CIP) plans. These include USA, Germany, United Kingdom, and Australia etc. Each plan aims at building a more secure and more resilient country with strengthened national preparedness with timely response and rapid recovery of critical infrastructures in the event of terrorist attacks, natural disasters, or other emergencies [2].

Power systems are always regarded as one of the most important infrastructures critical to the national security across the world. Consequently the vulnerability (security) analysis [3, 6] of power systems plays an essential role in the development of the electrical power industry. Traditionally, power system security is implemented via the well-established methodologies, which are mainly some criteria guiding the decision-makers on how to organize the prevention, response and recovery from a usual failure. For example, the widely used N-1 criterion [4] and the high risk N-k criterion [5] are among those criteria used for the power industry. Those criteria can handle disruptions resulted from accidents or random acts of the nature. Technically, they identify and deal with sets of events that are most likely to disrupt the systems and when those events happen so as to ensure the secure or normal operation of power systems. For economy, electrical companies usually ignore such events with sufficiently low probabilities of occurring.

However, with the boosting of a broad range of terrorism motivations, power systems, as one of the most important critical infrastructures, might became the target of terrorists [2, 7]. As a result, the traditional security framework of power systems is facing an immense challenge, because terrorists are often considered as fully intelligent and strategic actors. They can deliberately trigger those low probability events which are lack of protection but can cause serious damage to the power system. If such malicious attack happens, the impact will be significant. Some researchers have studied the power grid security problems under terrorist attacks. By studying how to attack power grids, they tried to explore new vulnerability measures of a power system. Salmeron et al. [8] firstly formulize the terrorism threat problem in power systems, in which terrorists try to maximize the load shed. Arroyo et al. [9] generalized Salmeron’ model to a more flexible bilevel programming problem. Moreover, Motto et al. [10] transformed the problems [8, 9] into a mixed integer bilevel programming model and presented a solution procedure. From [8-10], it can be observed that in the new context where terrorists come into play, traditionally robust power systems have become vulnerable. Therefore, seeking new methodologies and security criteria for defending power systems under potential terrorism threat is an urgent and important work.

Game theory [11] treats actors as fully strategic and has been successfully applied to many disciplines including economics [12-13], political science [14] and military [15-16], where multiple players with different objectives can compete and interact with each other. Recently, Holmegren et al. [17] proposed a static two-player zero-sum game model for studying the strategies of defending electric power systems under terrorist attacks. In the model, simultaneously, defenders deploy a strategy with limited budget for protecting each element of power systems and terrorists choose a target to attack. Furthermore, they studied a number of attack strategies and found that a dominant defense strategy1 did not exist. For every attack strategy, there exists

G. Chen and D. J. Hill are with the College of Engineering and Computer Science, The Australian National University, Canberra, Australia (e-mail: guo.chen@anu.edu.au and David.Hill@anu.edu.au).
Z. Y. Dong is with the Department of Electrical Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (e-mail: ezzydong@polyu.edu.hk)
Y. S. Xue is with the State Grid Electric Power Research Institute and NARI, Nanjing, China (e-mail: yxue@sgepri.com).

1 A dominant defense strategy means that for all attack scenarios, the strategy is always optimal.
an optimized defense strategy against it. This is an initial attempt for power system protection under terrorism threat and game theory inaugurates a new dimension for potential solutions. However, it is obvious that successful application of those optimized defense strategies requires priori defenses knowledge on the terrorists’ attack strategies. Otherwise, those optimized strategies might not be effective. For example, Table I which is abstracted from ref. [17], shows the loss of defenders in different combinations of attack strategies and defense strategies.

<table>
<thead>
<tr>
<th>Attack Strategy</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
<th>D6</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>2.0</td>
<td>2.5</td>
<td>5.0</td>
<td>2.5</td>
<td>3.0</td>
<td>3.2</td>
</tr>
<tr>
<td>D5</td>
<td>314</td>
<td>432</td>
<td>641</td>
<td>220</td>
<td>121</td>
<td>189</td>
</tr>
<tr>
<td>D6</td>
<td>703</td>
<td>966</td>
<td>1187</td>
<td>435</td>
<td>423</td>
<td>559</td>
</tr>
</tbody>
</table>

From Table I, we can see that defense strategy D1 for the attack strategy A1 is optimized, which can limit the loss of defenders to a minimum level compared with all other defense strategies. Nevertheless, if terrorists employ A5 against D1, the loss of defenders will increase significantly. If defenders employ D5 for defending A5, the loss can fall to a minimum level again, but will increase sharply again if terrorists adopt another strategy A11 against D5. In this kind of game, because both defenders and terrorists do not know the strategy of each other, employing optimized strategies D1 and D5 are not reliable and the loss of defenders is not predictable. Moreover, a risk of large loss always exists. To eliminate the risk, we need to explore a reliable strategy of defenders, under which large loss can be avoided and the final loss can be estimated no matter what attack scenario terrorists will play.

On the other hand, in real world applications, a more reliable situation is not like the static model but rather a dynamic (sequence) game model. Under a dynamic game model, before starting a strike, terrorists will make sufficient preparation and gather inside information about the object by all means, i.e. “how do defenders protect the object?”; “which parts are more vulnerable?”; “which parts can cause serious damage?” etc. That is to say, in the dynamic game, defenders should deploy a strategy first and then terrorists decide a target to attack after seeing the defender’s action. In this situation, those optimized strategies D1-D6 in Table I will fail to withstand. For example, if defenders deploy strategy D5, intelligent terrorists must choose attack strategy A11 in order to achieve the maximum loss of defenders. In this game, because terrorists have known the defenders’ strategy, they must seek a more effective attack scenario which can maximize the damage. This is an extreme situation between defenders and terrorists. Another possible situation is that terrorists can not obtain all information needed. That is, they only know partially the strategy of the defenders. Accordingly, the game between the defenders and the terrorists are manifold and to the best of our knowledge, few papers have discussed them in power system security analysis.

In this paper, we propose a comprehensive game framework which includes the static game model firstly proposed in [17] and several new dynamic versions extended by us. Furthermore, risk management theory will be introduced to analyze the framework and reliable strategies of defenders will be explored. Finally, two effective algorithms will be developed to achieve reliable strategies for the following two common defense problems.

1. When the defenders have a limited budget, how do they allocate the budget to deploy a reliable strategy?
2. When the defenders want to limit their loss to an expected value, how much budget do they need in order to deploy a reliable defending strategy?

The rest of this paper is organized as follows. Section II introduces a defender-attacker modeling approach. The game framework is described and analyzed in Section III. The two effective algorithms to obtain reliable strategies for defenders are formed in Section IV, while some illustrative examples are given in Section V. The conclusions are drawn in Section VI.

II. DEFENDER-ATTACKER MODELING

The defender-attacker model of electrical power systems, previously reported in [17], is provided below and we make some improvement for its wider application. The defenders are governments who have limited budget to protect power systems as much as possible. The attackers are terrorists who have a capability to launch a successful attack on a target with different sizes. For example, a single terrorist can break a transmission line or a transformer; a terrorism organization can disable several key elements of a power system. Thus, [17] assumes that a combination of elements of power systems can be considered as a target. A successful attack will lead to a failure of a target, which may cause a loss to utilities. Usually, many factors determine the loss, such as the price of failed equipments, the expense on maintenance, the reliability cost of interruption of energy supply to customers, as well as many other resultant direct and indirect costs. Moreover, it is obvious that losses are proportional to the time required to restore the services after such attacks. Thus, without loss of generality, before deploying a defense strategy, let \( x_j \geq 0 \) be the expected loss ratio of target \( j \). That is, if target \( j \) is attacked, the loss will be \( x_j \) per hour. Let \( T_j \) be the recovery time (h) that target \( j \) is totally restored. Accordingly, the total expected loss of target \( j \) without defending, defined as \( y_j \), can be calculated as \( y_j = x_j \cdot T_j \).

A. Defending Formulation

Assume that a power system is composed of \( N \) elements including generators, transmission lines, substation devices, transformers, etc.. Defenders need to decide how to allocate their limited budget \( R_0 \) for the protection of the \( N \) elements. Every element \( i \) has a protection function described by \( p_i(c_i) \) [17] as follows,
where \( c_i \) is the allocated budget for protecting element \( i \). Function \( p(f(c)) \) in (1) represents a decreasing function of a successful attack against element \( i \). For example, \( c_i = 0 \) means that there is no budget distributed on element \( i \), thus the probability of successful attack on \( i \) is 1. When \( c_i \) increases, the corresponding \( p(f(c)) \) should decrease. That is, it is harder to attack successfully. Correspondingly, \( p(f(c)) \) can be formulated as a continuous decreasing function. Defenders allocate budget among the \( N \) elements in a power system. Protection can be described by the following protection function vector \[ P = (p_1, p_2, \ldots, p_N) \]

In addition, usually electrical companies have a basic recovery capacity for maintaining and for repairing the failure of each element. \( t_{i \text{ave}} \) is the time it takes to repair element \( i \) when no budget is spent on the recovery [17]. If defenders allocate budget on recovery, the recovery time \( t_i \) will decrease, i.e.

\[ t_i = t_{i \text{ave}} \times f_i(c_{\text{recovery}}) \]

where \( f_i(c_{\text{recovery}}) \) is a continuous decreasing function and \( c_{\text{recovery}} \) is a variable as the allocation of budget for recovery. If a target is comprised of \( n \) elements, defenders can employ different repair schemes depending on the available budget [17]. Anyway, the recovery time can be roughly formulated as

\[ T_j = \sum_{i=1}^{n} t_{i \text{ave}} \times f_i(c_{\text{recovery}}) \]

Let the total budget be \( R_o \) and the strategy of defenders be \( c = (c_1, \ldots, c_N, c_{\text{recovery}}) \). It is easy to obtain the following:

\[ \sum_{i=1}^{N} c_i + c_{\text{recovery}} = R_o, i = 1, \ldots, N \]

where \( c_i \geq 0, c_{\text{recovery}} \geq 0 \) and \( N \) represents \( N \) elements.

### B. Attacking Formulation

As discussed previously, different terrorists (individuals and groups) can launch different size of attack. Suppose that terrorists can successfully attack a target consisting of \( n \) elements, where the \( n \) is limited by the capability of terrorists. For a target composed of \( n \) elements, the total number of targets \( M \) can be calculated as \[ M = \binom{N}{n} = \frac{N!}{n!(N-n)!} \].

Because defenders do not know the strategy that terrorists will choose, a reasonable assumption is that terrorists use a mixed strategy, i.e. randomize over those targets. Thus, terrorists’ strategy is a vector of probabilities \( q \) of dimension \( M \). \( q_j \) corresponds to the probability that target \( j \) is attacked, \[ q = (q_1, \ldots, q_M) \]

To make an attack on a target, the terrorists would choose, a reasonable assumption is that terrorists use a mixed strategy, i.e. randomize over those targets. Thus, terrorists’ strategy is a vector of probabilities \( q \) of dimension \( M \). \( q_j \) corresponds to the probability that target \( j \) is attacked, \[ q = (q_1, \ldots, q_M) \]

\[ 0 \leq q_j \leq 1, \quad j = 1, \ldots, M \]

\[ \sum_{j=1}^{M} q_j = 1 \]

### III. The Comprehensive Game Framework

#### A. The payoff between defenders and terrorists

Defenders and terrorists are strictly adversaries of each other and there is no cooperation between them. The aim of defenders is to decrease their loss as much as possible while terrorists intend to increase the loss to the greatest possible extent. Thus, the interaction between them can be seen as a two-person zero-sum game. That is, terrorists’ gain equals to defenders’ loss which is called the payoff between them.

Accordingly, considering the descriptions in in Section II, the payoff (the loss of defenders or the gain of terrorists) can be defined as

\[ L = \sum_{j=1}^{M} L_j(c, q) = \sum_{j=1}^{M} q_j U_j(c) \]

where \( q_j \) is the probability that target \( j \) is attacked and \( U_j(c) \) is the expected loss under defending strategy \( c \). For an attack on a target \( j \) which is comprised of a single element \( j \),

\[ U_j(c) = p_1(c_1)p_2(c_2)y_{1j} + p_1(c_1)(1-p_2(c_2))y_{1j} + p_2(c_2)(1-p_2(c_2))y_{1j} \]

Generally speaking, if a target \( j \) consisting of \( n \) elements, i.e. \( i_1, \ldots, i_n \), is attacked, the loss \( U_j(c) \) of defenders can be defined as

\[ U_j(c) = p_1(c_1)p_2(c_2)y_{1j} + \cdots + p_1(c_1)\cdots p_n(c_n)y_{1j} \]

\[ (1/p_1(c_1))y_{b_1} + \cdots + p_1(c_1)\cdots p_n(c_n)(1-p_n(c_n))y_{b_1} \]

\[ \cdots (1/p_n(c_n))y_{b_1} \]

\[ U_j(c) \text{ is a sum of } c_n^n + c_n^{n-1} + \cdots + c_n^1 = 2^n - 1 \text{ items.} \]

On the other hand, it should be noted that the expected loss of each target \( j \) without defending, i.e. \( y_{[S]} \) where \( S \) is the subset of target \( j \), will change with time. For example, the depreciation of equipments; the maintenance cost should increase if the payment of employees or other fees rise; the load is also increasing every year with increasing demand for electricity in most nations. However, usually the fluctuation of \( y_{[S]} \) in a period is not distinct. Therefore, we set a period as the validity of the strategy of defenders. Depending on the requirement of defenders for different precision, the period can be a quarter, half a year or a year. In a period, \( y_{[S]} \) is...
considered as a constant. After the period, defenders should change the strategy according to the new $y_{t(5)}$. In real world, it is also common that governments would adjust the appropriation of budget periodically through various approaches/projects such as annual planning, operations planning, system planning and revenue resets.

Admittedly, the computational time will increase if defenders recalculate $y_{t(5)}$ regularly. However, governments have the capability to employ some high performance computing equipments to accurately calculate $y_{t(5)}$ or approximately estimate $y_{t(5)}$ by empirical data. Given the focus of this paper, we do not discuss how to obtain $y_{t(5)}$. The objective of the paper is to develop methodologies such that when defenders have known $y_{t(5)}$ beforehand, the defenders know how to deploy a reliable strategy against all possible attacks by intelligent and strategic terrorists. In fact, estimating $y_{t(5)}$ should be a prerequisite for defenders against potential terrorist attacks. That is, if defenders do not know $y_{t(5)}$, it is impossible to protect power systems effectively. Specifically, we attempt to acquire a universal strategy $q$ and no matter what strategy $q$ terrorists plays, the playoff $L$, namely the loss of defenders, can be estimated and limited to a minimum level. In order to do so, firstly we need polish the relationship between defenders and terrorists. According to game theory [18], the interaction between defenders and terrorists can be divided into several situations. In the following section, we will describe those situations and then combine them into a new comprehensive framework.

**B. Types of game between defenders and terrorists**

1) Static game [17]

Simultaneously, defenders deploy a strategy $c$ (allocation plan of budget for $N$ elements of a power system) to defend and terrorists choose a strategy $q$ to launch an attack. Simultaneity [17, 18] includes another equivalent case: they do not move at the same time, but the later player is unaware of earlier player’s action. In the game of defenders and terrorists, the later player must be terrorists. Otherwise if terrorists attack first, the later defense is useless$^2$.

2) Dynamic game

The static game assumes that terrorists know nothing about defenders’ strategy. However, in real situation, terrorists can try their best to seek the information they need. For example, they can use threat, blackmail, torture and bribery to acquire the protection information of power systems, namely the strategy of defenders and the accurate time when the strategy will be altered. Therefore in order to better include those factors, we extend the static model to a dynamic version:

\[ \max_{q \in Q} \sum_{j=1}^{M} q_j U_j (c'). \]  

(14)

where $Q$ is the strategy space of terrorists. This programming is implemented with constraints (1)–(9). Let $q^*$ be the solution and $L'$ be the objective function value, namely the maximum. If terrorists do not know $c'$, there is a possibility (risk) that defenders will lose the maximum $L'$ when terrorists play strategy $q'$ by chance, which is the pessimistic situation under strategy $c'$ and provide an appropriate benchmark indicating the risk. Moreover, if

In the beginning of every period, defenders deploy a strategy $c$ first. Terrorists can see the action $c$ and then choose a strategy $q$ to launch an attack in the period.

(3) Manifoldness of games

The static and dynamic models are two extreme cases which assume that terrorists know nothing or fully know the strategy of defenders. Sometimes terrorists can only partly know the strategy and we can form many cases based on how much terrorists have known the strategy. Consequently, the game models between defenders and terrorists are manifold, which make the problem quite complicated if we discuss them one by one. For facilitating the analysis, we generalize the diversity into the following comprehensive framework.

**C. The comprehensive framework**

1) Framework description

From the previous subsection, we can see that there is no difference for defenders to play those different kinds of games. They must deploy a strategy first and they do not know the strategy of terrorists. The manifoldness is only resulted from the diversity of terrorists. Accordingly the framework can be formed by the following two phases.

Phase 1: Periodically, defenders deploy a strategy first and they know nothing about the strategy of terrorists.

Phase 2: Terrorists choose a strategy to attack in the period. There are three kinds of terrorists. They know nothing about the strategy of defenders which is equivalent to the static model; they partly know the strategy, or they fully know the strategy, which is equivalent to the dynamic model.

2) Framework Analysis

In Introduction section, we have seen that while deploying those optimized strategy $D1-D6$ [17] in the static model, there is a risk or possibility with a large loss against the strategy. With the increase of information collecting activities about the strategy of defenders, which corresponds to terrorists partly know the strategy, the risk will rise and finally reach 100% of the dynamic model. According to risk management theory [19], for each strategy played by defenders, a maximum loss of defenders under the strategy always exists. The maximum loss is a pessimistic situation. For example, assume that defenders deploy a strategy $c' \in C$, where $C$ is the strategy space of defenders, the maximum loss or pessimistic loss can be achieved by the following mathematical programming

\[ \max_{q \in Q} \sum_{j=1}^{M} q_j U_j (c'). \]  

where $Q$ is the strategy space of terrorists. This programming is implemented with constraints (1)–(9). Let $q^*$ be the solution and $L'$ be the objective function value, namely the maximum. If terrorists do not know $c'$, there is a possibility (risk) that defenders will lose the maximum $L'$ when terrorists play strategy $q'$ by chance, which is the pessimistic situation under strategy $c'$ and provide an appropriate benchmark indicating the risk. Moreover, if

\[ \text{If terrorists attack first, they can obtain a large playoff and the later defense can not change the result which has been gained by terrorists.} \]
terrorists know $c^t$, rational and intelligent terrorists must play strategy $q^t$ for pursuing the maximum gain $L^t$.

From the example, we can see that although there are several kinds of terrorists considered in the proposed framework, they can only increase the risk or possibility of pessimistic situation and they can not increase the pessimistic loss, which implies that we can seek a universal reliable strategy against all kinds of terrorists by decreasing the pessimistic loss as much as possible.

For each strategy of defenders $c \in C$, we can have a pessimistic loss by maximum programming (14). It is obvious that there exists a minimum pessimistic loss on the strategy space of defenders, which can be formed by the following minmax programming.

$$
\min_{c \in C} \max_{q \in Q} \sum_{j=1}^{M} q_j U_j(c).
$$

Eq (15) is also implemented with constraints (1)–(9). Let pair $(q^0,c^0)$ be the solution and $L^0$ be the objective function value. That is, when defenders deploy strategy $c^0$ and terrorists choose strategy $q^0$, the loss of defenders is the minimum pessimistic loss $L^0$. Because $L^0$ is a pessimistic loss of defenders under strategy $c^0$, as long as they deploy strategy $c^0$, no matter what strategies terrorists play, the loss of defenders can not exceed $L^0$. Moreover, strategy $c^0$ can decrease the pessimistic loss to a minimum. Obviously, $c^0$ is a reliable strategy of defenders. Thus, we can derive the following Criterion for reliable strategy design.

**Reliable Strategy Criterion:** In the proposed framework, the reliable strategy of defenders is $c^0$ and the minimum pessimistic loss of defenders is $L^0$, which can be obtained respectively by

$$
c^0 = \arg \min_{c \in C} \max_{q \in Q} \sum_{j=1}^{M} q_j U_j(c). \tag{16}
$$

and

$$
L^0 = \min_{c \in C} \max_{q \in Q} \sum_{j=1}^{M} q_j U_j(c). \tag{17}
$$

This criterion reveals that as long as defenders deploy strategy $c^0$, they can guarantee that their loss will be not more than $L^0$. It should be noted that the final payoff can not be exactly achieved before terrorists launch an attack because of the diversity of terrorists. In fact, $L^0$ is the least upper bound of defenders’ loss, namely the minimum loss in the worst case.

**IV. THE PROPOSED ALLOCATION ALGORITHMS**

**A. Budget allocation analysis**

From the criterion, we know that the reliable strategy of defenders is $c^0$ and the minimum pessimistic loss that defenders can guarantee to themselves is $L^0$. They can be achieved by solving (15). Previous attempts [20-23] to solve the minmax programming have mostly focused on the saddle point, which needs equation (18) to be hold.

$$
\max_{q \in Q} \min_{c \in C} \sum_{j=1}^{M} q_j U_j(c) = \min_{c \in C} \max_{q \in Q} \sum_{j=1}^{M} q_j U_j(c) \tag{18}
$$

However, Eq. (18) requires the playoff function satisfying some strict conditions [20-23]. In the proposed framework, we just point out how to design the protection functions $p_j(c)$ and the recovery functions $f_j(c_q,c_p,e)$, but we do not exactly specify the details of them. This approach provides sufficient flexibility for defenders to design their protection and recovery functions depending on their specific situations. Thus, the condition of (18) can not always hold. To avoid such difficulties, in this section, we will reveal that the playoff function has a special structure which implies that we can develop a general effective algorithm to obtain the reliable strategy $c^0$ and the minimum pessimistic loss $L^0$, regardless of whether equation (18) holds or not.

The playoff function is

$$
L(c) = \sum_{j=1}^{M} L_j(c) = \sum_{j=1}^{M} q_j U_j(c)
$$

where $\sum_{j=1}^{M} q_j = 1$. Obviously, $L(c)$ can be seen as a normalized weighted mean of $M$ items and each item $U_j(c)$ represents the expected loss of defenders if target $j$ is attacked with defending under strategy $c$. For a fixed $c$, the maximum of the normalized weighted mean, namely

$$
\max_{q \in Q} \sum_{j=1}^{M} q_j U_j(c), \text{ must be } \max \{U_j(c) \mid c \in C\}.
$$

That is to say, if only one target has the maximum expected loss among all targets, the maximum playoff can be obtained by terrorists attacking the target. If there are $k$ targets with the same maximum expected loss, the maximum playoff is achieved by terrorists attacking one of them. Anyway, for any strategy $c$ deployed by defenders, the maximum playoff $M(c)$ always exists, which is defined as

$$
M(c) = \max \{U_j(c) \mid c \in C\} \tag{19}
$$

Obviously, Eq. (19) is equivalent to

$$
\max_{q \in Q} \sum_{j=1}^{M} q_j U_j(c).
$$

Accordingly, minmax programming (15), namely

$$
\min_{c \in C} \max_{q \in Q} \sum_{j=1}^{M} q_j U_j(c),
$$

can be transformed into

$$
\min_{c \in C} M(c) \tag{20}
$$

By the equivalent transformation, we can eliminate the vector $q$ in the initial mathematical programming (15). This largely reduces the complexity of the programming. In fact, the new programming (20) can be considered as a pure allocating problem: i.e. to allocate budget to $N+1$ variables
(c₁,⋯,cₙ,c_{recovery}) such as M(c) becomes minimum. The final allocation plan, i.e. (c₁,⋯,cₙ,c_{recovery}) is the reliable strategy of defenders and the final M(c) is the minimum pessimistic loss of defenders.

To make the allocation more clear, firstly, we consider the simplest case. It is known that every currency has a minimum unit. For example, the minimum of USD is 1 cent and the minimum of AUD is 5 cents. If the budget is only the minimum, it must be totally allocated to a single variable in (c₁,⋯,cₙ,c_{recovery}). The expected loss of a target which is comprised of n elements, i.e. i₁⋯iₙ, can be expressed as

\[ U_j(c) = p_j(c) - \sum_{i} p_i(c) y_{i,j} (y_{i,j} - d) \]

\[ = (1 - p_j(c)) y_{i,j} - \sum_{i} p_i(c) y_{i,j} \]

\[ + (1 - p_j(c)) y_{i,j} - \sum_{i} p_i(c) y_{i,j} \]

\[ + \cdots + (1 - p_j(c)) y_{i,j} - \sum_{i} p_i(c) y_{i,j} \]

\[ (c₁,⋯,cₙ,c_{recovery}) \]

(2) Assume that \( U_j(c) \) is the minimum pessimistic loss of defenders.

Step 2: Define protection functions \( f_j(c) \) and recovery functions \( f_j(c) \) for N elements.

Step 3: While \( r < R_0 \) do

1) According to current allocation \( (c₁,⋯,cₙ,c_{recovery}) \), compute the expected loss for all M targets and let the maximum one be \( U'_j(c) \). If there are two or more targets owning the maximum expected loss, randomly choose one, i.e. \( U'_j(c) = \max \{U_j(c), U'_M(c)\} \).

2) Compute all first-order partial derivatives of the function \( U'_j(c) \) for \( n+1 \) variables at current allocation \( (c₁,⋯,cₙ,c_{recovery}) \) and let the fastest descent direction be \( c'\). (If there are two or more variables owning the fastest descent direction, randomly choose one).

3) Allocate a \( \Delta R_0 \) to the variable in \( (c₁,⋯,cₙ,c_{recovery}) \) which corresponds to \( c'\) and update the \( r \).

Step 4: Output \( (c₁,⋯,cₙ,c_{recovery}) \) and \( U_1(c) \) to \( U_M(c) \).

The output \( (c₁,⋯,cₙ,c_{recovery}) \) is the final allocation plan, namely the reliable strategy of defenders. \( U_1(c) \) to \( U_M(c) \) are the final expected loss of all targets after deploy the reliable strategy. The maximum value, namely \( \max \{U_1(c),U_M(c)\} \) is the minimum pessimistic loss of defenders. \( \Delta R_0 \) is a minimum unit of the distribution.

According to the requirement of defenders for different precision, they can adopt different \( \Delta R_0 \), e.g. the minimum of the currency or some other units of the currency. The smaller the \( \Delta R_0 \), the more time need to implement the algorithm, but more accurate the result will be.

C. Allocation algorithm B

Step 1: Let \( (c₁,⋯,cₙ,c_{recovery}) = 0 \) and \( r = c₁ + c₂ + \cdots + cₙ + c_{recovery} \) and let \( U \) be the minimum pessimistic loss that defenders want to obtain.

Step 2: Define protection functions \( p_j(c) \) and recovery functions \( f_j(c) \) for \( N \) elements. Compute the initial expected loss for all M targets and let the maximum be \( U'_j(c) \).

Step 3: While \( (U'_j(c) > U) \) do

1) Compute all first-order partial derivatives of the function \( U'_j(c) \) for \( n+1 \) variables at current allocation \( (c₁,⋯,cₙ,c_{recovery}) \) and let the fastest descent direction be \( c'\).

2) Allocate a \( \Delta R_0 \) to the variable in \( (c₁,⋯,cₙ,c_{recovery}) \) which corresponds to \( c' \) and update the \( r \).

3) According to current allocation \( (c₁,⋯,cₙ,c_{recovery}) \), compute the expected loss for all targets and let the maximum be \( U'_j(c) \).

Step 4: Output \( (c₁,⋯,cₙ,c_{recovery}) \), \( U_1(c) \) to \( U_M(c) \) and \( r \).

The output \( (c₁,⋯,cₙ,c_{recovery}) \) is the reliable strategy of defenders. \( U_1(c) \) to \( U_M(c) \) are the final expected loss of all targets. The maximum among all targets must be not more
than the specified value $U$ and the $r$ is the final budget that is necessary to deploy the reliable strategy.

D. The implementation of defense strategy

Admittedly, it should be noted that for attack size $n=1$, the time of “calculating the expected loss for all $M$ targets” will exponentially increase with the growth of system size $N$. In order to effectively implement the two algorithms, here a pruning strategy is introduced. We set a threshold of damage and only consider those targets which can cause a loss more than the threshold. By this way, defenders can adjust the computing time.

In addition, the proposed framework provides sufficient flexibility for defenders to deploy a reliable strategy and some details should be designated beforehand by the defenders. They can be summarized as follows:

(P1) Defenders can consider several kinds of elements, i.e. transmission lines, generators, transformers, etc. being attacked;

(P2) Defenders can estimate $y_{[S]}$ in different concerns and many factors can be included, such as the price of failed equipments, the expense on maintenance, the cost of energy loss of supply to customers etc;

(P3) Defenders can define different protection functions and recovery functions for each element and choose arbitrary currency as the budget; and

(P4) Defenders can specify different thresholds in order to adjust the computing time.

Given the flexibility of the framework, the proposed criterion and the two algorithms are universal. That is, as long as defenders specify those details (P1)-(P4), the two proposed algorithms can be used to obtain reliable strategies.

V. CASE STUDY

For clarity, conciseness and easy illustration of this case study, we only consider transmission lines being attacked and $y_{[S]}$ is the energy loss only. Moreover, budget $R_0$ is considered as a dimensionless quantity, i.e. a quantity without any physical units. In addition, the two functions (21-22) used for transmission lines in [17] are also employed here.

\[
p_i(c_i) = 1 - \frac{c_i}{4 + c_i}
\]

\[
f_i(c_{re cov} \text{ MWh}) = \frac{10}{10 + c_{re cov} \text{ MWh}}
\]

The base recovery time $t_{base}$ is one hour [17]. By simplification, we can easily illustrate that how the proposed algorithms are carried out to explore reliable strategies in a period for different cases.

It should be noted that we just illustrate some simple cases in this simulation. Defenders can further develop the details (P1)-(P4) with more protection functions, recovery functions as well as $y_{[S]}$ for each target according to their real situation and precision requirement.

A. Five-bus system

The five-bus system [9] is a small system with five buses and six transmission lines, which can be applied to enable a clear illustrating the procedure of the proposed algorithms, the reliable strategies of defenders and the final playoffs in different cases. five-bus test system and its demands are displayed in Fig. 1. The line reactance is expressed on a base of 100 MVA and 138 kV and line capacities are all 100 MW. All generators have lower and upper generation bounds from 0 to 150 MW.

![Five-bus system](image)

**V. CASE STUDY**

For clarity, conciseness and easy illustration of this case study, we only consider transmission lines being attacked and $y_{[S]}$ is the energy loss only. Moreover, budget $R_0$ is considered as a dimensionless quantity, i.e. a quantity without any physical units. In addition, the two functions (21-22) used for transmission lines in [17] are also employed here.

**TABLE II**

<table>
<thead>
<tr>
<th>Targets</th>
<th>Expected Energy Loss (MWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1: 4-5</td>
<td>50</td>
</tr>
<tr>
<td>U2: 3-5</td>
<td>50</td>
</tr>
<tr>
<td>U3: 1-2</td>
<td>0</td>
</tr>
<tr>
<td>U4: 1-3</td>
<td>0</td>
</tr>
<tr>
<td>U5: 2-3</td>
<td>0</td>
</tr>
<tr>
<td>U6: 1-4</td>
<td>0</td>
</tr>
</tbody>
</table>

From Table III, we can see that when budget $R_0=10$, the reliable strategy of defenders is $c_1=5, c_2=5$, and others are 0, namely $c = (5, 5, 0, 0, 0, 0)$. Table IV shows that under this strategy, the maximum expected energy loss among all targets is $U_1 = U_2 = 22.22 \text{ MWh}$, which is the minimum pessimistic loss of defenders. It is obvious that the playoff function $L = \sum_{j=1}^{6} q_j U_j(c) \leq 22.22$ and the equality holds
when \( q = (q_1, q_2, 0, 0, 0, 0) \) where \( q_1 + q_2 = 1 \). If terrorists fully know the strategy of defenders, that is, they can know that \( U_1 \) and \( U_2 \) have the maximum expected energy loss, they must play the strategy \( q = (q_1, q_2, 0, 0, 0, 0) \) for pursuing the maximum gain 22.22 MWh. Otherwise, if terrorists play any others strategies, the playoff can not reach 22.22 MWh. Therefore, as long as defenders deploy strategy \( c = (5, 5, 0, 0, 0, 0) \), they can guarantee that their loss will not exceed the minimum pessimistic loss of 22.22 MWh. Likewise, from Tables III and IV, we can obtain other reliable strategies in different budgets and corresponding minimum pessimistic loss of defenders.

Furthermore, it can be noted that in Table IV, expected energy loss for \( U_1 \) and \( U_2 \) are both the same under different budgets. The reason is that at the beginning, \( U_1 \) and \( U_2 \) have the same expected loss without defending. Firstly, minimum unit \( \Delta R_0 \) is allocated to one of them e.g. \( U_1 \). At the next step, \( U_2 \) has the highest expected loss, \( \Delta R_0 \) should be allocated to \( U_2 \). Likewise, defenders must distribute budget to both of them alternately in order to keep them always the same. Otherwise, terrorists will attack the one owning the higher expected loss. In this case, the budget allocated on the lower one is a waste.

On the other hand, if we want to limit the minimum pessimistic loss to 5% of the initial pessimistic loss, namely 300*3% = 9 MWh, applying algorithm \( B \), we can achieve the reliable strategy of defenders as \( c_1=16, \ c_2=16, \ c_3=0.7809, \ c_4=0, \ c_5=0.7809, \ c_6=0, \ c_7=21.1111 \) with a total budget of 54.67.

### Case 2: \( n=2 \)

The target is comprised of two lines and there are 15 targets according to formula (6). Table V illustrates the expected energy loss (MWh) of each target after it is attacked without defending.

<table>
<thead>
<tr>
<th>Targets</th>
<th>Expected Energy Loss (MWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1: 3-5, 4-5</td>
<td>300</td>
</tr>
<tr>
<td>U2: 2-3, 4-5</td>
<td>100</td>
</tr>
<tr>
<td>U3: 1-4, 4-5</td>
<td>100</td>
</tr>
<tr>
<td>U4: 1-2, 4-5</td>
<td>100</td>
</tr>
<tr>
<td>U5: 1-3, 4-5</td>
<td>100</td>
</tr>
<tr>
<td>U6: 1-2, 3-5</td>
<td>100</td>
</tr>
<tr>
<td>U7: 1-3, 3-5</td>
<td>100</td>
</tr>
<tr>
<td>U8: 1-4, 3-5</td>
<td>100</td>
</tr>
<tr>
<td>U9: 2-3, 3-5</td>
<td>100</td>
</tr>
<tr>
<td>U10: 1-2, 2-3</td>
<td>40</td>
</tr>
<tr>
<td>U11: 1-2, 1-3</td>
<td>20</td>
</tr>
<tr>
<td>U12: 1-3, 2-3</td>
<td>0</td>
</tr>
<tr>
<td>U13: 1-2, 1-4</td>
<td>0</td>
</tr>
<tr>
<td>U14: 1-4, 2-3</td>
<td>0</td>
</tr>
<tr>
<td>U15: 1-3, 1-4</td>
<td>0</td>
</tr>
</tbody>
</table>

### B. IEEE Reliability test systems

The IEEE reliability test system (RTS) [24] has 24 buses, 32 generators, 38 lines and 17 loads. It is used for case study 3 as a more complex system compared with the five bus system of Fig1.
Case 3: \( n=1 \), i.e. the target is comprised of one transmission line. According to (6), there are 38 targets. Like Tables II and V, it is easy to obtain a table which represents the expected energy loss for the 38 targets without defending. For conciseness, we do not display it here. Generally speaking, for each budget, we can have a reliable strategy and a minimum pessimistic loss i.e. the least upper bound. With the increase of budget, the least upper bound will decrease, which can form a damage frontier. Fig.2 displays the frontier with the budget from 0 to 200. The red circles comprise the frontier, representing the minimum pessimistic loss of defenders. The final loss depends on the terrorists’ strategies. The region in blue stands for the possible loss of defenders.

![Fig.2 Damage frontier with the budget from 0 to 200](image)

VI. CONCLUSIONS

Power systems are among the most important critical infrastructures for a country. Severe power system blackouts may result in huge billion dollar losses. Furthermore, the failures of power systems usually will propagate into other critical infrastructures such as communications, water supply, natural gas and transportation etc., which will cause a even larger disturbance of a society as well as panic and fear among its citizens. Power systems reliability and security are essential for the electrical industry.

In recent years, with the extensively growth of terrorism activities, power systems probably become the target of terrorists. However, the current reliability and security framework is vulnerable against terrorist attacks, because terrorists can be highly intelligent and/or they can even hire scientists and power engineers to seek the vulnerability of power systems and then launch a vital attack. If this happens, the impact and the loss of a society can be immense.

This paper presents a new comprehensive and quantitative mathematic framework to study the new power systems security problem under potential terrorism threats. The interactions between the defenders and terrorists are formed as different games. Game theory is a useful mathematic tool by which terrorists can be modeled as fully intelligent and strategic players. We also derived a new criterion for reliable strategy design and two effective algorithms are also developed to acquire reliable strategies against terrorist attacks. When defenders deploy the strategy before terrorists launch an attack, the loss can be predictable and limited to a minimum level.

APPENDIX

The DC power flow equation can be described as [4,25]

\[ F = H \times P \]  

(A1)

where \( F = (F_1, F_2, \ldots, F_m)^T \) is the real power in \( m \) transmission lines. \( P \) is a vector whose components are the power of each node. \( H \) is a constant matrix. The reference node is not included in the vector \( P \) to avoid singularity of \( H \). In DC power flow model the susceptance matrix \( B \) is [25]

\[ B = A^T \cdot b \cdot A \]  

(A2)

where \( A \) is the network adjacency matrix \( A_{non} \) and \( b \) is a diagonal matrix with each entry representing the susceptance of each transmission line. Let \( \Theta \) be the voltage angle vector. It is easy to obtain the following two relations

\[ P = B \cdot \Theta \]  

(A3)

\[ F = b \cdot A \cdot \Theta \]  

(A4)

Combining (A2)-(A4) we can have

\[ H = b \cdot A \cdot B^{-1} \]  

(A5)

Normally, a power system is in a stationary state in which it operates with a feasible solution of power flow equations. When a target is attacked, some lines might be overloaded. In this case, it is necessary to redispatch the injected power to obey the system constraints and if those constraints can not be satisfied, load has to be shed to reach a new feasible solution. Furthermore, it is known that shed load counts for the energy loss of power to supply to customers. The loss should be minimized. Therefore, the redispatch of power flow can be formulated as a linear programming (LP) problem. The objective function of the problem, namely the load shedded is defined as [25]

\[ f = \min \sum_{j=load} c_j \]  

(A6)

which subjects to the equation (A1) and overall power balance [25]

\[ \sum_{i=generators} p_i + \sum_{j=loads} c_j - \sum_{j=loads} d_j = 0 \]  

(A7)

where \( p_i \) is the generated power for generator node \( i \), \( c_j \) is the load shedding for load node \( j \) and \( d_j \) is the initial load of \( j \). In addition, this minimization is implemented with the following constraints:

(a) Generation capacity limits for generator \( i \)

\[ p_i^{min} \leq p_i \leq p_i^{max} \]  

(A8)

(b) The constraints of load shedding limits for load \( j \)

\[ 0 \leq c_j \leq d_j \]  

(A9)

(c) The line flow limits

\[ |F_{ik}| \leq F_{ik}^{max} \]  

(A10)

where \( F_{ik}^{max} \) is the maximum line power flow of line \( k \).
REFERENCES


Zhaoyang Dong (M’99, SM’06) obtained his PhD from The University of Sydney, Australia in 1999. He is now an Associate Professor at The Hong Kong Polytechnic University, Kowloon, Hong Kong. He previously held academic positions with The University of Queensland, Australia. He also held industrial positions with Transend Networks, Tasmania, Australia. His research interest includes power system planning, power system security, load modeling, electricity market, and computational intelligence and its application in power engineering.

David J. Hill (M’76-SM’91-F’93) received the BE (Electrical Engineering) and BSc (Mathematics) degrees from the University of Queensland, Australia, in 1972 and 1974, respectively. He received the PhD degree in Electrical Engineering from the University of Newcastle, Australia, in 1976. He is currently an Australian Research Council Federation Fellow in the Research School of Information Sciences and Engineering at The Australian National University. He is also a Chief Investigator of the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems. He held academic and substantial visiting positions at the universities of Melbourne, California (Berkeley), Newcastle (Australia), Lund (Sweden), Sydney and Hong Kong (City University). He currently holds honorary professorships at the University of Sydney (Australia), City University of Hong Kong, South China University of Technology, Wuhan University and Northeastern University (China). His research interests are in network systems, stability analysis, nonlinear and distributed control and applications. He is a Fellow of the Institution of Engineers, Australia, the Society for Industrial and Applied Mathematics, USA, and the Australian Academy of Science; he is also a Foreign Member of the Royal Swedish Academy of Engineering Sciences.

YuSheng Xue (M’1987) obtained his PhD in Electrical Engineering from the University of Liege (Belgium) in 1987. He became a Member of Chinese Academy of Engineering in 1995 and has been the Chief Engineer at Nanjing Automation Research Institute (NARI), China since 1993. His research interests include nonlinear stability, control and power system automation.