Evaluating Strategies for Defending Electric Power Networks against Antagonistic Attacks

Åke J. Holmgren, Erik Jenelius, and Jonas Westin

Published paper available electronically at http://dx.doi.org/10.1109/TPWRS.2006.889080

***Abstract*—We show how concepts from game theory can be used to find and evaluate strategies for defending an electric power system against antagonistic attacks. Consequently, the interaction between the antagonist and the defender of the system is envisaged as a game. In a numerical example, we study the performance of different defense strategies against a number of attack scenarios. Particularly, we study whether there is a dominant defense strategy, and an optimal allocation of resources between protection of components, and recovery.**

***Index Terms*—Homeland security, blackout, game theory.**

I. Introduction

ELECTRIC power systems, from the beginning superimposed on the civilization, are today an indispensable part of the fabric of a modern society. Consequently, these systems make attractive targets to terrorists with a broad range of motives. Major power outages in the last decade clearly show that natural disasters, human factors, and technical failures, also need consideration. Here, we will, however, emphasize the national security ("Homeland security" in the U.S.) aspects of critical infrastructure protection, a topic that has received extensive interest lately (see e.g. [1]). Thus, we will consider physical antagonistic attacks against electric power grids (i.e. we will not include "cyber attacks").

During war and armed conflicts (e.g. in Iraq, and former Yugoslavia), military, or semi-military, attacks against power grids have resulted in major power outages. To the authors' knowledge, no major power outage in the Western world has originated from an antagonistic attack. For example, empirical data from the Swedish national transmission grid (1993–2003), and the Stockholm distribution grid (1998–2003), display a few recorded minor power outages caused by sabotage [2]. Further, there are few publicly reported sabotage attempts (near-misses). However, in 1996 police and the British Security Service (MI5) arrested members of a militant subgroup of the Irish Republican Army (IRA). The objective of their planned attack was a number of transformer stations, critical for the power distribution in London and south England. The investigations that followed showed that only a small amount of highly effective explosives would have been required (less than three kilo per station), and that the result could have been a power outage with duration of several weeks [3].

Since there is an interaction between the attacker and the defender of a system, studies of antagonistic attacks embrace a game situation rather than a decision situation. The measures applied to defend a power grid will affect an antagonist's course of action, which again will affect the defense, etc. In defense analysis, game theory is widely used to analyze effects of selecting alternative strategies to achieve a military objective. Further, penetration testing ("red teaming") is conducted to seek out technical and structural vulnerabilities in computer systems, but also for studying attack approaches and consequences of attacks.

In this paper, we introduce a model for studying strategies for defending electric power systems subject to different types

This work was supported in part by the Swedish Emergency Management Agency (SEMA) under Grant dnr. KBM 0054/2002.

Åke J. Holmgren is with the Swedish Defence Research Agency (FOI), S-164 90 Stockholm, Sweden (phone: +46-8-5550-8038; fax: +46-8-5550-3866; e-mail: ake.holmgren@foi.se), and the Div. of Safety Research, Royal Institute of Technology (KTH).

Erik Jenelius is with the Div. of Transport and Location Analysis, Royal Institute of Technology (KTH), Stockholm, Sweden (e-mail: jenelius@infra.kth.se).

Jonas Westin is with the Swedish Defence Research Agency (FOI), Stockholm, Sweden (e-mail: jonas.westin@foi.se).

of antagonistic threats. The interaction between the defender of the system and the antagonists is envisaged as a game. There has been some previous work done on related topics. For example, a similar systems defense game is formulated in [4]. Optimal protection strategies for simple attack scenarios and systems, with components either in series or in parallel, are derived in [5]. None of these models, however, considers the time dimension, and the defender cannot reduce the potential damage by reducing the time for recovery. In [6], a game is set up between a router, who seeks to minimize the travel cost for data packets, or vehicles, by choosing routes in a network, and an antagonist, who seeks to maximize the travel cost by destroying edges. A similar technique, using an electric power flow model, is used to identify critical components of a power grid in [7], is optimally solved in [8], and generalized in [9]. Finally, in [10], an overarching model for setting priorities among threats and countermeasures, based on probabilistic risk analysis, decision analysis and elements of game theory, is presented.

The rest of the paper is organized as follows. First, we introduce a simple electric power systems model. In Sections III and IV, we model the defense of the power grid, and the antagonist. Then, in Section V we describe the structure of the game, and formulate the optimization problem. After that, we provide a practical example, using an idealized version of the Swedish national transmission grid. Finally, we discuss the model, in particular its implications for decision-making, and provide some conclusions regarding possible applications and extensions.

II. The Electric Power Systems Model

The structure of an electric power system is described as an undirected and connected graph $G = (V, E)$, i.e. a vertex set $V(G)$, an edge set $E(G)$, and a relation that associates with each edge two vertices. The vertices in the graph represent power generation, stations (switching, transformation etc.), and users. Edges correspond to power lines. For simplicity, we will consider a lossless electric power grid with N elements (vertices and edges). A formulation of the maximum-flow optimization problem is given in the Appendix.

Since the methodology presented here is modular, the electric power systems model can be replaced with a more realistic one; compare with [7]–[9]. Also, professional, or educational, software packages for power systems analysis and design can be employed to calculate the consequences of attacks.

The network is maintained by a power systems operator (defender), and is subject to attacks by antagonists. An attack results in disabled elements in the network, which in turn may lead to loss of power for users. Let $x_i \geq 0$ be the power loss (MW) resulting from disabling element i. Depending on the structure of the network, the resulting power loss caused by disabling multiple elements is, in general, not additive, but always at least as great as for any subset of those elements. That is, if we let x_S denote the power loss caused by disabling element set S, then $x_S \geq x_{S'}$ for every subset $S' \subset S$. Let t_i be the recovery time (h), i.e., the time until element i is fully restored. As shown below, t_i can be controlled by the defender. The total consequence y_i is measured as the energy loss (MWh), which is approximated as the power loss times the recovery time, i.e., $y_i = x_i \cdot t_i$.

III. Defending the Electric Power System

A. Measures for Defending Electric Power Systems

The management of power outages consists of a number of phases, for example prevention, mitigation, response, recovery, and learning. Measures for prevention aim at reducing the likelihood of an attack or avoiding an attack. Mitigation aims at minimizing the negative consequences of an attack. Response includes the measures performed during the acute crisis phase in order to minimize the negative consequences of the attack. Finally, recovery involves all measures carried out to bring back the system to a normal state after an attack.

Some general defensive tactics for prevention and mitigation are: barriers and fortification; redundancy (to introduce additional, equivalent, components); diversity (applied to equipment, functions, and staff); training, quality control, and procedures review; preventive maintenance; monitoring, surveillance, testing and inspection [11]. The response to a power outage (emergency control) can to a large extent be based on the same principles as normal electric power system operations. Recovery (power systems restoration) includes determining the state of the system, preparing the equipment for restoration to service, reintegrating and rebuilding the system, and balancing generation and load as they are brought back to a normal level [12], [13].

B. A Mathematical Model of the Defense

We will assume that the defender can only spend resources on increasing component protection (e.g. fortification), and/or decreasing the recovery time after an attack (e.g. repair teams). We do not consider the possibility of adding components to the network. Let c_{total} be the total amount of resources available to the defender, thus

$$c_{\text{total}} = c_{\text{prevent}} + c_{\text{recovery}} , \tag{1}$$

where c_{prevent} is the budget for protecting components, and c_{recovery} is the budget for restoration in order to decrease the recovery time.

1) Protection: Every element i has a protection described by the parameter p_i. This parameter corresponds to the probability that an attack against element i fails. Consequently, we assume that the probability that the attack succeeds only depends on the defense of the attacked element p_i and not on the attacker (i.e. the attacker has enough resources and competence for performing a successful attack, see section IV). The probability that an attack against element i is successful is, therefore, equal to $(1 - p_i)$. We assume that a successful attack disables the element completely, and that all elements are disabled independently of each other.

The protection p_i of element i is a function of the resources c_i spent on protecting that element. The defender distributes the resources for protection between the N elements in the network. Thus, the protection is described by the vector $\mathbf{p} = (p_1, p_2,\ldots, p_N)$:

$$p_i = p_i(c_i), \ i = 1,\ldots, N \qquad (2)$$

$$0 \le p_i \le 1, \quad i = 1,\ldots, N \qquad (3)$$

$$\sum_{i=1}^{N} c_i \le c_{\text{prevent}} \qquad (4)$$

For the sake of the analysis, we assume that the defense functions $p_i(c_i)$ are continuous increasing functions, and that the marginal utility of spending resources on protection is diminishing. That is, the marginal cost of the defense p_i is an increasing function.

2) Recovery: The repair time of element *i* depends on the resources spent on recovery, as well as the type of the disabled element, and the attack method. We assume that the defender has a basic recovery capacity for maintenance and for repairing minor failures. We therefore let the constant t_i^{base} correspond to the time it takes to repair an element when no extra resources is spent on the recovery. Thus, if the defender chooses to spend extra resources on recovery, the recovery time t_i decreases, i.e.

$$t_i = t_i^{\text{base}} \cdot f_i(c_{\text{recovery}}) , \qquad (5)$$

where $f_i(\cdot)$ is a continuous decreasing function. The marginal cost of decreasing the recovery time is assumed to be an increasing function. If several elements are disabled, the defender can use different repair schemes depending on the available resources for, e.g., parallel repairs. These possibilities will not be modeled explicitly here, but we will assume that elements are repaired one at a time in the order that minimizes the total negative consequence. Different schemes may produce somewhat different results, but will not affect the analysis method itself, since the total consequence will be a convex decreasing function of c_{recovery} independently of whether elements are repaired simultaneously or one at a time.

The total allocation of defense resources can thus be described by the vector $\mathbf{c} = (c_1,\ldots, c_N, c_{\text{recovery}})$.

IV. The Antagonist

A. The Nature of the Antagonist

Antagonistic attacks form a broad category of threats that spans from insiders and saboteurs, to crime syndicates and transnational terrorist organizations. Even warfare can be included in this category. Attacks are different from random failures in the sense that the antagonist chooses what parts of the network that are attacked, and also the time of the attack. Antagonists can be classified according to a number of criteria (not necessary mutually exclusive), e.g. goals and motivation, tactics and modes of operation, resources, group size, knowledge and competence, origin, ideology, ethical constraints etc. The antagonist's degree of rationality and determination are other important factors (compare with [6]).

The purpose of an antagonistic attack can be to cause severe damage to a power grid in an attempt to disable important functions of a society. However, the goal with an attack might not always be to maximize the negative consequences. Instead, the objective of the attack can be to make a symbolic demonstration, or to cause a large enough consequence, in order to achieve a psychological effect such as a spread of fear and anxiety. Also, a threat does not always need to be realized, and sometimes it can be enough just to make a threat, without any real intention of realizing it. The attack can, rather than being a goal in itself, be seen as a mean to reach a higher goal (political, religious, economical etc.). What this higher goal is, and how the attack is supposed to help in achieving the goal is, nevertheless, not always obvious.

B. The Attack Model

We will only consider qualified antagonists. That is, determined, well-informed, and competent antagonists with access to enough resources to perform a successful attack against an electric power system. It is possible to construct an advanced mathematical model of the antagonist. In theory, it might be achievable to model the antagonist's behavior as a utility function that also captures the ethical restrictions, resources, knowledge, etc. In practice, however, we lack much of the necessary information about the antagonists.

In this paper we will, thus, use a more realistic approach in which antagonistic attacks against a system are captured by a broad set of attack scenarios that describes the attack strategy, tactics, and modes of operation.

1) Attack Strategy: The attacker's problem consists of choosing one of the available attack targets, by which we mean every combination of elements considered possible to attack. Each choice of target constitutes a (pure) strategy of the attacker. Let *T* be the set of targets and *M* the number of targets. If we consider only attacks on single elements, then *T* is the set of all elements and $M = N$. If we consider simultaneous attacks on exactly $n > 1$ elements, then *T* is the set of all unique combinations of *n* elements $\{i_1, i_2,\ldots, i_n\}$, and

$$M = \binom{N}{n} = \frac{N!}{n!(N-n)!}. \qquad (6)$$

We do not have to restrict the attacker to use only pure strategies. That is, the antagonist is allowed to randomize between which targets to attack. Let q_j correspond to the probability that target *j* is attacked, given that an attack is made. The vector $\mathbf{q}$ of dimension *M* then describes the mixed strategy, i.e.

$$q_j = P(\text{target } j \text{ is attacked} \mid \text{attack}) \qquad (7)$$

$$0 \le q_j \le 1, \ j = 1,\ldots, M \qquad (8)$$

$$\sum_{j=1}^{M} q_j = 1 \qquad (9)$$

The outcome of an attack against target *j* depends on the protection of the elements within the target and can be described as a stochastic variable Y_j. For an attack on a single element i,

$$\begin{cases} P(Y_i = y_i) = 1 - p_i \\ P(Y_i = 0) = p_i \end{cases} \tag{10}$$

where y_i is the consequences of disabling element i (Section II). For an attack on $n > 1$ elements, one must account for the possibility that only a subset of the target is destroyed. Since the consequences of destroying several elements, in general, are not additive, every subset of elements must be considered individually. Let T_j be the set of elements in target j, let S denote a subset of T_j and let y_S be the consequences of disabling the elements in S. Then

$$P(Y_j = y_S) = \prod_{i \in S} (1 - p_i) \prod_{i \notin S} p_i . \tag{11}$$

Based on the discussion above (Section A), we will consider the following three different classes of attacks:

i. *Worst-Case Attack:* The antagonist chooses the target that maximizes the expected negative consequences $\mu_j = E(Y_j)$ of the attack.

ii. *Probability-Based Attack:* The antagonist tries to maximize the probability that the outcome Y_j of an attack is over a certain magnitude $y_{\min}$, i.e. $P(Y_j > y_{\min})$.

iii. *Random Attack:* The antagonist chooses the attack target randomly. Each target is attacked with equal probability.

The attack models of [7]–[9] differ from the present one in that they are completely deterministic. That is, an attack against an element will disable it for certain, and the attacker does not randomize over which target to attack. In [7] and [8] the objective of the attackers is to maximize the negative consequences. In [9] various objectives can be analyzed, including maximizing the consequences or minimizing the resources required to achieve consequences above a certain level.

In the Swedish dataset [2] mentioned above, there are only a few minor disturbances caused by sabotage. In all the cases in point, the antagonists have targeted installations that are easy to access or do not require any specific knowledge about the electric power system. Thus, from an electric power systems view, these attacks could be seen as a more or less random selection of attack target.

2) Tactics and Modes of Operation: In order to capture the antagonist's course of action in more detail, we will also make it possible to specify tactics and modes of operation.

An attack scenario is constructed by specifying the attack strategy class, and a few additional parameters. The aim is to make the attack scenario more realistic by adding a few conditions and restrictions. For example, an attack against an overhead power line requires less resources, and competence, than an attack against a station or a generator. Also, this makes it possible to assign a different recovery time to an element depending on the attack method etc. In our model, attack tactics and modes of operation are captured by three parameters: type of attack method, type of attack target, and attack size (n). Examples will be given in Section VI (Table 2).

V. The Interaction Between Defender and Antagonist

For the Worst-Case and Probability-Based Attack classes, the interaction between the defender and the antagonist can be described as a two-player zero-sum (strictly competitive) game, where, simultaneously, the defender chooses an allocation of defense resources, and the antagonist chooses a target to attack. Thus, we assume that the defender's payoff is the negative value of the attacker's payoff. We will also assume that "all cards are open", i.e., both the defender and the antagonist have complete information about the system, and the resources and preferences of the other. The situation where the defender and/or the antagonist have no, or limited, information about the other's preferences is briefly discussed under Section 4 below.

We now formulate the problems corresponding to each of the three attack classes.

1) Worst-Case Attack: The situation where the attacker tries to maximize and the defender tries to minimize the total expected damage can be translated into the following optimization problem, with the restrictions given by (1)–(5) and (7)–(9):

$$\max_{\mathbf{q}} \left[\min_{\mathbf{c}} \sum_{j=1}^{M} \mu_j(\mathbf{c}) \cdot q_j \right] \tag{12}$$

For an attack on a single element i, $\mu_i = (1 - p_i) \cdot y_i$. Following (11), we can calculate the result for attack size $n > 1$. For example, with $n = 2$ and an attack on target j consisting of elements i_1 and $i_{2,}$

$$\mu_j = (1 - p_{i_1})(1 - p_{i_2}) y_{\{i_1, i_2\}} + (1 - p_{i_1}) p_{i_2} y_{i_1} + p_{i_1}(1 - p_{i_2}) y_{i_2} \tag{13}$$

In this type of game there exists a Nash equilibrium where neither the attacker nor the defender can increase their payoff by choosing another strategy. This follows from that **q** and **c** are compact, convex subsets of a Euclidian space, and the payoff functions of the defender and the antagonist described by (12) are quasi-concave and continuous [14].

2) Probability-Based Attack: The situation where the attacker tries to maximize, and the defender minimize, the probability that the consequences are above a certain threshold $y_{\min}$ can be formulated similarly to the Worst-Case Attack problem:

$$\max_{\mathbf{q}} \left[\min_{\mathbf{c}} \sum_{j=1}^{M} P(Y_j > y_{\min}) \cdot q_j \right] \tag{14}$$

Let S denote a subset of elements within the target. Then let the indicator variable I_S be 1 if the consequences of destroying the element set S are larger than $y_{\min}$ and 0 otherwise, i.e.,

$$I_S = \begin{cases} 1 & \text{if } y_S > y_{\min} \\ 0 & \text{if } y_S \le y_{\min} \end{cases} \tag{15}$$

By substituting every y_S in the Worst-Case Attack with I_S, we obtain the desired problem formulation. The defender can control I_S by varying the proportion of resources spent on recovery. For this problem, we can, however, not guarantee the existence of a Nash equilibrium, since the payoff function is not continuous in c_{recovery}, compare with (15).

3) Random Attack: When the antagonist chooses the attack target randomly, there is no interaction between the defense and attack strategies. For the defender, the situation thus changes from a game to a decision problem. We assume that the defender wishes to minimize the total expected consequences of the attack. The attack strategy is fixed to $q_j = 1/M$, $j = 1,\dots, M$, and the problem becomes

$$\min_{\mathbf{c}} \frac{1}{M}\sum_{j=1}^{M}\mu_j(\mathbf{c}). \qquad (16)$$

This problem can be seen as a special case of the Worst-Case Attack problem, and has a well-defined solution.

4) Limited Information: It is often the case that neither the attacker, nor the defender, has full information about each other's choice of strategies or the consequences of different attacks. Also, an antagonist might act irrational, or in other ways not correspond to the payoff maximizing rational player that is assumed in game theory. Some of these situations can be described as games with incomplete information where a so-called Bayesian Nash equilibrium can be applied [14]. Instead of trying to model these situations, the expected consequences of the Worst-Case Attack and the Random Attack (*ceteris paribus*) can be used as a span between which the expected consequences of an unknown attack strategy will lie. Accordingly, by studying how these two boundaries are affected by different defense measures, we can make a rough evaluation of defense strategies.

VI. Application to the Swedish Transmission Grid

A. General Premises

1) Rationale of the Example: In this section, we will illustrate how strategies for defending a network can be evaluated. We will use the Swedish national high voltage transmission system (400 kV and 220 kV voltage levels) as a practical example. The purpose of the example is, however, not to evaluate this particular power system. Svenska Kraftnät (the state utility that manages and operates the national electric grid) has provided us with basic information about the network, and allowed us to disclose some results. We will not use authentic data on capacities and lengths of the power lines, generation, or power transmission to regional distribution grids. However, the numerical assumptions in the example are validated by Svenska Kraftnät, and can thus be seen as somewhat reasonable expert assessments.

General information regarding the Swedish national electric grid, including the approximate location of major generators, power lines, and stations, in northern Europe, is also published on Svenska Kraftnät's website [15].

2) The Network Model: The vertices represent major generators (source vertices), and regional power grids (sink vertices) that deliver the electricity to the customers. Edges correspond to overhead power lines (Section II and Appendix). The location of the generation, and the energy mix (mainly hydropower and nuclear power), is realistic. The demand has been spread out over most of the country, where the demand of a normal regional power grid is set to one power unit (Figure 1).

The total load on the system, i.e. the sum of all users' demand, describes the operational situation. Let u be the maximal possible supply capacity, that is, when every generator produces at maximum, there is full import of power, and the power lines are used at recommended maximal capacity. We differentiate between three different operational situations: i) "Normal conditions" ($0.75\cdot u$), ii) "Cold winter" ($0.95\cdot u$), and iii) "Extreme winter" (u).

In the model, a regional power grid is connected to the national transmission grid via one infeed point, which represents the entire regional grid (a sink vertex). In reality, users are of course also supplied from regional and local power plants, and measures can be taken to reduce the power demand if the grid becomes unstable. These two factors can to some extent compensate for a small supply shortage from the national transmission grid, and increase the ability of a regional grid to stay connected during a disturbance.

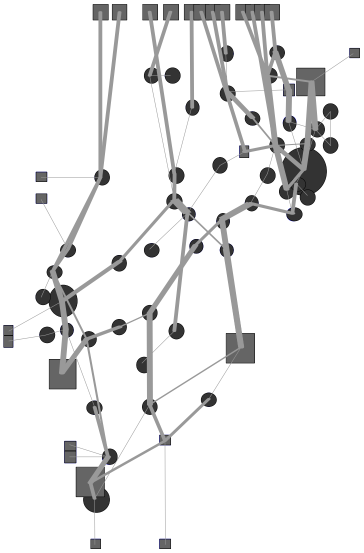

Fig. 1. Supply, demand, and edge capacity in the network. The circles correspond to the consumers' demand, the squares are the supply of generators, and the lines are the edge capacities. The size of the markers corresponds to the capacity of generators and consumers' demand. The thickness of a line is related to the edge capacity.

Thus, we allow the supply s_i to a sink vertex i (regional power grid) to be lower than the demand D_i. If s_i is lower than $\alpha \cdot D_i$, where $0 \le \alpha \le 1$, the regional grid will be disconnected in an attempt to avoid further disturbances. Here, we will assume that $\alpha = 0.95$.

B. Assumptions About the Defense and the Antagonist

1) Defense Cost Functions: We lack an empirical protection function, and will use the following function as an example:

$$p_i(c_i) = \frac{c_i}{k_i^p + c_i} \tag{17}$$

By changing the protection cost parameter k_i^p, we can to some degree account for the different costs of protecting different types of elements. Here, we set k_i^p to 2 for regional power grids (sink vertices), 3 for generators and 4 for power lines. This is related to the difficulty of protecting overhead power lines against attacks. This function is chosen mainly because it is one of the simplest functions satisfying the conditions in Section III. There are of course numerous other reasonable functions, such as the exponential form $p_i(c_i) = 1 - e^{-\lambda_i c_i}$, where λ_i is a parameter. Given that the function is suitably fitted to data, we believe that the results are fairly insensitive to this choice.

We model the recovery time t_i with the following function, subject to the characteristics described in Section III:

$$t_i(c_{\text{recovery}}) = t_i^{\text{base}} \cdot \frac{k_i^t}{k_i^t + c_{\text{recovery}}} \tag{18}$$

The parameter k_i^t is set to 30 for generators, 20 for users, and 10 for power lines. Further, the basic recovery time t_i^{base} depends on the type of element, and the antagonist's tactics and modes of operation. The numbers in Table 1 are obtained by means of expert assessments.

TABLE 1
BASIC RECOVERY TIME (h) FOR THREE DIFFERENT ATTACK METHODS

Element	**Low Damage**	**Moderate Damage**	**High Damage**
Power line	1	10	24
Station	2	20	96
Generator	3	30	192

2) Attack Scenarios and Defense Strategies: We will not consider multiple attacks separated in time, which can be a way of wearing down the defense. A vertex in our model (i.e. a bus) consists of several technical components (transformers, busbars, protective and control equipment, etc.). Consequently, to practically disable a vertex would most likely involve targeting several different local facilities. We will assume that the antagonist is capable of coordinating at most an attack of size n = 2.

In Table 2, all possible attack scenarios are summarized. The head of each column contains a parameter, and in the rows below the conditions that the parameters can assume are shown. A scenario is described by selecting one element from each column. Every possible combination of elements does not have to be realistic. In order to illustrate different aspects of the model, we will select 12 of the scenarios (Table 3).

TABLE 2
ATTACK SCENARIO PARAMETERS AND POSSIBLE VALUES

Operational Situation	**Attack Strategy**	**Attack Method**	**Attack Target**	**Attack Size**
Normal	Random	Low damage	Power Line	1
Cold winter	Probability-Based	Moderate damage	Station	2
Extreme winter	Worst-Case	High damage	Generator	
			Combination	

TABLE 3
SELECTED ATTACK SCENARIOS (FOR ALL SCENARIOS THE ATTACK METHOD IS "MODERATE DAMAGE", AND THE ATTACK TARGET IS "COMBINATION")

Attack Scenario	**Operational Situation**	**Attack Strategy**	**Attack Size (n)**
A1	Normal	Random	1
A2	Normal	Worst-Case	1
A3	Normal	Probability-Based	1
A4	Extreme	Random	1
A5	Extreme	Worst-Case	1
A6	Extreme	Probability-Based	1
A7	Normal	Random	2
A8	Normal	Worst-Case	2
A9	Normal	Probability-Based	2
A10	Extreme	Random	2
A11	Extreme	Worst-Case	2
A12	Extreme	Probability-Based	2

C. Simulation Results

1) Technical Notes: Calculating the maximum flow of the network is a linear programming problem. We want to maximize the flow between the source vertices and the sink vertices (Appendix). This problem can be solved with the network simplex method. After we have calculated the negative consequences of all considered attack combinations, the Nash equilibrium can be found using a min-max solver. For simplicity, we have used the Optimization Toolbox in MATLAB 6.5 [16].

2) Optimal Defense Strategies: It is possible to calculate optimal defense strategies for all attack scenarios (A1,…, A12). That is, a strategy with an outcome that, given the attack scenario, is at least as good as that of any other strategy. A common design criterion for transmission grids is the so-called "*N*–1 Criterion", i.e. the whole system must be capable of operating normally even if one major failure occurs. (To calculate the optimal defense against scenarios involving multiple attack targets is also very time consuming for a large network). Consequently, we have only calculated the optimal defense for scenarios involving attacks of size n = 1. The optimal defense strategy against attack scenario A1 is denoted D1, the optimal strategy against A2 is denoted D2, and so on.

Table 4 shows the expected consequences μ of attack scenarios (A1,…, A12) for the different defense strategies (D1,…, D6), given the budget $c_{\text{total}} = 100$. The lowest value for each attack scenario is marked in bold. A dominant strategy is a defense strategy with lower expected negative consequence against every attack scenario than every other defense strategy. In Table 4 we can see that no such strategy exists, which is to be expected.

TABLE 4
EXPECTED CONSEQUENCES (POWER UNITS) FOR DIFFERENT COMBINATIONS OF ATTACK SCENARIO AND DEFENSE STRATEGY

Attack	Defense Strategy					
Scenario	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**
A1	**2.0**	2.5	5.0	2.5	3.0	3.2
A2	33	**15**	166	73	61	65
A3	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
A4	36	50	65	**32**	37	40
A5	314	432	641	220	**121**	189
A6	192	208	170	172	121	**64**
A7	**4.8**	6.1	11.3	5.6	7.0	7.4
A8	190	260	340	135	**127**	136
A9	190	260	166	135	84	**66**
A10	81	112	144	**71**	83	91
A11	703	966	1187	435	**423**	559
A12	122	65	**46**	189	189	189

In Fig. 2, we show the expected negative consequence μ as a function of the total amount of resources spent on protection c_{total}. The defense strategies D4 and D5 are evaluated against the two attack scenarios A4 and A5. The marginal decrease in expected negative consequence is a decreasing function of the resources spent on the protection c_{total}. The difference between defending against a Worst-Case Attack and a Random Attack can be illustrated as the span between e.g. (D4;A5), and (D4;A4). (Compare with the discussion about limited information in Section V.)

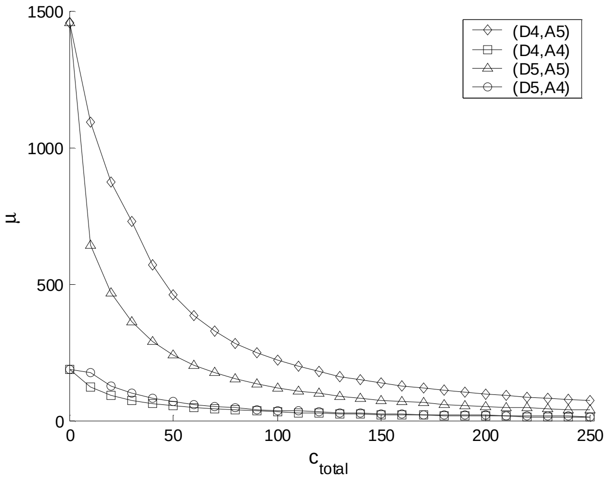

Fig. 2. Expected negative consequences μ as a function of the total defense budget c_{total} for different combinations of attack scenario and defense strategy.

3) Balancing Prevention and Recovery: There exists an optimal allocation between measures for protection and recovery for the scenarios above. This proportion depends on the total amount of resources c_{total} and the attack scenario. During an extreme situation there are more elements whose failure will cause large negative consequences compared to the normal situation. In this situation it is therefore more effective to spend a larger fraction of the resources on recovery than during the normal situation.

For the Worst-Case scenarios, e.g. (D2;A2) and (D5;A5), more elements become likely targets when the defense budget increases. That is, more and more elements will yield the same expected negative consequence if attacked. Therefore, it will be increasingly interesting to spend resources on recovery (Fig. 3). In the Random Attack on the other hand, e.g. (D1;A1) and (D4;A4), every component is a possible target. Because of the large number of components that need protection, it will be more cost-effective to spend resources on recovery. When the budget increases, the marginal gain from the extra resources spend on recovery will be lower and lower. As a result, the fraction of resources spent on recovery decreases when the total amount of resources c_{total} increases.

VII. CHOOSING DEFENSE STRATEGY

It is well known that results in game theory depend significantly on how the problem is framed, i.e. the structure of the game. As shown in the example above, a defense optimized against, e.g., the Worst-Case Attack strategy will not necessarily provide an optimal defense against other attack scenarios. Here, we will discuss this dilemma further, and also how the defender can choose between different strategies.

In Fig. 4, the optimal defense strategy against a Worst-Case Attack is shown. Since the antagonist will strike at the elements that yield the largest expected consequences μ, the defender can lower μ by placing the protection on these elements. The more resources the defender spends on the total protection, the more elements will be protected, and all of these elements will cause the same expected negative consequences. The optimal defense against this attack strategy is, thus, to protect the elements in the order of their criticality, starting with the element that causes the largest negative consequences if disabled.

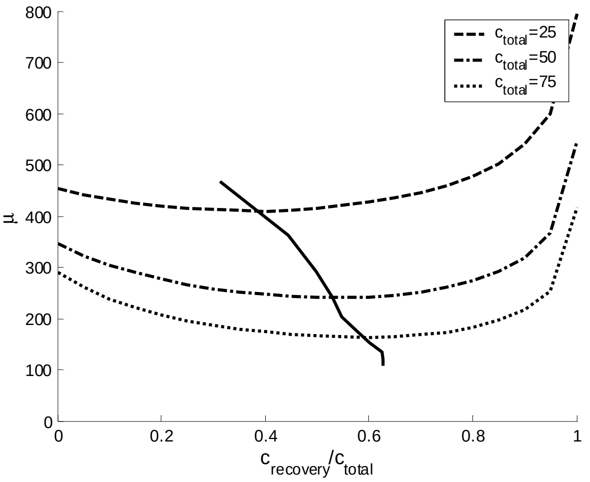

Fig. 3. Balance between protection and recovery under (D5;A5). The dotted lines show the expected negative consequences for three different total amount of resources c_{total} as a function of the fraction $c_{\text{recovery}}/c_{\text{total}}$. The solid line shows the optimal distribution between protection and recovery for different budgets c_{total}, i.e. the minimum of the dotted lines. Extra calculations have been made to find the optimal distribution for c_{total} between the horizontal lines.

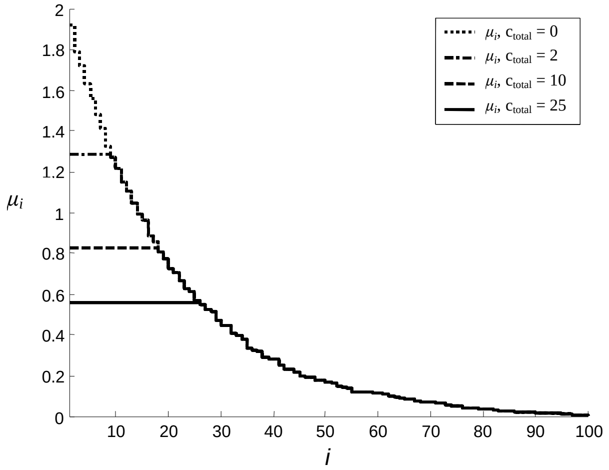

Fig. 4. The expected consequences μ_i of disabled elements in a fictitious network for different defense budgets c_{total}. The horizontal axis contains the different elements i, sorted in order of possible consequence y_i.

The optimal defense strategy against a Worst-Case Attack will, however, not give the optimal defense against a Probability-Based Attack (Fig. 5). Under the assumption that there exists elements for which $y_i > y_{\min}$ and that the defender cannot afford to protect all those elements equally well, the antagonist will choose to attack that with the lowest protection of those elements. Since the Worst-Case defense strategy above dictates that elements are to be protected according to their degree of criticality, it is likely that the antagonist in this situation will choose to attack the "first" unprotected element. Consequently, distributing the protection over a larger number of elements will give a better protection against a Probability-Based Attack than concentrating the protection on the most critical elements.

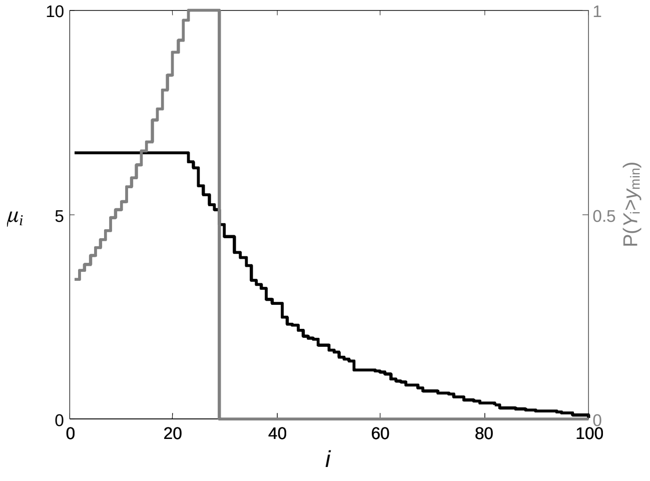

Fig. 5. Expected consequences μ_i of disabled elements in a fictitious network. The horizontal axis contains the different elements i, sorted in order of possible consequence y_i. Defending the network against Worst-Case Attacks, the antagonist using a Probability-Based Attack strategy will attack the first unprotected element ($y_i > y_{\min}$).

As shown in Table 4, there is in general no dominant defense strategy. An important question, thus, is how to choose a defense strategy. We can use a number of statistical methods to give a ranking of the different defense strategies. A problem with this approach, however, is that the relative likelihood of each attack scenario is highly uncertain. Accordingly, we cannot calculate the total expected consequence of the defense strategies (D1,…, D6) above. We will discuss three different ways of comparing the different defense strategies against each other.

In the first comparison, we study the order of the defense strategies. A simple method is to use the sum of the rankings as an indicator of the total ranking, and assign an equal weight to every attack scenario in the summation. But it would also be possible to put different weights to the different attack scenarios, e.g. based on estimations of their likelihood or the size of the expected negative consequence.

By using ordered statistics we overlook the relative difference between the different defense strategies within a certain attack scenario. In the second comparison, we calculate by how many percent the expected consequences of each defense strategy differ from the mean expected consequences in that attack scenario. By doing this we can measure the relative effectiveness of the defense strategies. If we take the sum of the relative difference, we can create a simple measure of the "best" defense strategy. As a third comparison we can also use the sum of the expected negative consequences μ for the different defense strategies over the 12 attack scenarios.

The three comparisons indicate that defense strategies optimized for the extreme operational situation is more effective than the defense strategies for the normal situation. This suggests that the defender should optimize the defenses for the worst situation rather than for the normal. The defense strategy D5, equivalent to minimizing the worst-case of a single attack during an extreme operational situation, has the lowest sum in the first test, the lowest sum of relative difference in the second test, and the lowest sum of μ for the attack scenarios (A1,…, A12) in the third test.

VIII. Discussion

In this paper, we have shown how concepts and models from game theory can be used when evaluating strategies for defending an electric power system against antagonistic attacks. The most important point of this paper, however, is not the particular game model itself, but rather the way to think about and formulate these issues. The game model has deliberately been kept simple in order to not obscure the general idea, and to guarantee the existence of optimal solutions.

In order to be able to use the model in practical decision-making, two main issues must be addressed. Firstly, a more realistic electric power model must be employed. Since the framework is modular, and there are suitable models described in the literature, we argue that this should be fairly simple (compare with Section II). Secondly, with a better understanding of the antagonistic threat it would perhaps be possible to assign probabilities to different attack scenarios. These probabilities could be based on the amount of resources and information that each attack strategy would require. However, studying antagonistic attacks, we will, to some degree, always face a genuine uncertainty. To make use of an even more elaborate model will not compensate for the lack of input data. That is, we will have to rely on expert judgments and sensitivity analysis when developing more detailed attack scenarios and estimating their corresponding parameters. The defense cost functions will have to be validated and calibrated. This is possible is theory since it is the power systems operator that makes the decisions about the allocation of the defense resources. It should also be noted that a more detailed modeling of the electric power system, including the defense cost functions, and the attack scenarios will most likely be classified, and thus cannot be reported in open sources.

The rationale for using a game theory model as described in this paper is threefold. Given the adjustments discussed above, the modeling framework can be useful in coarse resource allocation planning. The analysis can, thus, be a first step in a screening process for finding areas where more detailed analysis is required. Further, the model can be used to study generic mechanisms in order to enhance the overall understanding of attacks against electric power systems. It is well known that theoretical results in game theory depend significantly on how the game situation is modeled (the set of players, the set of strategies for each player, the order in which decisions are made, etc.). However, to use concepts and general models from game theory can be a powerful way of framing the problem. Finally, it is important to point out that an important contribution of all kinds of risk analysis is the actual work process itself. That is, the mathematical attack modeling creates a tangible result that can facilitate the thought process, bring together different stakeholders in the strategic planning process, and raise the awareness of these issues in the organization.

IX. Future Work

Given what has been said above, there are a number of possibilities for future technical refinements of the model. For example, the objectives of the defender and the attacker need not be the complete opposite of each other. Also, the assumption that the attacker has complete knowledge of the electric power system and the defender's resources may be weakened. The uncertainty regarding the outcome of an attack could be represented using stochastic variables. A way for the systems operator to reduce the consequences of an attack is to increase the redundancy of the network, i.e., to add new components. This option could be included in the model, given that costs are assigned to all such possible reinforcements. Further, mixing the game theoretical approach with a bit of chance can be a way of reducing the degree of rationality of an attacker.

Finally, it is important to emphasize that beside the antagonistic threats, electric power networks are subject to technical component failures and weather related disruptions. How to minimize the consequences in these cases is overlapping the present problem. For example, increasing the redundancy of the network would likely decrease the consequences of technical failures and extreme weather also. Similarly, decreasing the recovery time would likely be beneficial in all cases. It is also likely that the resources available for defense against all these threats are constrained by the same budget. Therefore, all sources of possible power disruptions should ideally be analyzed within a common framework. Then an overall best approach to reducing the risk of major power outages could be found.

Acknowledgment

The authors would like to thank Torbjörn Thedéen (KTH), Lars-Göran Mattsson (KTH), and Stefan Arnborg (Svenska Kraftnät) for valuable contributions and discussions.

Appendix

The standard maximum-flow optimization problem must be adjusted for the fact that elements may be disabled and, subsequently, repaired. The problem is thus time-dependent. We consider a lossless electric power grid with m vertices, power input ϕ_{it}^{in} and output ϕ_{it}^{out} at vertex i, and power flow ϕ_{ijt} (MW) from vertex i to j, at time t. Let Φ_t denote the matrix (ϕ_{ijt}), and $\boldsymbol{\varphi}_t^{\text{in}} = (\phi_{1t}^{\text{in}},\ldots,\phi_{mt}^{\text{in}})$. We also introduce the parameters δ_{it}, which is 1 (or 0) if vertex i is functioning (or disabled) at time t, and δ_{ijt}, which is 1 (or 0) if edge (i, j) is functioning (or disabled) at time t.

The time-dependent maximum-flow problem can be formulated as:

$$\max_{\Phi_t,\boldsymbol{\varphi}_t^{\text{in}}} \sum_{i=1}^{m} \phi_{it}^{\text{out}} \tag{A.1}$$

Subject to:

$$\phi_{it}^{\text{in}} - \phi_{it}^{\text{out}} = \sum_{j\neq i} \phi_{ijt}, \qquad i = 1,\ldots,m \tag{A.2}$$

$$\sum_{i=1}^{m} (\phi_{it}^{\text{in}} - \phi_{it}^{\text{out}}) = 0 \tag{A.3}$$

$$\phi_{ijt} = -\phi_{jit}, \qquad i, j = 1,\ldots,m \tag{A.4}$$

$$0 \le \phi_{it}^{\text{out}} \le \delta_{it} D_i, \qquad i = 1,\ldots,m \tag{A.5}$$

$$0 \le \phi_{it}^{\text{in}} \le \delta_{it} S_i, \qquad i = 1,\ldots,m \tag{A.6}$$

$$\left|\phi_{ijt}\right| \le \delta_{ijt} L_{ij}, \qquad i, j = 1,\ldots,m \tag{A.7}$$

Equations (A.2) correspond to Kirchoff's first rule, (A.3) stand for the conservation of energy, and (A.4) is the skew symmetry. Inequalities (A.5) are demand constraints, (A.6) are supply constraints, and (A.7) are the capacity constraints in the edges. D_i, S_i, and L_{ij} are nonnegative constants. Note that ϕ_{it}^{out} corresponds to s_i in Section VI.

References

[1] D. G. Kamien (ed.), *The McGraw-Hill Homeland Security Handbook*. New York: McGraw-Hill, 2006.

[2] Å. J. Holmgren and S. Molin, "Using disturbance data to assess vulnerability of electric power delivery systems," to appear (accepted October 2005) in *J. Infrastruct. Syst*.

[3] H. Christiansson and G. Fischer, “Cyberterrorism – en ny terrortyp,” in G. Jervas (ed.), *Terrorismens tid* (in Swedish). Stockholm: SNS Förlag.

[4] M. Shubik and R. J. Weber, “Systems defense games: Colonel Blotto, Command and Control,” *Naval Res. Logistics Quarterly*, vol. 28, no. 2, pp. 281-287, 1981.

[5] W. M. Bier, A. Nagaraj and V. Abhichandani, “Protection of simple series and parallel systems with components of different values,” *Rel. Eng. & Syst. Safety*, vol. 87, no. 3, pp. 315-323, 2005.

[6] M. G. H. Bell, “The use of game theory to measure the vulnerability of stochastic networks,” *IEEE Trans. Rel.*, vol. 52, no. 1, pp. 63-68, 2003.

[7] J. Salmeron, K. Wood, and R. Baldick, “Analysis of electric grid security under terrorist threat,” *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 905-912, 2004.

[8] A. L. Motto, J. M. Arroyo, and F. D. Galiana, “A Mixed-Integer LP Procedure for the Analysis of Electric Grid Security under Disruptive Threat,” *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1357-1365, 2005.

[9] J. M. Arroyo and F. D. Galiana, “On the solution of the bilevel programming formulation of the terrorist threat problem,” *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 789-797, 2005.

[10] E. Paté-Cornell and S. Guikema, “Probabilistic modeling of terrorist threats: a systems analysis approach to setting priorities among countermeasures,” *Military Oper. Res.*, vol. 7, no. 4, pp. 5-20, 2002.

[11] G. W. Parry, “Common cause failure analysis: a critique and some suggestions,” *Rel. Eng. & Syst. Safety*, vol. 34, no. 3, pp. 309–326, 1991.

[12] J. J. Ancona, “A framework for power system restoration following a major power failure,” *IEEE Trans. Power Syst.*, vol. 10, no. 3, pp. 1480–1485, 1995.

[13] M. M. Adibi and L. H. Fink, “Power system restoration planning,” *IEEE Trans. Power Syst.*, vol. 9, no. 1, pp. 22–28, 1994.

[14] D. Fudenberg and J. Tirole, Game theory. Cambridge: MIT Press, 1991.

[15] Svenska Kraftnät (SvK): http://www.svk.se

[16] Documentation for MATLAB 6.5, The MathWorks Inc.