Security Games for Voltage Control in Smart Grid

Yee Wei Law Tansu Alpcan Marimuthu Palaniswami
Department of Electrical & Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia
Email: {ywlaw, tansu.alpcan, palani}@unimelb.edu.au

***Abstract*—Information and communication technologies bring significant improvements to power grid and help building a "smart grid". At the same time, they cause novel vulnerabilities making the power grid, which is a critical infrastructure, susceptible to malicious cyber attacks such as false data injection. This paper develops a game-theoretic approach to smart grid security by combining quantitative risk concepts with decision making on protective measures. Specifically, the interaction between malicious attackers and grid defense systems is modeled as a security game, where the attackers choose the intensity of false data injection and defenders determine the detection threshold level. The consequences of data injection attacks are quantified using a risk assessment process based on realistic system simulations. The simulation results are used as an input to a stochastic game model, where the decisions on defensive measures are made taking into account resource constraints represented by cost values. Thus, security games provide a framework for choosing the best response strategies against attackers in order to minimize potential risks. The framework developed is also useful to analyse different types of attacks and defensive measures. The theoretical results obtained are demonstrated using numerical examples.**

I. Introduction

The *smart grid* is "the integration of power, communications, and information technologies for an improved electric power infrastructure serving loads while providing for an ongoing evolution of end-use applications" [1]. Being a critical infrastructure, a smart grid must be protected against potential threats. While system security[1] is an important issue for grid operators, real world constraints such as resource limitations necessarily force adoption of a risk management approach to the problem. Protective measures are usually taken based on a cost-benefit analysis balancing available defensive resources with perceived security risks.

This paper investigates cyberphysical security of smart grid by focusing on the important class of false data injection attacks which directly affect the operation of voltage control systems and potentially lead to blackouts. The problem is formulated within a quantitative risk context and then as a stochastic (Markov) security game. The resulting game analysis helps smart grid operators to make informed decisions on their security strategies while taking into account their resource constraints. Although the paper focuses on a certain type of attack and subsystem, the approach can be applied to similar security problems in smart grid, and hence, can be extended to develop the foundation of a systematic framework for smart grid security.

[1]This is to be differentiated from "power system security", which refers to the ability of a power system to survive plausible contingencies without interruption to supply.

A simple but elegant definition of risk is "the probability and magnitude of a loss, disaster, or other undesirable event" [2]. **Security risk analysis** can be defined as "the process of identifying the risks to system security and determining the likelihood of occurrence, the resulting impact, and the additional safeguards that mitigate this impact" [3]. Most smart grid standards and guidelines, e.g. IEC 62351-1, NISTIR 7628, identify risk assessment as a critical part of a security framework. For instance, the Australian Government advocates the use of the Australian and New Zealand Standard for Risk Management (AS/NZS ISO 31000:2009) by owners and operators of critical infrastructure [4]. However, the standard ISO 31000:2009 is "not mathematically based", and has "little to say about probability, data, and models" [5].

Security games provide an analytical framework for modeling the interaction between malicious attackers, who aim to compromise smart grid, and operators defending them. The "game" is played on smart grids, which are complex and interconnected systems. The rich mathematical basis provided by the field of game theory facilitates formalising the strategic struggle between attackers and defenders for the control of the smart grid [6]. Utilising the risk framework and some of the concepts of earlier studies [7]–[9], this work applies game theory to the modeling of attacks on and defenses for a critical voltage control component called the **static synchronous compensator (STATCOM)**, which will be discussed in Section III.

The **main contributions** of this work include

- Assessment and identification of risks faced by the static synchronous compensator, which constitutes an important part of a smart grid's voltage control system, due to false data injection attacks.
- A discussion of the security threat model, potential attacks, and countermeasures.
- A stochastic (Markov) security game for analysis of best defensive actions building upon the risk analysis conducted and under resource limitations.
- A numerical study illustrating the framework developed.

The rest of the paper is organized as follows. Section II discusses related work. Section III states the problem of assessing the cyberphysical security risks of voltage control by static synchronous compensation. Section IV presents our game and risk model. In Section V, we specify an informal threat model; we also discuss attack and defense actions under this threat model. In Section VI, we apply the game and risk model to voltage control, and present our simulation results. Section VII concludes this paper.

II. Related Work

Smart grid cyberphysical security is an emerging area. Substantial research effort is still being dedicated to exploring cyber attacks and their effects on power grids. Stamp et al. [10] develop a cyber-to-physical modeling approach called *Reliability Impacts from Cyber Attack*, for quantifying the degradation of system reliability for a given probability of cyber attack. Several metrics are investigated, including frequency of interruption, loss of load expectancy, load curtailed per interruption, etc. Kundur et al. [11] present two simulation studies on the effects of attacks against a single-generator system and a 13-bus system by injecting false data into a sensor in the systems. Sridhar et al. [12] propose a technique for determining the voltage control device in a grid that makes the highest impact under false data injection attacks, but they do not use any specific device model. Esfahani et al. [13] design elaborate schemes for controlling maliciously injected control signal to maximally disrupt a frequency control system.

Risk assessment has been garnering a lot of attention lately. We note that some authors erroneously refer to risk assessment as *vulnerability assessment*, which is a different concept. *Attack trees* or attack graphs is a common starting point for most work in this area. An attack tree represents attacks against a system in a tree structure, with the goal as the root node and different ways of achieving that goal as leaf nodes. Ten et al. [14] propose a framework based on attack trees for evaluating system security. They focus on attacks originating from substations connecting to the control center through a *virtual private network*. They limit cyber intrusions to firewall penetration and password cracking, singling out password policies and port auditing as the two most important security measures – these assumptions are used in other work by the same research team [15], [16]. Their framework define three vulnerability indices: the *system vulnerability index* is the maximum of *scenario vulnerability indices*, which are products of *leaf vulnerability indices*, which in turn depend on subjective definitions of port vulnerability and password strength. Liu et al. [17] take an attack tree as input, and assign a "difficulty level" to each action on the tree using Analytic Hierarchy Process. Their methodology produces a *vulnerability factor*, an artificial measure of the success probability of an attack. Analytic Hierarchy Process is a decision making methodology that is often applied to risk management, but for its reliance on subjective scoring and failure to satisfy several statistical axioms (e.g., transitivity), the risk management community is divided regarding its validity [2]. In comparison, only empirical evidence is used in this work.

The limitation of attack trees is not unrecognized. Sommestad et al. [18] propose *defense graphs* as an alternative to attack graphs, to take into account the countermeasures already in place within a system. They model defense graphs using *influence diagrams*, which are essentially Bayesian networks enhanced with indicators that express *beliefs* on *likelihood* values. The output of their assessment methodology is the expected loss associated with a successful attack. Hahn et al. [19] propose *privilege graphs* to model the privilege states in a system and the paths exploitable by an attacker. They propose an algorithm for computing an *exposure metric*, that takes into account (i) the number of attack paths through the security mechanisms protecting a target asset, and (ii) the path length representing the effort required to exploit a path.

Ten et al. [15] model attacks using *stochastic Petri Nets*, which encapsulate the probability and risk of attacks. They define the metric *system vulnerability* which is the maximum of all *scenario vulnerability* values, and the metric *impact factor* w.r.t to a substation disconnected by a successful attack. Sridhar et al. [16] use stochastic Petri Nets to model computers, firewalls and intrusion protection systems. To assess the *steady-state impact* of attacks on the power system itself, they present the impact study of six coordinated attack scenarios, where "coordination" means targeting multiple power system components at the same time. They define risk as the product of the probability of a successful attack and the resultant shed load; we adopt this definition of risk. With the exception of [16], most risk assessment work discussed so far is ICT-centric, and does not consider the impact of cyber attacks on the power system itself. In comparison, our work involves the detailed modeling and simulation of attacks on the STATCOM.

III. Voltage Control in Smart Grid

The most critical aspect of a power system is stability, and one of the most important parameters to stabilize is voltage. *Voltage instability* refers to the inability of a power system to maintain steady voltages at all buses in the system after being subjected to a disturbance from a given initial operating condition [20]. Voltage instability can lead to loss of load in one or more areas, or tripping of transmission lines and other elements by their protective systems, which in turn leads to cascading outages. The term *voltage collapse* is often used to refer to a sequence of events leading to a blackout or abnormally low voltages in a significant part of a power grid. Ultimately, the root cause of voltage instability is a system's inability to meet reactive power demand [21]. The common goal of countermeasures against voltage instability is therefore to control the production, absorption or flow of reactive power in *all segments* of the power system.

In the generation segment, the voltage level at the terminals of a generator is controlled by controlling the field excitation, through an *automatic voltage regulator*.

In the transmission and distribution segments, voltage can be controlled through a wide range of devices that either inject, absorb or redirect reactive power flow. Among the most prevalent of these devices are the *load tap changer* and *shunt capacitor*. A load tap changer is a component that changes the ratio of a transformer by adding or subtracting turns/taps from either the primary or the secondary winding (a tap provides a 1% voltage regulation). Shunt capacitors boost local voltages by injecting reactive power. Since load tap changers only redirect reactive power flow, and capacitors only inject fixed amounts of reactive power, these devices by

themselves are incapable of arresting voltage collapse. Devices that provide *active compensation* (inject/absorb reactive power in adjustable amounts) are necessary. Among devices providing active compensation, our focus here is the **static synchronous compensator (STATCOM)**. A STATCOM is a controlled reactive power source and a key member of the power electronics-based Flexible AC Transmission System family of devices, that improves voltage stability by adsorbing or generating reactive power as required. Its advantages are:

- Compared to a *synchronous condenser*, a STATCOM offers better dynamics, a lower investment cost, and lower operating and maintenance costs [22].
- Compared to a *static VAR compensator*, a STATCOM is capable of providing reactive power at low voltage, faster response and lower harmonic emission [23].

Furthermore, a STATCOM occupies less space and produces less audible noise. As power grids operate ever closer to their stability limits, the superb controllability of STATCOM becomes more valuable. Since the STATCOM's introduction in the 1990s, the number of STATCOM installations (in substations) has been steadily increasing [24]. Although our focus here is the STATCOM, it must be said that our framework is applicable to the synchronous condenser, the static VAR compensator and other types of active compensators.

In extreme circumstances, even the changing of AVR setpoints and switching of reactive devices may fail to arrest voltage collapse. **Undervoltage load shedding (UVLS)** is the ultimate countermeasure for short-term voltage instability [25], and is understandably the last resort since it leads to revenue loss. *Our aim is to model and quantify the risks posed by an attacker whose intention is to inflict revenue loss on the electricity provider by injecting false data to a STATCOM in the hope of triggering load shedding.*

IV. Security Game Model

The security game model presented is based on Alpcan and Başar's framework [6]. The concept of **risk states** is combined with this model. A system has a set of states, and a different level of risk is associated with each state. In this work, we define risk as *the product of the probability of a successful attack and the resultant shed load (in the unit of power)*. Clearly under this definition, risk ranges from 0 to the maximum sheddable load. As a starting point, we partition this risk space into only two states: s_0 where risk is zero (no load is shed), and s_1 where risk is nonzero (some load is shed). We model the state to evolve probabilistically according to a stochastic process with the Markov property. Accordingly, we model the interactions between an attacker and a defender using stochastic or Markov *security games*.

As a general basis for Markov security games, consider a 2-player (attacker vs. defender) zero-sum Markov game played on a finite state space, where each player has a finite number of actions to choose from. Let the attacker's action space be $\mathcal{A}^A \stackrel{\text{def}}{=} \{a_1, \ldots, a_{N_A}\}$, the defender's action space be $\mathcal{A}^D \stackrel{\text{def}}{=} \{d_1, \ldots, d_{N_D}\}$, and the state space be $\mathcal{S} \stackrel{\text{def}}{=} \{s_1, \ldots, s_{N_S}\}$. It is assumed that the state evolves according to a discrete-time finite-state Markov chain which enables utilization of well-established analytical tools to study the problem. Let $\boldsymbol{p}^{\mathcal{S}}(t)$ be the probability distribution on the state space $\mathcal{S}$, i.e.,

$$\boldsymbol{p}^{\mathcal{S}}(t) \stackrel{\text{def}}{=} \begin{bmatrix} \Pr[s(t) = s_1] & \cdots & \Pr[s(t) = s_{N_S}] \end{bmatrix}^T,$$

where $t \geq 1$ denotes the discrete time (stage) of the repeated Markov game. Let $\boldsymbol{M}(a, d) = [M_{s_i,s_j}(a, d)]_{N_S \times N_S}$ be the state transition matrix which is parameterized by $a \in \mathcal{A}^A$ and $d \in \mathcal{A}^D$, such that

$$\boldsymbol{p}^S(t+1) = \boldsymbol{M}(a, d)\boldsymbol{p}^S(t). \tag{1}$$

The matrix entry $M_{s_i,s_j}(a, d)$ represents the probability of state s_i transitioning to state s_j under attacker action a and defender action d.

In each state $s \in \mathcal{S}$, the attacker and defender play a zero-sum game represented by matrix $\boldsymbol{G}(s)$. More precisely, given a state $s(t) \in \mathcal{S}$ at a stage t, the players play the zero-sum game $\boldsymbol{G}(s(t)) = [G_{a,d}(s(t))]_{N_A \times N_D}$. The matrix entry $G_{a,d}(s)$ represents the attacker's gain from risk state s by taking action a when the defender action is d. As a simplifying assumption, actions have no cost other than their "contribution" to load shedding, so $G_{a,d}(s)$ is the *expected total load shed* in state s under attacker action a and defender action d. Due to the adopted zero-sum Markov game formulation, the attacker's gain (loss) equals the defender's loss (gain).

The attacker's (mixed) strategy is defined as a probability distribution on $\mathcal{A}^A$ for a give state s, i.e., $p^A(s) \stackrel{\text{def}}{=} \begin{bmatrix} \Pr[a(s) = a_1] & \cdots & \Pr[a(s) = a_{N_A}] \end{bmatrix}^T$. The defender's strategy is similarly defined. For the zero-sum Markov game formulation here, the defender aims to minimize its own expected total cost, Q, in response to the attacker who tries to maximize it. The reverse is true for the attacker due to the zero-sum nature of the game. Hence, it is sufficient to describe the solution algorithm for only one player, which is the defender in this case.

The game is played in stages over an infinite horizon. Using the future-discounted cost model, the defender's Q at the end of a game is the sum of all realized stage costs discounted by a scalar *discount factor* $\gamma \in [0, 1)$:

$$Q \stackrel{\text{def}}{=} \sum_{t=0}^{\infty} \gamma^t G_{a(t),d(t)}(s(t)), \tag{2}$$

where $a(t) \in \mathcal{A}^A$, $d(t) \in \mathcal{A}^D$, $s(t) \in \mathcal{S}$, $G_{a(t),d(t)}(s(t))$ is the $(a(t), d(t))$-th element of the stage-t game matrix $\boldsymbol{G}(s(t))$. The defender can theoretically choose a different strategy $p^D(s(t))$ at each stage t of the game to minimize Q in (2). Fortunately, this complex problem can be simplified significantly. First, it can be shown that a stationary strategy $p^D(s) = p^D(s(t)), \forall t$ is optimal, and hence there is no need to compute a separate optimal strategy for each stage. Second, the problem can be solved recursively using *dynamic programming* to obtain the stationary optimal strategy (solving a zero-sum matrix game at each stage). The optimal strategy can be mixed, i.e., stochastic for each state s. At stage $t + 1$,

the optimal cost $Q_{t+1}(s,a,d)$ (the dependency of s, a and d on t is omitted for notational brevity) can be expressed with the *Bellman equations*:

$$Q_{t+1}(s,a,d) = G_{a,d}(s) + \gamma \sum_{s' \in \mathcal{S}} M_{s,s'}(a,d) \cdot V(s'), \quad (3)$$

$$V(s') = \min_{p^D(s')} \max_{a} \sum_{d \in \mathcal{A}^D} Q_t(s',a,d) p_d^D(s'), \quad (4)$$

for $t = 0, 1, \ldots$ and a given initial condition Q_0. In (4), $p_d^D(s')$ is the element of $p^D(s')$ that corresponds to d. Changing the equal signs in (3) and (4) to assignment operators "$\leftarrow$" gives us the *value iteration* algorithm, which converges to the optimal Q^* as $t \to \infty$ [26]. To take advantage of increasingly more efficient linear program solvers, it is common to formulate (4) as a linear program:

$$\min_{p^D(s)} V(s) \quad (5)$$
$$\text{s.t. } V(s) \geq \sum_{d \in \mathcal{A}^D} Q_t(s,a,d) p_d^D(s), \forall a \in \mathcal{A}^A,$$
$$p_d^D(s) \geq 0, \sum_d p_d^D(s) = 1, \forall d \in \mathcal{A}^D.$$

The strategy $p^D(s), \forall s \in \mathcal{S}$ computed from (5) is the *minimax* strategy w.r.t. Q. The fixed points of (4) and (3), V^* and Q^*, then lead to the optimal minimax solution for the defender. For this work, we use the value iteration algorithm in the form of (5) and (3).

V. Threat Analysis

Fig. 1 shows the communication architecture of a substation based on the international standard IEC 61850. The Merging Units are responsible for combining current and voltage measurements, and transmitting them through the IEC 61850-9-2 Process Bus to all subscribing devices, including STATCOMs. Access to the substation's control system is typically enabled through a virtual private network (VPN) [27]. Some authors [14] equate the compromise of an entire substation to the successful cracking of a VPN access password and the penetration of an Internet-facing firewall (see Fig. 1). This strong attacker model is not entirely unrealistic, however, our goal is to investigate the strategy of an attacker that has successfully penetrated the protected network but whose actions within the control system are bounded by several resource constraints. We assume the following resource constraints:

- The attacker cannot tamper with the EMS.
- The attacker cannot directly trip generators and transmission lines (by opening circuit breakers).
- The attacker cannot tamper with turbine governors.
- The attacker cannot tamper with automatic voltage regulators (AVRs). Since a generator and its associated AVR are usually installed side-by-side, voltage measurements at the generator's terminals are directly wired to the AVR.
- The attacker cannot tamper with overvoltage protection and undervoltage load shedding relays.
- The attacker cannot tamper with underfrequency load shedding relays. Some commercial relays (e.g., SEL-387E) have an integrated frequency meter, and are therefore not subject to false frequency data injection attacks.
- The attacker cannot tamper with STATCOMs.

Without the above constraints, it is a trivial exercise for any attacker that has successfully penetrated the protected network to trigger cascading failures across the power grid. It is therefore conceivable that an energy provider would make protecting its EMS, generators, circuit breakers, turbine governors, AVRs, overvoltage protection relays, undervoltage load shedding relays, underfrequency load shedding relays, STATCOMs its foremost priority. Despite the above constraints, an attacker can forge and send false voltage data through a compromised Merging Unit to a STATCOM (see Fig. 1). In the spirit of stealthy attacks as embodied by Stuxnet, Duqu and Flame, it is also conceivable that a persistent attacker would adopt this subtle and stealthy strategy. It is up to the STATCOM software to detect this attack.

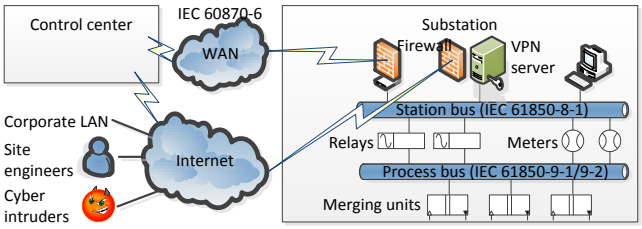


Fig. 1. Accessibility of a substation's control system from the Internet. In our threat model, an attacker can feed false data to a STATCOM through a compromised Merging Unit connected to the process bus.

Basic attacks: The purpose of a false data injection attack on a STATCOM is presumably to cause undervoltage or overvoltage to trigger shedding of loads or tripping of generators. It is impossible to exhaust all injection patterns, but a basic attack pattern is to change voltage measurements from v to $\boxed{kv+b}$, where v is the true voltage, k and b are multiplicative and additive constants determined by the attacker. The effect of this basic injection pattern varies with k's value:

- Case $k < 0$: Misled by the false voltage values which are 180 degrees out of phase with the true values, the STATCOM adsorbs reactive power when it should generate, and generate when it should adsorb. Consequences include pronounced high-order oscillations, in some cases load shedding, and in the worst case voltage collapse.
- Case $k = 0$: The STATCOM receives a constant input b, and does nothing (to regulate a supposedly constant voltage). This attack in effect disables the STATCOM.
- Case $0 < k < 1$: The STATCOM gets partial regulatory capability. Other values of k cause more severe attacks.
- Case $1 < k$: This attack causes the STATCOM to adsorb or generate more reactive power than necessary. Consequences include oscillations, in some cases load shedding, and in the worst case voltage collapse.

Simulation results showing the impact of the attacks in (6) are given in the next section. For a security game where the attacker substitutes every true voltage value v with kv, the attacker action space can be defined in terms of values of k:

$$\mathcal{A}^A \stackrel{\text{def}}{=} \{k_1, k_2, \ldots, k_{N_A}\}, \tag{6}$$

where $k_1, k_2, \ldots, k_{N_A}$ are constants.

Basic defenses: Both general and STATCOM-specific countermeasures are applicable:

- Redundancy: Measurement redundancy is routinely provisioned for critical voltage measurements [28]. Multiple Merging Units of different grades can be installed, so that the likelihood of all units being compromised is small and the STATCOM has a non-zero chance of getting genuine voltage measurements.
- Saturation filter: We can constrain instantaneous voltage input to a STATCOM to a certain value range, e.g., $[-1.2, 1.2]$ p.u. (i.e., passing the input through a saturation filter), since most systems do not tolerate a rise of more than 20% [28], [29]. This mitigates the effect of an attack that uses a large k.
- Detection: Measurement redundancy and saturation filtering only limit the effect of an attack, stopping an attack requires the attack to be detected and the source be removed. A STATCOM's internal variables can be analyzed for signs of intrusions. For example, a threshold-based algorithm can be constructed to count the number of times a control variable crosses zero; a count exceeding a predefined threshold within a predefined time span indicates anomalous operation. One such control variable is $I_q - I_{qref}$, where I_q is the current flowing through the current regulator of a STATCOM controller, and I_{qref} is the reference current of the current regulator [30, Chapter 5]. In normal circumstances, $I_q - I_{qref}$ stabilizes around zero, but in anomalous circumstances, $I_q - I_{qref}$ takes longer to converge to zero if at all.

For a security game where the defender monitors the number of zero crossings of $I_q - I_{qref}$ within a fixed time frame, the defender action space can be defined in terms of the zero crossing thresholds:

$$\mathcal{A}^D \stackrel{\text{def}}{=} \{\tau_1, \tau_2, \ldots, \tau_{N_D}\}, \tag{7}$$

where $\tau_1, \ldots, \tau_{N_D}$ are constants.

There are unlimited ways to improve upon the basic attacks to defeat the basic defenses. Correspondingly, there are unlimited ways to detect these improved attacks with varying accuracy, and it is plausible that there are more advanced STATCOM controllers that are less susceptible to these attacks. Nevertheless, our interest is not on the design of attacks, defenses or the controller, but on the modeling of system risk dynamics under the actions of the attacker and defender for any given system.

VI. Voltage Control Games

Our simulation study consists of two parts:

1) Section VI-A assesses the impact and characteristics of basic attacks, the results of which allow us to determine suitable values for the constants in the attacker action space and defender action space (see (6) and (7));
2) Section VI-B simulates attacks and defenses on a test system, the results of which allow us to compute the optimal attack and defense strategies through the value iteration algorithm (see (5) and (3)).

For simulations, we use the 1-machine 4-bus distribution system in Fig. 2. The loads, connected to bus 2 (B2) and bus 4 (B4), are each connected to a UVLS relay. The UVLS relays implement the following rules adapted from Hydro-Québec's [28] ($V_{\text{B}i} \stackrel{\text{def}}{=}$ voltage magnitude at bus i):

- If voltage $V_{\text{B4}} < 0.94$ p.u. for t_1 s, shed L1 (2 MW).
- If voltage $V_{\text{B4}} < 0.92$ p.u. for t_2 s, shed L2 (2 MW).
- If voltage $V_{\text{B4}} < 0.90$ p.u. for t_3 s, shed L3 (3.5 MW).
- If voltage $V_{\text{B2}} < 0.90$ p.u. for t_3 s, shed L0 (3 MW + 0.2 MVAr).

Above, "p.u." stands for "per unit" and is simply the ratio of an absolute value in some unit to a base/reference value in the same unit; for example, at bus B4, 1 p.u. $\equiv$ 600 V. The values t_1, t_2 and t_3 are set to 0.4 s, 0.3 s and 0.2 s respectively, which for simulation efficiency are shorter than in practice. We note that the following simulations take several actual minutes to simulate 1 virtual second on an Intel Core 2 Duo processor.

A. Characteristics of basic attacks

The perfect cover for a stealthy attack is a disturbance. As such, we program the test system to experience a disturbance every 0.5 s, such that the STATCOM not only has to cope with the disturbances but also attacks. We simulate disturbances at a high occurrence frequency for computational efficiency and illustrative purposes, but frequent disturbances can be viewed as sporadic disturbances in a compressed timeline.

The disturbances originate in the generator and take the form of an undervoltage (overvoltage) peak followed by an overvoltage (undervoltage) peak. The peaks are separated from each other by up to 0.1 s, and each peak has a magnitude uniformly distributed between 0.9 p.u. and 1.1 p.u.

Fig. 3 plots the voltage magnitude at Bus 4 under attacks with $k \in \{-1.2, -0.8, 1.1, 1.2\}$ (these values are chosen to show that even a small multiplier has a sizeable impact, and because large values of k are easier to detect). Fig. 4 plots the same variables but with load L_3 changed from a 3.5 MW load to a variable load whose apparent power varies between 1 MVA and 5.2 MVA at 5 Hz. Except for the case $k = -1.2$, all attacks trigger load shedding.

Both Fig. 3 and Fig. 4 validate the potentiality of using the control variable $I_q - I_{qref}$ of a STATCOM controller for intrusion detection. As can be seen, I_q has trouble matching I_{qref} under the influence of attacks. In other words, a zero-crossing count of $I_q - I_{qref}$ that is above a certain threshold provides an indication of possible attacks. However, a variable load narrows the range of valid thresholds, and makes this detection method more prone to false positives.

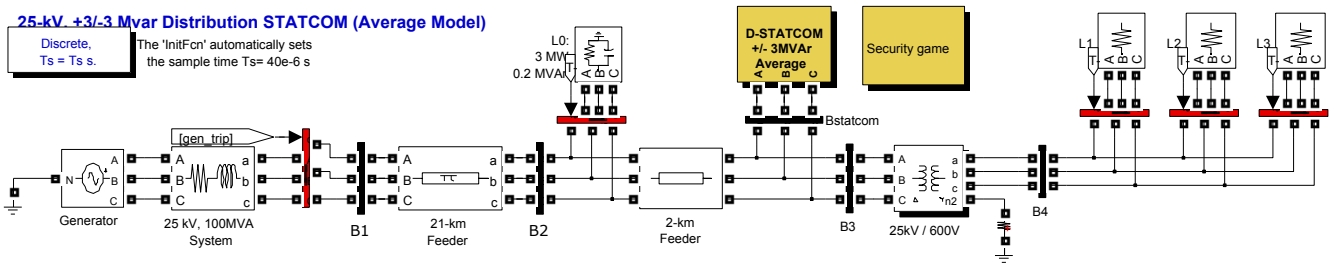


Fig. 2. The Simulink representation and simulation parameters for a 1-machine 4-bus distribution system adapted from the SimPowerSystems sample "power_dstatcom_avg". A Distribution STATCOM is connected to B3 for voltage regulation.

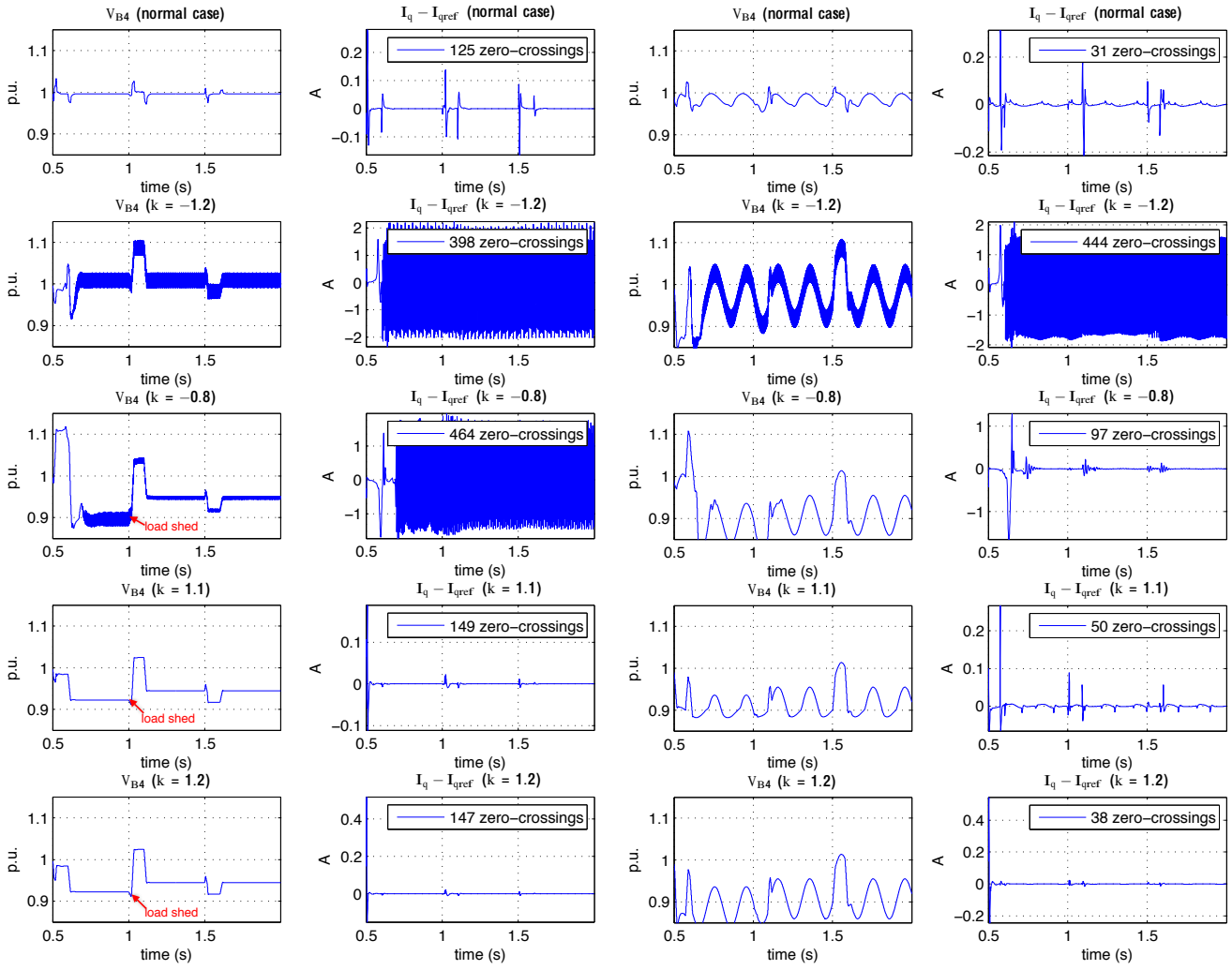


Fig. 3. Plots of the voltage magnitude at Bus 4 (left column) and the internal variable $I_q - I_{qref}$ of the STATCOM (right column), when L_3 in Fig. 2 is fixed load of 3.5 MW. All attacks except for $k = -1.2$ trigger load shedding.

Fig. 4. Plots of the voltage magnitude at Bus 4 (left column) and the internal variable $I_q - I_{qref}$ of the STATCOM (right column), when L_3 in Fig. 2 is a variable load (1∼5.2 MVA at 5 Hz).

B. Security games

The 4-bus system in Fig. 2, where a STATCOM operates defensively against an attacker, is simulated in order to observe the state transition matrix $\boldsymbol{M}(a,d) = [M_{s_i,s_j}(a,d)]_{N_S \times N_S}$, and the game matrix $\boldsymbol{G}(s) = [G_{a,d}(s(t))]_{N_A \times N_D}$. $M_{s_i,s_j}(a,d)$ is readily obtained by fixing attacker action at a, defender action at d, and measuring the probability of a session starting in state s_i ends in state s_j. Based on our assumption that actions have no cost other than their "contribution" to load shedding, $\boldsymbol{G}(s_0) = \mathbf{0}$; $\boldsymbol{G}(s_1)$ is the expected total load

shed in state s_1. To obtain $G_{a,d}(s_1)$, we fix attacker action at a, defender action at d, measure the total energy shed throughout the game E_{s_1}, measure the combined duration of load shedding T_{s_1}, and then compute $G_{a,d}(s_1) = E_{s_1}/T_{s_1}$.

There are two meters in the system, with Meter 1 being online and compromised from the start, and Meter 2 being offline and intact. The STATCOM reads from Meter 1 and executes a detection algorithm every 0.5 s until it detects Meter 1's compromised state, at which point

1) it will attempt to bring Meter 2 online, and switch to Meter 2;
2) if Meter 2 is in the middle of a disinfection, it will wait until the disinfection is completed;
3) upon successful switching to Meter 2, it will take Meter 1 offline and disinfect Meter 1.

Disinfection, for example in the form of refreshing the firmware including cryptographic keys in the Flash memory, takes $T_{\text{disinfect}}$ seconds. At the end of a disinfection, any shed load is reconnected. Once brought online, a meter becomes compromised after T_{infect} seconds. The game proceeds as such with the STATCOM switching between Meter 1 and Meter 2 as necessary. We simulate two scenarios depending on (i) the nature of the load L_3 in Fig. 2, (ii) the value of $T_{\text{disinfect}}$, (iii) the value of T_{infect}, (iv) attacker actions, (v) defender actions, and (vi) whether the cost of false positives is accounted for. Using MATLAB/Simulink, each scenario is simulated for 25.5 virtual seconds (which take more than an actual hour). The obtained $\boldsymbol{M}$ and $\boldsymbol{G}$ are fed into the value iteration algorithm.

Scenario 1: L_3 is a fixed load of 3 MW. $T_{\text{disinfect}} = 1$. $T_{\text{infect}} = 0.01$. Fig. 3 shows that attacks with $k = 1.1$ and $k = 1.2$ are not only effective in causing load shedding but also more stealthy, so we define the attacker action space as

$$\mathcal{A}^A \stackrel{\text{def}}{=} \{a_1, a_2\} = \{1.1, 1.2\}. \tag{8}$$

Fig. 3 also shows the normal case has 125/3 zero crossings per 0.5 s, and the $k = 1.1$ case has 147/3 zero crossings per 0.5 s, so we define the defender action space as

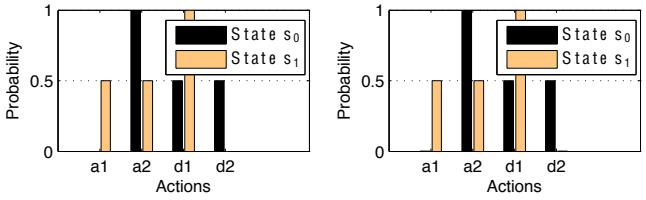
$$\mathcal{A}^D \stackrel{\text{def}}{=} \{d_1, d_2\} = \{\lceil 125/3 \rceil, \lfloor 147/3 \rfloor\} = \{42, 49\}. \tag{9}$$

The cost of false positives is ignored by setting $\boldsymbol{G}(s_0) = \boldsymbol{0}$. Note that near-instantaneous infection ($T_{\text{infect}} = 0.01$) rules out the possibility of false positives anyway.

The results in Fig. 5 suggest, for the attacker, adopting a pure strategy in state s_0, but a mixed strategy in state s_1; for the defender, adopting a mixed strategy in state s_0, but a pure strategy in state s_1. The results show a_2 is a stronger attack, and d_1 is a stronger defense. While these results confirm intuition, in practice, unnecessarily disinfecting a meter due to a false positive incurs cost in terms of decreased meter lifespan, and introduces the risk of switching failures. Therefore, in Scenario 2, we take the cost of false positives into account.

Scenario 2: L_3 is a variable load whose apparent power varies between 1 MVA and 5 MVA at 5 Hz. $T_{\text{disinfect}} = 0.5$. $T_{\text{infect}} = 1$. Fig. 4 shows that the case $k = 1.1$ has the highest

$$\boldsymbol{M}(a_1, d_1) = \begin{bmatrix} 6/14 & 8/14 \\ 7/35 & 28/35 \end{bmatrix} \quad \boldsymbol{M}(a_1, d_2) = \begin{bmatrix} 4/9 & 5/9 \\ 4/40 & 36/40 \end{bmatrix}$$
$$\boldsymbol{M}(a_2, d_1) = \begin{bmatrix} 6/14 & 8/14 \\ 7/35 & 28/35 \end{bmatrix} \quad \boldsymbol{M}(a_2, d_2) = \begin{bmatrix} 6/14 & 8/14 \\ 7/35 & 28/35 \end{bmatrix}$$
$$\boldsymbol{G}(s_0) = \boldsymbol{0} \qquad \boldsymbol{G}(s_1) = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$



(a) $\gamma = 0.1$ (b) $\gamma = 0.9$

Fig. 5. Optimal attack and defense strategies for Scenario 1.

tendency to cause false positives, whereas $k = -0.8$ is able to cause load shedding and yet not likely to cause false positives, so we define the attacker action space as

$$\mathcal{A}^A \stackrel{\text{def}}{=} \{a_1, a_2\} = \{1.1, -0.8\}. \tag{10}$$

Fig. 4 also shows the normal case has 31/3 zero crossings per 0.5 s, and the $k = -0.8$ case has 97/3 zero crossings per 0.5 s, so we define the defender action space as

$$\mathcal{A}^D \stackrel{\text{def}}{=} \{d_1, d_2\} = \{\lceil 31/3 \rceil, \lfloor 97/3 \rfloor\} = \{11, 32\}. \tag{11}$$

The cost of false positives is accounted for by setting $G_{a,d}(s_0) = c_{\text{fp}} p_{\text{fp}}(a, d)$, where c_{fp} is the cost of false positive in the same unit as load shed, and $p_{\text{fp}}(a, d)$ is the false positive probability due to attacker action a and defender action d. Note that false positives could occur due to a large T_{infect}.

The results in Fig. 6(a-b) suggest that when c_{fp} is two orders of magnitude smaller than cost of load shed (i.e., negligible), a_1 is the optimal attack and d_1 is the optimal defense. Fig. 6(c-d) shows that when c_{fp} has the same order of magnitude as the cost of load shed, the cost of false positive makes a_2 and d_2 more appealing to the attacker and defender respectively.

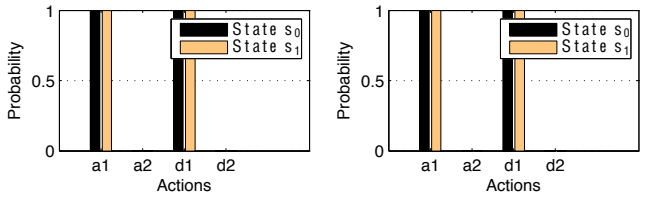
VII. Conclusion and Future Work

This work investigates an important class of false data injection attacks on a critical voltage control component called static synchronous compensator (STATCOM), under an explicit security threat model. We assess the characteristics of these attacks and propose several countermeasures including a detection algorithm. The interaction between an attacker and a grid defense system is modeled as a security game and simulated as part of a power system, where the attacker chooses the intensity of false data injection and the defender determines the detection threshold level. The simulation results are used as an input to a stochastic game model, where the decisions on defensive measures are made taking into account resource constraints represented by cost values. Thus, security games provide a framework for choosing the best response strategies against attackers in order to minimize potential risks.

$$\boldsymbol{M}(a_1, d_1) = \begin{bmatrix} 43/45 & 2/45 \\ 2/4 & 2/4 \end{bmatrix} \qquad \boldsymbol{M}(a_1, d_2) = \begin{bmatrix} 0/2 & 2/2 \\ 1/47 & 46/47 \end{bmatrix}$$

$$\boldsymbol{M}(a_2, d_1) = \begin{bmatrix} 48/49 & 1/49 \\ 0 & 0 \end{bmatrix} \qquad \boldsymbol{M}(a_2, d_2) = \begin{bmatrix} 25/32 & 7/32 \\ 7/17 & 10/17 \end{bmatrix}$$

$$\boldsymbol{G}(s_0) = c_{\text{fp}} \cdot \begin{bmatrix} \frac{44}{46} & 0 \\ \frac{42}{49} & \frac{24}{33} \end{bmatrix} \qquad \boldsymbol{G}(s_1) = \begin{bmatrix} 2 & 2.50 \\ 2 & 2.15 \end{bmatrix}$$



(a) $c_{\text{fp}} = 0.01$, $\gamma = 0.1$ (b) $c_{\text{fp}} = 0.01$, $\gamma = 0.9$

(c) $c_{\text{fp}} = 4$, $\gamma = 0.1$ (d) $c_{\text{fp}} = 4$, $\gamma = 0.9$

Fig. 6. Optimal attack and defense strategies for Scenario 2. In (d), $p^A(s_1) = \{0.9975, 0.0025\}$ and $p^D(s_1) = \{0.0046, 0.9954\}$ are mixed strategies.

For numerical simplicity, we define only two attacker actions and two defender actions, when in fact our framework is applicable to any number of attacker and defender actions.

For our preliminary study, we have adopted a risk-neutral framework, such that the expected loss from a blackout tends to conceal the significance of rare events at the tail-end of a probability distribution. Financial risk measures (e.g., conditional value-at-risk) have been proposed to account for these rare events [31], and are being explored in ongoing work.

REFERENCES

[1] "IEEE Guide for Smart Grid Interoperability of Energy Technology and Information Technology Operation with the Electric Power System (EPS), End-Use Applications, and Loads," *IEEE Std 2030-2011*, pp. 1–126, Sep. 2011.

[2] D. W. Hubbard, *The Failure of Risk Management: Why It's Broken and How to Fix It*. Wiley, 2009.

[3] NIST, "Glossary of key information security terms," IR 7298 Revision 1, Feb. 2011.

[4] Australian Government, "Critical infrastructure resilience strategy," ISBN 978-1-921725-25-8, http://www.tisn.gov.au/, 2010.

[5] M. Leitch, "ISO 31000:2009—The New International Standard on Risk Management," *Risk Analysis*, vol. 30, no. 6, pp. 887–892, 2010.

[6] T. Alpcan and T. Başar, *Network Security: A Decision and Game Theoretic Approach*. Cambridge University Press, 2011.

[7] J. Mounzer, T. Alpcan, and N. Bambos, "Dynamic Control and Mitigation of Interdependent IT Security Risks," in *2010 IEEE International Conference on Communications (ICC)*, May 2010, pp. 1–6.

[8] P. Bommannavar, T. Alpcan, and N. Bambos, "Security risk management via dynamic games with learning," in *2011 IEEE International Conference on Communications (ICC)*, Jun. 2011, pp. 1–6.

[9] Y. W. Law, T. Alpcan, M. Palaniswami, and S. Dey, "Security games and risk minimization for automatic generation control in smart grid," in *Proc. 3rd Conference on Decision and Game Theory for Security (GameSec 2012)*, ser. LNCS, J. Grossklags and J. Walrand, Eds., vol. 7638. Springer Heidelberg, 2012, pp. 281–295.

[10] J. Stamp, A. McIntyre, and B. Ricardson, "Reliability impacts from cyber attack on electric power systems," in *IEEE/PES Power Systems Conference and Exposition (PSCE '09)*, Mar. 2009, pp. 1–8.

[11] D. Kundur, X. Feng, S. Mashayekh, S. Liu, T. Zourntos, and K. L. Butler-Purry, "Towards modelling the impact of cyber attacks on a smart grid," *International Journal of Security and Networks*, vol. 6, no. 1/2011, pp. 2–13, 2011.

[12] S. Sridhar and G. Manimaran, "Data integrity attack and its impacts on voltage control loop in power grid," in *Power and Energy Society General Meeting, 2011 IEEE*, Jul. 2011, pp. 1–6.

[13] P. M. Esfahani, M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "A Robust Policy for Automatic Generation Control Cyber Attack in Two Area Power Network," in *IEEE Conference on Decision and Control*, Dec. 2010.

[14] C.-W. Ten, G. Manimaran, and C.-C. Liu, "Cybersecurity for critical infrastructures: Attack and defense modeling," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 40, no. 4, pp. 853–865, Jul. 2010.

[15] C.-W. Ten, C.-C. Liu, and G. Manimaran, "Vulnerability Assessment of Cybersecurity for SCADA Systems," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1836–1846, Nov. 2008.

[16] S. Sridhar, M. Govindarasu, and C.-C. Liu, "Risk analysis of coordinated cyber attacks on power grid," in *Control and Optimization Methods for Electric Smart Grids*, ser. Power Electronics and Power Systems. Springer US, 2012, vol. 3, pp. 275–294.

[17] N. Liu, J. Zhang, H. Zhang, and W. Liu, "Security Assessment for Communication Networks of Power Control Systems Using Attack Graph and MCDM," *IEEE Trans. Power Del.*, vol. 25, no. 3, pp. 1492–1500, Jul. 2010.

[18] T. Sommestad, M. Ekstedt, and L. Nordstrom, "Modeling security of power communication systems using defense graphs and influence diagrams," *IEEE Trans. Power Del.*, vol. 24, no. 4, pp. 1801–1808, Oct. 2009.

[19] A. Hahn and M. Govindarasu, "Cyber attack exposure evaluation framework for the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 835–843, Dec. 2011.

[20] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, and V. Vittal, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004.

[21] P. Kundur, *Power System Stability and Control*. McGraw-Hill Professional, 1994.

[22] X.-P. Zhang, C. Rehtanz, and B. Pal, *Flexible AC Transmission Systems: Modelling and Control*. Springer, 2012.

[23] P. Giroux, G. Sybille, and H. Le-Huy, "Modeling and simulation of a distribution STATCOM using Simulink's Power System Blockset," in *The 27th Annual Conference of the IEEE Industrial Electronics Society (IECON '01)*, vol. 2, 2001, pp. 990–994.

[24] J. Paserba, "Secondary voltage-VAr controls applied to static compensators (STATCOMs) for fast voltage control and long term VAr management," in *2002 IEEE Power Engineering Society Summer Meeting*, vol. 2, Jul. 2002, pp. 753–761.

[25] T. van Cutsem and C. Vournas, *Voltage Stability of Electric Power Systems*. Springer, 2007.

[26] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2009.

[27] J. Wiles, T. Claypoole, P. A. Henry, P. Drake, and S. Lowther, *Techno Security's Guide to Securing SCADA: A Comprehensive Handbook On Protecting The Critical Infrastructure*. Syngress, 2008.

[28] D. Lefebvre, S. Bernard, and T. Van Cutsem, "Undervoltage load shedding scheme for the Hydro-Québec system," in *IEEE Power Engineering Society General Meeting*, vol. 2, Jun. 2004, pp. 1619–1624.

[29] V. V. Thong, D. Van Dommelen, and R. Belmans, "Penetration level of distributed energy resources with anti-islanding criteria and secure operation of power system," in *IEEE Power Engineering Society General Meeting*, 2006.

[30] H. Québec, *SimPowerSystems™ User's Guide R2012a*, version 5.6 ed., Hydro-Québec and The MathWorks, Inc., The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098, Mar. 2012.

[31] P. Varaiya, F. Wu, and J. Bialek, "Smart operation of smart grid: Risk-limiting dispatch," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 40–57, Jan. 2011.