INVITED PAPER

Cyber–Physical Security of a Smart Grid Infrastructure

By Yilin Mo, Tiffany Hyun-Jin Kim, Kenneth Brancik, Dona Dickinson, Heejo Lee, Adrian Perrig, and Bruno Sinopoli

ABSTRACT | It is often appealing to assume that existing solutions can be directly applied to emerging engineering domains. Unfortunately, careful investigation of the unique challenges presented by new domains exposes its idiosyncrasies, thus often requiring new approaches and solutions. In this paper, we argue that the "smart" grid, replacing its incredibly successful and reliable predecessor, poses a series of new security challenges, among others, that require novel approaches to the field of cyber security. We will call this new field cyber–physical security. The tight coupling between information and communication technologies and physical systems introduces new security concerns, requiring a rethinking of the commonly used objectives and methods. Existing security approaches are either inapplicable, not viable, insufficiently scalable, incompatible, or simply inadequate to address the challenges posed by highly complex environments such as the smart grid. A concerted effort by the entire industry, the research community, and the policy makers is required to achieve the vision of a secure smart grid infrastructure.

KEYWORDS | Cyber–physical systems; security; smart grids

I. INTRODUCTION

The electric grid is arguably the world's largest engineered system. Vital to human life, its reliability is a major and often understated accomplishment of humankind. It is the motor of the economy and the major driver of progress. In its current state, the grid consists of four major components: 1) *generation* produces electric energy in different manners, e.g., by burning fossil fuels, inducing nuclear reaction, harnessing water (hydro-electric dams), wind, solar, and tidal forces; 2) *transmission* moves electricity via a very high voltage infrastructure; 3) *distribution* steps down current and spreads out for consumption; and 4) *consumption*, i.e., industrial, commercial, and residential, uses the electric energy in a multitude of ways.

Given the wide variety of systems, their numerous owners, and a diverse range of regulators, a number of weaknesses have emerged. Outages are often recognized only after consumers report. Matching generation to demand is challenging because utilities do not have clear cut methods to predict demand and to request demand reduction (load shedding). As a consequence, they need to overgenerate power for peak demand—which is expensive and contributes to Green-house Gas (GhG) emissions. For similar reasons it is difficult to incorporate variable generation, such as wind and solar power, into the grid. Last, there is a dearth of information available for consumers to determine how and when to use energy.

To address these challenges, the smart grid concept has evolved. The smart grid uses communications and information technologies to provide better "situational awareness" to utilities regarding the state of the grid. Smart grid provides numerous benefits [1]–[4]. Using intelligent communications, load shedding can be implemented so that peak demand can be flattened, which reduces the need to bring additional (expensive) generation plants online. Using information systems to perform predictive analysis, including when wind and solar resources will produce less power, the utilities can keep power appropriately balanced. As new storage technologies emerge at

Manuscript received April 17, 2011; revised June 22, 2011; accepted June 23, 2011.
Y. Mo, **T. H.-J. Kim**, **A. Perrig**, and **B. Sinopoli** are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: ymo@ece.cmu.edu; hyunjin1@ece.cmu.edu; adrian@ece.cmu.edu; brunos@ece.cmu.edu).
K. Brancik and **D. Dickinson** are with Northrop Grumman Corporation, McLean, VA 22102 USA (e-mail: kenneth.brancik@ngc.com; dona.dickinson@ngc.com).
H. Lee was with the CyLab, Carnegie Mellon University, Pittsburgh, PA 15213 USA. He is now with Division of Computer and Communication Engineering, Korea University, Seoul 136-701, Korea (e-mail: heejo@korea.ac.kr).
Digital Object Identifier: 10.1109/JPROC.2011.2161428

the utility scale, incorporation of these devices will likewise benefit from intelligent demand prediction. Last, the ability for consumers to receive and respond to price signals will help them manage their energy costs, while helping utilities avoid building additional generation plants.

With all these approaches, the smart grid enables a drastic cost reduction for both power generation and consumption.

Dynamic pricing and distributed generation with local generators can significantly reduce the electricity bill. Fig. 1(a) shows how to use electricity during off-peak periods when the price is low. Conversely, Fig. 1(b) shows load shedding during peak times and utilization of energy storage to meet customer demand. The effect of peak demand reduction by "demand management" is shown in Fig. 2. Pilot projects in the states of California and Washington [1] indicate that scheduling appliances based on price information can reduce electricity costs by 10% for consumers. More advanced smart grid technologies promise to provide even larger savings.

To establish the smart grid vision, widespread sensing and communications between all grid components (generation, transmission, distribution, storage) and consumers must be created and managed by information technology systems. Furthermore, sophisticated estimation, control, and pricing algorithms need to be implemented to support the increasing functionality of the grid while maintaining reliable operations. It is the greatly increased incorporation of IT systems that supports the vision, but unfortunately also creates exploitable vulnerabilities for the grid and its users.

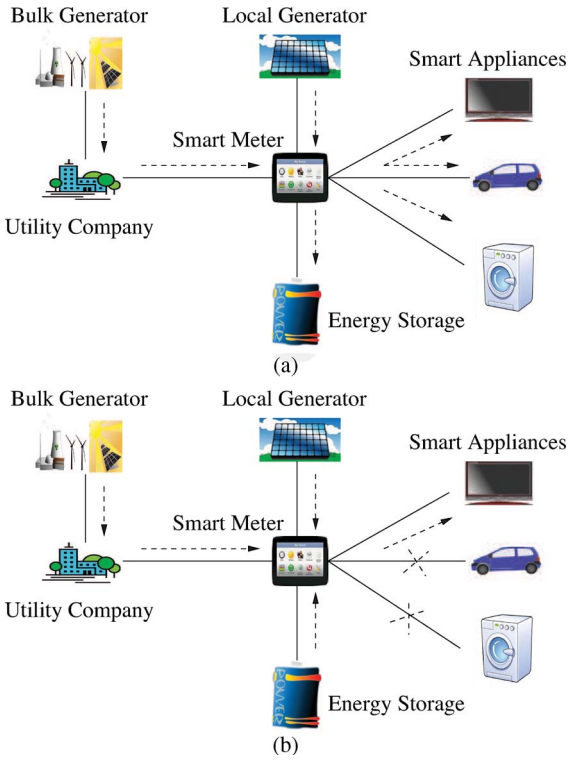


Fig. 1. *During off-peak time periods, inexpensive electric power can be used without restrictions (e.g., diverted to energy storage). During peak time periods, some appliances will be temporarily turned off, and stored energy is used. (a) Power usage during off-peak time period. (b) Power usage during peak time period.*

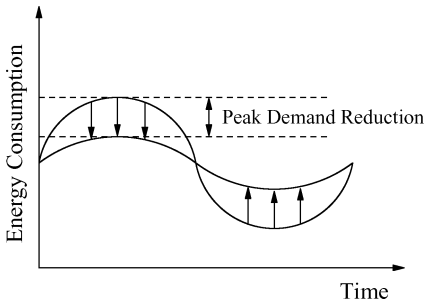


Fig. 2. *The peak demand for electricity will be reduced by the use of smart appliances, local generators, and/or local energy storage.*

A. A Cyber–Physical Approach to Smart Grid Security

A wide variety of motivations exist for launching an attack on the power grid, ranging from economic reasons (e.g., reducing electricity bills), to pranks, and all the way to terrorism (e.g., threatening people by controlling electricity and other life-critical resources). The emerging smart grid, while benefiting the benign participants (consumers, utility companies), also provides powerful tools for adversaries.

The smart grid will reach every house and building, giving potential attackers easy access to some of the grid components. While incorporating information technology (IT) systems and networks, the smart grid will be exposed to a wide range of security threats [5]. Its large scale also makes it nearly impossible to guarantee security for every single subsystem. Furthermore, the smart grid will be not only large but also very complex. It needs to connect different systems and networks, from generation facilities and distribution equipment to intelligent end points and communication networks, which are possibly deregulated and owned by several entities. It can be expected that the heterogeneity, diversity, and complexity of smart grid components may introduce new vulnerabilities, in addition to the common ones in interconnected networks and stand-alone microgrids [3]. To make the situation even worse, the sophisticated control, estimation, and pricing algorithms incorporated in the grid may also create additional vulnerabilities.

The first-ever control system malware called Stuxnet was found in July 2010. This malware, targeting vulnerable SCADA systems, raises new questions about power grid security [6]. SCADA systems are currently isolated, preventing external access. Malware, however, can spread using USB drives and can be specifically crafted to

Table 1 Taxonomy of Attacks and Consequences in Cyber and Physical Systems

Attack \ Consequence	Cyber	Physical
Cyber	Eavesdropping of private information	Stuxnet
Physical	Meter bypassing	Instability due to physical destructions

sabotage SCADA systems that control electric grids. Furthermore, increasingly interconnected smart grids will unfortunately provide external access which in turn can lead to compromise and infection of components.

Many warnings concerning the security of smart grids are appearing [7]–[12] and some guidelines have been published, such as NISTIR 7628 [3] and NIST SP 1108 [13]. This paper argues that a new approach to security, bringing together cyber security and system theory under the name of cyber–physical security (CPS), is needed to address the requirements of complex, large-scale infrastructures like the smart grid. In such systems, cyber attacks can cause disruptions that transcend the cyber realm and affect the physical world. Stuxnet is a clear example of a cyber attack used to induce physical consequences. Conversely physical attacks can affect the cyber system. For example, the integrity of a meter can be compromised by using a shunt to bypass it. Secrecy can be broken by placing a compromised sensor beside a legitimate one. As physical protection of all assets of large-scale physical systems, such as the smart grid, is economically infeasible, there arises the need to develop methods and algorithms that can detect and counter hybrid attacks. Based on the discussions at the Army Research Office workshop on CPS security in 2009, we classify current attacks on cyber–physical systems into four categories and provide examples to illustrate our classification in Table 1. Although cyber security and system theory have achieved remarkable success in defending against pure cyber or pure physical attacks, neither of them alone is sufficient to ensure smart grid security, due to hybrid attacks. Cyber security is not equipped to provide an analysis of the possible consequences of attacks on physical systems. System theory is usually concerned with properties such as performance, stability, and safety of physical systems. Its theoretical framework, while well consolidated, does not provide a complete modeling of the IT infrastructure.

In this paper, we propose to combine system theory and cyber security to ultimately build a science of cyber–physical security. Toward this goal, it is important to develop cyber–physical security models capable of integrating dynamic systems and threat models within a unified framework. We believe that cyber–physical security can not only address problems that cannot be currently solved but provide new improved solutions for detection, response, reconfiguration, and restoration of system functionalities while keeping the system operating. We also believe that some existing modeling formalisms can be used as a starting point toward a systematic treatment of cyber–physical security. Game theory [14] can capture the adversarial nature of the interaction between an attacker and a defender. Networked control systems [15] aim at integrating computing and communication technologies with system theory, providing a common modeling framework for cyber–physical systems. Finally, hybrid dynamic systems [16] can capture the discrete nature of events such as attacks on control systems.

The rest of the paper motivates the need for cyber–physical security in the context of the smart grid. Section II reviews cyber threats and countermeasures. Section III describes system-theoretic approaches to contingency analysis and detection of anomalies in the sensory system. Section IV shows how methods from either domain may be incapable to address specific security threats. Section V provides examples of the unique features of cyber–physical security. Finally, Section VI concludes the paper with future research directions.

II. CYBER SECURITY APPROACHES

This section delineates cyber security approaches to smart grid security.

A. System Model

As Fig. 3 shows, smart grids consist of four components: generation, transmission, distribution, and consumption. In the consumption component, customers use

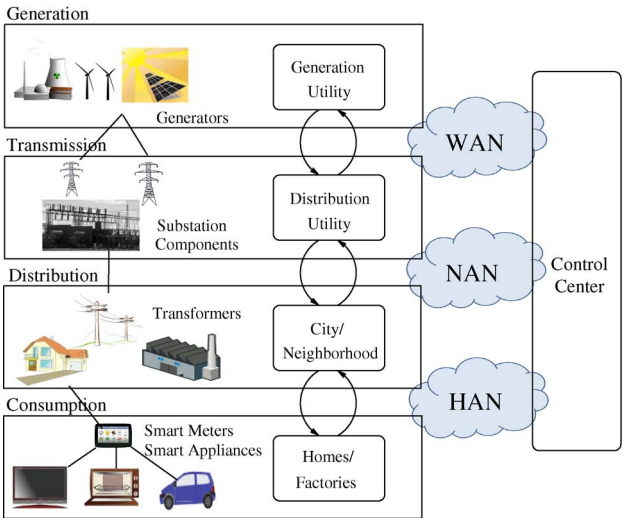


Fig. 3. *A cyber security view of smart grid.*

electric devices (e.g., smart appliances, electric vehicles), and their usage of electricity will be measured by an enhanced metering device, called a smart meter. The smart meter is one of the core components of the advanced metering infrastructure (AMI) [17]. The meter can be collocated and interact with a gateway of a home-area network (HAN) or a business-area network (BAN). For simple illustration, we denote a smart meter in the figure as a gateway of a HAN. A neighbor-area network (NAN) is formed under one substation, where multiple HANs are hosted. Finally, a utility company may leverage a wide-area network (WAN) to connect distributed NANs.

B. Cyber Security Requirements

In this section, we analyze the information security requirements for smart grids. In general, information security requirements for a system include three main security properties: confidentiality, integrity, and availability. Confidentiality prevents an unauthorized user from obtaining secret or private information. Integrity prevents an unauthorized user from modifying the information. Availability ensures that the resource can be used when requested.

As shown in Fig. 4, price information, meter data, and control commands are the core information exchanged in smart grids which we consider in this paper.

While more types of information are exchanged in reality, these core information types provide a comprehensive sample of security issues.

We now examine the importance of protecting the core information types with respect to the main security properties. The degree of importance for price information, control commands, and meter data is equivalent to the use cases of NISTIR 7628 [3], to which we added the degree of importance for software. The most important requirement for protecting smart grids are outlined below.

- *Confidentiality of power usage*: Confidentiality of meter data is important, because power usage data provides information about the usage patterns for individual appliances, which can reveal personal activities through nonintrusive appliance monitoring [18]. Confidentiality of price information and control commands are not important in cases where it is public knowledge. Confidentiality of software should not be critical, because the security of the system should not rely on the secrecy of the software, but only on the secrecy of the keys, according to Kerckhoffs's principle [19].
- *Integrity of data, commands, and software*: Integrity of price information is critical. For instance, negative prices injected by an attacker can cause an electricity utilization spike as numerous devices would simultaneously turn on to take advantage of the low price. Although integrity of meter data and commands is important, their impact is mostly limited to revenue loss. On the other hand, integrity of software is critical since compromised software or malware can control any device and grid component.
- *Availability against DoS/DDoS attacks*: Denial-of-service (DoS) attacks are resource consumption attacks that send fake requests to a server or a network, and distributed DoS (DDoS) attacks are accomplished by utilizing distributed attacking sources such as compromised smart meters and appliances. In smart grids, availability of information and power is a key aspect [20]. More specifically, availability of price information is critical due to serious financial and possibly legal implications. Moreover, outdated price information can adversely affect demand. Availability of commands is also important, especially when turning a meter back on after completing the payment of an electric bill. On the other hand, availability of meter data (e.g., power usage) may not be as critical because the data can usually be read at a later point.

From the above discussion, we can summarize the importance of data, commands, and software, which are shown in Table 2. "High" risk implies that a property of certain information is very important/critical, and "medium" and "low" risks classify properties that are important and noncritical, respectively. This classification enables prioritization of risks, to focus effort on the most critical aspects first. For example, integrity of price information is more important than its confidentiality; consequently, we need to focus on efficient cryptographic authentication mechanisms before encryption.

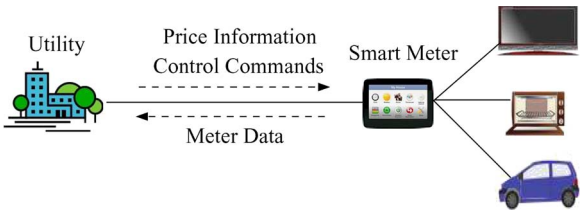


Fig. 4. ***Information flows to/from a smart meter including price information, control commands, and meter data.***

C. Attack Model

To launch an attack, an adversary must first exploit entry points, and upon successful entry, an adversary can deliver specific cyber attacks on the smart grid infrastructure. In the following sections, we describe this attacker model in detail.

1) Attack Entry Points: In general, strong perimeter defense is used to prevent external adversaries from accessing information or devices within the trusted grid zone.

Table 2 The Importance of Security Properties for Data, Commands, and Software

	Price information	Control command	Meter data	Software
Confidentiality	Low	Low	Medium	Low
Integrity	High	High	High	High
Availability	High	High	Low	N/A

Unfortunately, the size and complexity of grid networks provide numerous potential entry points as follows.

- *Inadvertent infiltration through infected devices*: Malicious media or devices may be inadvertently infiltrated inside the trusted perimeter by personnel. For example, *USB memory sticks* have become a popular tool to circumvent perimeter defenses: a few stray USB sticks left in public spaces are picked up by employees and plugged into previously secure devices inside the trusted perimeter, enabling malware on the USB sticks to immediately infect the devices. Similarly, *devices used both inside and outside the trusted perimeter* can get infected with malware when outside, and infiltrate that malware when used inside. Common examples are corporate laptops that are privately used at home over the weekend.
- *Network-based intrusion*: Perhaps the most common mechanism to penetrate a trusted perimeter is through a network-based attack vector. Exploiting *poorly configured firewalls* for both misconfigured inbound and faulty outbound rules is a common entry point, enabling an adversary to insert a malicious payload onto the control system.

Backdoors and holes in the network perimeter may be caused by components of the IT infrastructure with vulnerabilities or misconfigurations. Networking devices at the perimeter (e.g., fax machines, forgotten but still connected modems) can be manipulated for bypassing proper access control mechanisms. In particular, *dialup access to remote terminal units (RTUs)* is used for remote management, and an adversary can directly dial into modems attached to field equipment, where many units do not require a password for authentication or have unchanged default passwords. Further, adversaries can exploit vulnerabilities of the devices and install backdoors for future access to the prohibited area. *Exploiting trusted peer utility links* is another potential network-based entry point.

An attacker could wait for a legitimate user to connect to the trusted control system network via VPN and then *hijack that VPN connection*. The network-based intrusions described above are particularly dangerous because they enable a remote adversary to enter the trusted control-system network.

- *Compromised supply chain*: An attacker can preinstall malicious codes or backdoors into a device prior to shipment to a target location, called *supply chain attacks*. Consequently, the need for security assurance in the development and manufacturing process for sourced software, firmware, and equipment is critical for safeguarding the cyber supply chain involving technology vendors and developers.
- *Malicious insider*: An employee or legitimate user who is authorized to access system resources can perform actions that are difficult to detect and prevent. Privileged insiders also have intimate knowledge of the deployed defense mechanisms, which they can often easily circumvent. Trivial accessibility to smart grid components will increase the possibility of escalating an authorized access to a powerful attack.

2) Adversary Actions: Once an adversary gains access to the power control network, he can perform a wide range of attacks. Table 3 lists actions that an adversary can perform to violate the main security properties (confidentiality, integrity, availability) for the core types of information. We classify more specific cyber attacks that lead to either cyber or physical consequences.

Cyber consequences:

- *Malware spreading and controlling devices*: An adversary can develop malware and spread it to infect smart meters [21] or company servers. Malware can be used to replace or add any function to a device or a system such as sending sensitive information or controlling devices.
- *Vulnerabilities in common protocols*: Smart grid components will use existing protocols, inheriting the vulnerabilities on the protocols. Common protocols may include TCP/IP, and remote procedure call (RPC).

Table 3 Threat Type Classification as Caused by Attacking Security Properties

	Price information	Control command	Meter data	Software
Confidentiality	Leakage of price info.	Exposure of control structure	Unauthorized access to meter data	Theft of proprietary software
Integrity	Incorrect price info.	Changes of control commands	Incorrect meter data	Malicious software
Availability	Unavailability of price info.	Inability to control grid	Unavailability of billing info.	N/A

- *Access through database links*: Control systems record their activities onto a database on the control system network then mirror logs into the business network. A skilled attacker can gain access to the database on the business network, and the business network gives a path to the control system network. Modern database architectures allow this type of attack if they are improperly configured.
- *Compromising communication equipments*: An attacker can potentially reconfigure or compromise some of the communication equipment, such as multiplexers.
- *Injecting false information on price and meter data*: An adversary can send packets to inject false information on current or future prices, or send wrong meter data to a utility company. Results of injecting false prices, such as negative pricing, will be power shortage or other significant damages on the target region. Results of sending wrong data include reduced electric bills for economic damages due to the loss of revenue of a utility company. Also, fake information can give huge financial impacts on electricity markets [12].
- *Eavesdropping attacks*: An adversary can obtain sensitive information by monitoring network traffic, which results in privacy breaches by stealing power usage, disclosure of the controlling structure of smart grids and future price information. Such eavesdropping can be used for gathering information to perpetrate further crimes. For example, an attacker can gather and examine network traffic to deduce information from communication patterns, and even encrypted communication can be susceptible to traffic analysis attacks.
- *Modbus security issues*: A SCADA protocol of noteworthy concern is the Modbus protocol [22], which is widely used in industrial control applications such as in water, oil, and gas infrastructures. The Modbus protocol defines the message structure and communication rules used by process control systems to exchange SCADA information for operating and controlling industrial processes. Modbus is a simple client-server protocol that was originally designed for low-speed serial communication in process control networks. Given that the Modbus protocol was not designed for highly security-critical environments, several attacks are possible.
  1) *Broadcast message spoofing*: This attack involves sending fake broadcast messages to slave devices.
  2) *Baseline response replay*: This attack involves recording genuine traffic between a master and a field device, and replaying some of the recorded messages back to the master.
  3) *Direct slave control*: This attack involves locking out a master and controlling one or more field devices.
  4) *Modbus network scanning*: This attack involves sending benign messages to all possible addresses on a Modbus network to obtain information about field devices.
  5) *Passive reconnaissance*: This attack involves passively reading Modbus messages or network traffic.
  6) *Response delay*: This attack involves delaying response messages so that the master receives out-of-date information from slave devices.
  7) *Rogue interloper*: This attack involves attacking a computer with the appropriate (serial or Ethernet) adapters to an unprotected communication link.

Physical consequences:

- *Interception of SCADA frames*: An attacker can use a protocol analysis tool for sniffing network traffic to intercept SCADA Distributed Network Protocol 3.0 (DNP3) frames and collect unencrypted plaintext frames that would provide valuable information, such as source and destination addresses. This intercepted data, which include control and setting information, could then be used at a later date on another SCADA system or intelligent equipment device (IED), thereby shutting services down at worst or at the minimum causing service disruptions.
- *Malware targeting industrial control systems*: An attacker can successfully inject worms into vulnerable control systems and reprogram industrial control systems. A well-known example is Stuxnet as discussed in Section I.
- *DoS/DDoS attacks on networks and servers*: An adversary can launch a DoS/DDoS attack against various grid components including smart meters, networking devices, communication links, and utility business servers. If the attack is successful, then electricity cannot be controlled in the target region. Furthermore, power supply can be stopped from the result of the attack.
- *Sending fake commands to smart meters in a region*: An adversary can send fake commands to a device or a group of devices in a target region. For example, sending disconnect messages to smart meters in a region will stop power delivery to that region. As well, invalid switching of electric devices can result in unsafe connections which may lead to burn the target place on fire. Thus, insecure communication in smart grids may be able to threaten human life.

The attacks mentioned above are not exhaustive, but they serve to illustrate risks to help develop secure grid

systems. Additional examples of SCADA threats are available at the web site of US-CERT.[1]

D. Countermeasures

1) Key Management: Key management is a fundamental approach for information security. Shared secret keys or authentic public keys can be used to achieve secrecy and authenticity for communication. Authenticity is especially important to verify the origin which in turn is key for access control.

The key setup in a system defines the root of trust. For example, a system based on public/private keys may define the public key of a trust center as the root of trust, and the trust center's private key is used to sign certificates and delegate trust to other public keys. In a symmetric-key system, each entity and the trust center would set up shared secret keys and establish additional trust relationships among other nodes by leveraging the trust center, as in Kerberos.

The challenge in this space is key management across a very broad and diverse infrastructure. As a recent NIST report documents [3], several dozens of secure communication scenarios are required, ranging from communication between the power distributor and the smart meter to communication between equipment and field crews. For all these communication scenarios, keys need to be set up to ensure secrecy and authenticity. Besides the tremendous diversity of equipment, there is also a wide variety of stakeholders: government, corporations, and consumers. Even secure e-mail communication among different corporations is a challenge today; yet the secure communication between equipment from one corporation and a field crew of another one poses numerous additional challenges. By adding a variety of key management operations to the mix (e.g., key refresh, key revocation, key backup, key recovery), the complexity of key management becomes truly formidable. Moreover, business, policy, and legal aspects also need to be considered, as a message signed by a private key can hold the key owner liable for the contents. A recent publication from NIST provides a good guideline for designing cryptographic key management systems to support an organization [23], but the diverse requirements of smart grid infrastructures are not considered.

2) Secure Communication Architecture: Designing a highly resilient communication architecture for a smart grid is critical to mitigate attacks while achieving high-level availability. Here are the required components.

- *Network topology design*: A network topology represents the connectivity structure among nodes, which can have an impact on the robustness against attacks [24]. Thus, connecting networking nodes to be highly resilient under attack can be the basis to build a secure communication architecture.
- *Secure routing protocol*: A routing protocol on a network is to build logical connectivity among nodes, and one simplest way to prevent communication is by attacking the routing protocol. By compromising a single router and by injecting bogus routes, all communication in the entire network can come to a standstill. Thus, we need to consider the security of a routing protocol running on top of a network topology.
- *Secure forwarding*: An adversary who controls a router can alter, drop, and delay existing data packets or inject new packets. Thus, securing individual routers and detecting malicious behaviors will be required to achieve secure forwarding.
- *End-to-end communication*: From end-to-end perspective, secrecy and authenticity of data are the most crucial properties. Secrecy prevents an eavesdropper from learning the data content, while authenticity (sometimes referred to as integrity) enables the receiver to verify that the data indeed originated from the sender, thus preventing an attacker from altering the data.

While numerous protocols exist (e.g., SSL/TLS, IPsec, SSH), some low-power devices may need lightweight protocols to perform the associated cryptography.

- *Secure broadcasting*: Many smart grid environments rely on broadcast communication. Especially for price dissemination, authenticity of the information is important, because an adversary could inject a negative cost and cause an electricity utilization to spike when numerous devices simultaneously turn on to take advantage of the low price.
- *DoS defense*: Given all the above mechanisms, an adversary can still prevent communication by mounting a DoS attack. For example, if an adversary controls many end points after compromising them, he can use these end points to send data to flood the network. Hence, enabling communication under these circumstances is crucial, for example to perform network management operations to defend against the attack. Moreover, electricity itself, rather than communication networks, can be a target of DoS attacks [25].
- *Jamming defense*: To prevent an external adversary from jamming the wireless network, jamming detection mechanisms can be used to detect attacks and raise alarms. A multitude of methods to counter jamming attacks has been developed [26], enabling operation during jamming.

3) System and Device Security: An important area is to address vulnerabilities that enable exploitation through software-based attacks, where an adversary either exploits a software vulnerability to inject malicious code into a

[1]http://www.us-cert.gov/control_systems/csvuls.html

system, or where a malicious insider uses administrative privileges to install and execute malicious code. The challenge in such an environment is to obtain "ground truth" when communicating with a potentially compromised system: Is the response sent by legitimate code or by malware? An illustration of this problem is when we attempt to run a virus scanner on a potentially compromised system—If the virus scanner returns the result that no virus is present, is that really because no virus could be identified or is it because the virus has disabled the virus scanner? A related problem is that current virus scanners contain an incomplete list of virus signatures, and the absence of a virus detection could be because the virus scanner does not yet recognize the new virus.

In the context of smart grids, researchers have proposed several techniques to provide prevention and detection mechanisms against malware. McLaughlin *et al.* have proposed diversity for embedded firmware [27] to avoid an apocalyptic scenario where malware pervasively compromises equipment, because each device executes different software, thus avoiding common vulnerabilities.

A promising new approach to provide remote code verification is a technology called *attestation*. Code attestation enables an external entity to inquire the software that is executing on a system in a way that prevents malware from hiding. Since attestation reveals a signature of executing code, even unknown malware will alter that signature and can thus be detected. In this direction, LeMay *et al.* have studied hardware-based approaches for attestation [28], [29]. Software-based attestation is an approach that does not rely on specialized hardware, but makes some assumptions that the verifier can uniquely communicate with the device under verification [30]. Shah *et al.* have demonstrated the feasibility of this concept on SCADA devices [31].

III. SYSTEM-THEORETIC APPROACHES

In this section, we want to focus on system-theoretic approaches to the real-time security of smart grids, which encompasses two main parts: contingency analysis (CA) and system monitoring [32].

A. System Model

Fig. 5 shows a typical system-theoretic view of an IEEE 14-bus system. The focus of such a view is the physical interactions between each component in the grid, while the cyber view focuses on the modeling of IT infrastructures.

Suppose the grid consists of N buses. Let us define the active power flow, reactive power flow, the voltage magnitude, and phase angle for each bus as P_i, Q_i, V_i, and θ_i, respectively.[2] Let us define vectors P, Q, V, and θ as the collections of P_i, Q_i, V_i, and θ_i, respectively.

[2]We assume that bus N is the reference bus and the phase angle of it is 0.

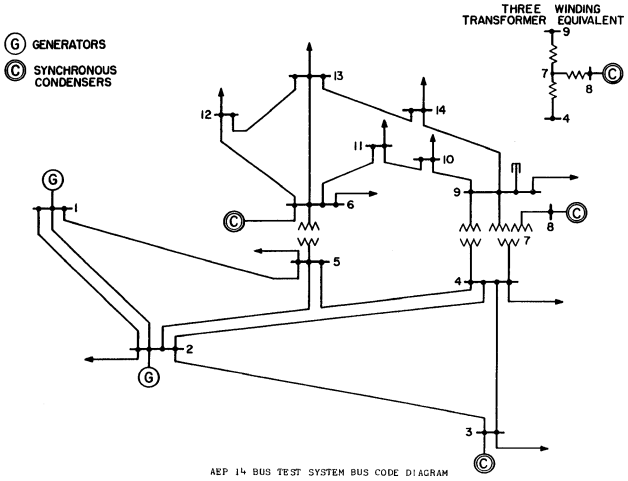


Fig. 5. ***A typical system-theoretic view of an IEEE standard 14-bus system.***

The relationship between node current I_k and voltage $V_k e^{j\theta_k}$ is given by the following linear equations [33]:

$$I_k = \sum_{i=1}^{N} Y_{ki} V_i e^{j\theta_i}$$

where Y_{ki} is the admittance between bus k and i. As a result, the active and reactive power at node k are given by

$$P_k + jQ_k = V_k e^{j\theta_k} \times \overline{I_k} = V_k e^{j\theta_k} \sum_{i=1}^{N} \overline{Y_{ki} V_i e^{j\theta_i}} \qquad (1)$$

where $\overline{I_k}$ means complex conjugate. It can be seen that V and θ are the states of the system since they completely determine power flow P and Q. Let us define the state[3] x as $x = [V', \theta_1, \ldots, \theta_{N-1}]' \in \mathbb{R}^{2N-1}$. The remote terminal units (RTUs) provide the system's measurements. Let us denote as $z \in \mathbb{R}^m$ the collection of all measurements, assumed to satisfy the following equation:

$$z = h(x) + v \qquad (2)$$

where $h : \mathbb{R}^{2N-1} \to \mathbb{R}^m$ represents the sensor model and $v \in \mathbb{R}^m$ denotes the measurement noise, which is further assumed to be Gaussian with mean 0 and covariance R.

Here we briefly introduce the weighted least square (WLS) estimator [34], as it is widely used in practice. Define the estimated state as $\hat{x}$, and the residue vector as $r = z - h(\hat{x})$, which measures the inconsistency between

[3]The state does not include θ_N as its phase angle is assumed to be 0.

state estimation $\hat{x}$ and measurements z. A WLS estimator tries to find the best estimation $\hat{x}$ with minimum inconsistency. In particular, the WLS estimator computes $\hat{x}$ based on the following minimization problem:

$$\hat{x} = \text{argmin}_{\hat{x}}\ r^T R^{-1} r. \tag{3}$$

B. Security Requirements

The U.S. Department of Energy (DoE) *Smart Grid System Report* [35] summarizes six characteristics of the smart grid, which were further developed from the seven characteristics of "Characteristics of the Modern Grid" [36] published by the National Energy Technology Laboratory (NETL). With respect to security, the most important characteristic identified by DoE is to *operate resiliently even during disturbances, attacks, and natural disasters*.

In real-time security settings, the following properties are essential for the resilience of smart grids:

1) the power system should withstand a prespecified list of contingencies;
2) the accuracy of state estimation should degrade gracefully with respect to sensor failures or attacks.

The first property is passive and prevention based. The second property enables the detection of attacks or abnormalities and helps the system operator actively mitigate the damage.

C. Attack Model

A contingency can usually be modeled as a change in vectors P, Q, V, θ (such as a loss of a generator) or as a change in the admittance Y_{ki} (such as an opening transmission line). For system monitoring, corrupted measurements can be modeled as an additional term in (2), i.e.,

$$z^a = z + u = h(x) + v + u \tag{4}$$

where $u = [u_1, \ldots, u_m]' \in \mathbb{R}^m$ and $u_i \neq 0$ only if the sensor i is corrupted.

D. Countermeasures

1) Contingency Analysis: Contingency analysis checks if the steady-state system is outside operating region for each contingency [32]. However, the number of potential contingencies is high for large power grids. Due to real-time constraints, it is impossible to evaluate each contingency. As a result, in practice, usually only "$N-1$" contingencies are evaluated, via considering single failure cases instead of multiple ones. Moreover, the list of possible contingencies is usually screened and ranked. After that, a selected number of contingencies is evaluated. If a violation occurs, the system needs to determine the control actions that can mitigate or completely eliminate the violation.

2) Bad Data Detection: Bad data detector such as χ^2 or largest normalized residue detector [34] detects the corruption in measurement z by checking the residue vector r. For uncorrupted measurements, it is expected that the residue vector r will be small since z should be consistent with (2). However, such a detection scheme has an inherent vulnerability as different z vectors can generate the same residue r. By exploiting this vulnerability, Liu *et al.* [10] show that an adversary can inject a stealthy input u into the measurements to change the state estimate $\hat{x}$ and fool the bad data detector at the same time. Sandberg *et al.* [37] consider how to find a sparse stealthy u, which enables the adversary to launch an attack with a minimum number of compromised sensors. To counter such a vulnerability, Kosut *et al.* [38] suggest using the prior knowledge of the state x to help detecting malicious sensors.

IV. THE NEED FOR CYBER–PHYSICAL SECURITY

Table 4 summarizes the discussion in Sections II and III. The cyber security approaches focus on the IT infrastructures of the smart grid while system-theoretic approaches focus more on the physical aspects. We argue that pure cyber or system-theoretic approaches are insufficient to

Table 4 Comparison Between Cyber and System-Theoretic Security

	Cyber Security	System Theoretic Security
System Model	WAN/NAN/HAN model	Power Flow Model Sensor Model
Requirements	Confidentiality Integrity Availability	Robust to Prespecified Contingency Accurate State Estimation
Attack Model	DoS attack Network-Based Intrusion ...	Contingencies Sensor Failures, False Data Injection
Countermeasures	Key Management Secure Communication System and Device Security	Contingency Analysis Bad Data Detection

guarantee security of the smart grid, for the following reasons:

1) *The system and attack models of both approaches are incomplete*: Cyber security does not model the physical system. Therefore, cyber security can hardly defend against physical attacks. For example, cyber security protects the integrity of measurements data by using secure devices and communication protocols.
   However, integrity of sensors can be broken by modifying the physical state of the system locally, e.g., shunt connectors can be placed in parallel with a meter to bypass it and cause energy theft. In that case, no purely cyber security method can be employed to effectively detect and counter such attacks, since the cyber portion of the system is not compromised. Thus, even the goals of cyber security cannot be achieved by pure cyber approaches in cyber–physical systems. Moreover, cyber security is not well equipped to predict the effect of cyber attacks and countermeasures on the physical system. For example, the DoS attacks can cause drops of measurements data and control command, which can lead to instability of the grid. A countermeasure to DoS attacks is to isolate some of the compromised nodes from the network, which may result in even more severe stability issues. Thus, an understanding of the physical system is crucial even for defending against cyber attacks.
   On the other hand, the system-theoretic model does not model the whole IT infrastructures, but usually just a high level abstraction. As a result of this oversimplification of the cyber world, it difficult to analyze the effect of cyber attacks on physical systems. For example, in DoS attacks, some control commands may be dropped due to limited bandwidth. However, the effect of the lossy communication cannot be evaluated in a pure power flow model.
2) *The security requirements of both approaches are incomplete and the security of the smart grid requires both of them*: System level concerns, such as stability, safety, and performance, have to be guaranteed in the event of cyber attacks. Cyber security metrics do not currently include the aforementioned metrics. On the other hand, system theory is not concerned with secrecy of information. Furthermore, it usually treats integrity and availability of information as intermediate steps to achieve stability, safety, or better performance. In the design of secure smart grid it is important to identify a set of metrics that combines and addresses the concerns of the two communities.
3) *The countermeasures of both approaches have drawbacks*: System-theoretic methods will not be able to detect any attack until it acts on the physical system. Furthermore, since system theory is based on approximate models and is subject to unknown disturbances, there will always be a discrepancy between the observed and the expected behavior. Most of the attack can bypass system theory-based intrusion detection algorithms with a small probability, which could be detrimental. Last, contingency analysis generally focuses on $N-1$ contingencies, which is usually enough for independent equipment failures. However, as we integrate the IT infrastructures into the smart grid, it is possible that several contingencies will happen simultaneously during an attack.
   On the other hand, cyber countermeasures alone are not sufficient to guarantee security of the smart grid. History has so far taught that cyber security is not always bulletproof. As operational continuity is essential, the system must be built to withstand and operate even in the event of zero-day vulnerabilities or insider threats, resorting to rapid reconfiguration to provide graceful degradation of performance in the face of an attack. As a large blackout can happen in a few minutes [39], it is questionable that pure cyber security approaches can react fast enough to withstand zero-day vulnerability exploits or insider attacks.

V. CYBER–PHYSICAL SECURITY

As shown in Section IV, both cyber and system-theoretic approaches are essential for the security of smart grids. In this section, we want to use two examples to show how the combination of cyber and system-theoretic approaches together can provide better security level than traditional methods. In the first example, we show how system-theoretic countermeasures can be used to defend against a replay attack, which is a cyber attack on the integrity of the measurement data. In the second example, we show how system theory can guide cyber security investment strategies.

A. Defense Against Replay Attacks

In this example, we consider defense against replay attack, where an adversary records a sequence of sensor measurements and replays the sequence afterwards. Replay attacks are cyber attacks which break the integrity or more precisely the freshness of measurements data. It is worth mentioning that Stuxnet [40] employed a replay attack of this type to cover its goal of damaging the centrifuges in a nuclear facility by inducing excessive vibrations or distortions. While acting on the physical system, the malware was reporting old measurements indicating normal operations. This integrity attack, clearly conceived and operated in the cyber realm, exploited four zero-day vulnerabilities to break the cyber infrastructures and it remained undiscovered for several months after its release. Therefore, a pure cyber approach to replay attacks may not be able to react fast enough before the system is damaged.

Next we develop the concept of physical authentication, a methodology that can detect such attacks independently of the type of attack used to gain access to the control system. This algorithm [41] was developed long before Stuxnet appeared and preceded it. We are reporting a summary below.

To achieve greater generality, the method is presented for a generic control system. We assume the sensors are monitoring a system with the following state dynamics:

$$x_{k+1} = Fx_k + Bu_k + w_k \tag{5}$$

where $x_k \in \mathbb{R}^n$ is the vector of state variables at time k, $w_k \in \mathbb{R}^n$ is the process noise at time k, and x_0 is the initial state. We assume w_k, x_0 are independent Gaussian random variables, $x_0 \sim \mathcal{N}(\bar{x}_0, \Sigma)$, $w_k \sim \mathcal{N}(0, Q)$.

For each sampling period k, the true measurement equation of the sensors can be written as

$$z_k = Hx_k + v_k \tag{6}$$

where $z_k \in \mathbb{R}^m$ is a collection of all the measurements from sensors at time k and $v_k \sim \mathcal{N}(0, R)$ is the measurement noise independent of x_0 and w_k.

We assume that an attacker records a sequence of measurements from time T_0 to time $T_0 + T - 1$ and replays it from time $T_0 + T$ to time $T_0 + 2T - 1$, where $T_0 \geq 0, T \geq 1$. As a result, the corrupted measurements z_k^a received by the system operator are

$$z_k^a = \begin{cases} z_k, & 0 \leq k \leq T_0 + T - 1 \\ z_{k-T}, & T_0 + T \leq k \leq T_0 + 2T - 1. \end{cases} \tag{7}$$

Our goal is to design an estimator, a controller and a detector such that:

1) the system is stable when there is no replay attack;
2) the detector can detect the replay attack with a high probability.

We propose the following design of a fixed gain estimator, a fixed gain controller with random disturbance and a χ^2 detector. In particular, our estimator takes the following form:

$$\hat{x}_{k+1} = F\hat{x}_k + Bu_k + Kr_{k+1}, \qquad \hat{x}_0 = \bar{x}_0. \tag{8}$$

where K is the observation gain matrix and the residue r_k is computed as

$$r_{k+1} = z_{k+1}^a - C(F\hat{x}_k + Bu_k). \tag{9}$$

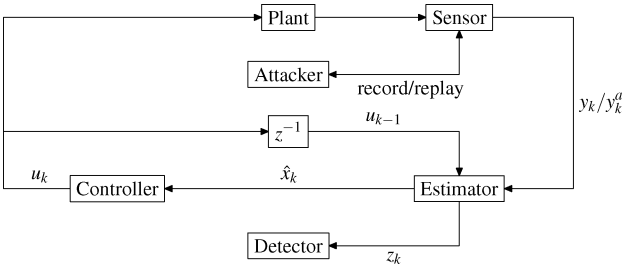


Fig. 6. ***System diagram.***

Our controller takes the following form:

$$u_k = L\hat{x}_k + \Delta u_k \tag{10}$$

where L is the control gain matrix and Δu_ks are independent identically distributed (i.i.d.) Gaussian noises generated by the controller, with zero mean and covariance $\mathcal{Q}$.

It can be easily shown that the residue r_k is a Gaussian random variable with zero mean when there is no replay. As a result, with large probability it cannot be far away from 0. Therefore, we design our filter to trigger an alarm at time k based on the following event:

$$\left\{ g_k = r_k' \mathcal{P} r_k \geq \text{ threshold} \right\} \tag{11}$$

where $\mathcal{P}$ is a predefined weight matrix. Fig. 6 shows the diagram of the proposed system.

We first consider the stability of the proposed system. It is well known that without Δu_k, the closed-loop system without replay is stable if and only if both $F - KCF$ and $F + BL$ are stable. Moreover, one can easily prove that adding Δu_k does not affect the stability of the system since Δu_k is i.i.d. Gaussian distributed. Hence, to ensure that the system is closed-loop stable without replay, we only need to make $F - KCF$ and $F + BL$ stable, which can be easily done as long as the system is both detectable and stabilizable.

Now we want to show our system design can successfully detect replay attacks. Consider the residue r_k, where $T_0 + T \leq k \leq T_0 + 2T - 1$, then one can prove that

$$r_k = r_{k-T} + C\mathcal{A}^{k-T_0-T}(I - KC)(\hat{x}_{T_0} - \hat{x}_{T_0+T}) + \sum_{i=0}^{k-T-T_0-1} C\mathcal{A}^i B(\Delta u_{k-T-1-i} - \Delta u_{k-1-i})$$

where $\mathcal{A} = (F + BL)(I - KC)$. The second term above converges to 0 exponentially fast if $\mathcal{A}$ is stable. As a result, if we do not introduce any random control disturbance, i.e., $\Delta u_k = 0$, then the third term vanishes

and the residue r_k under replay attack converges to the residue r_{k-T} when no replay attack is present. Therefore, the detection rate of the replay attack will be the same as the false alarm rate. In other words, the detector cannot distinguish between healthy and corrupted measurements. However, if $\Delta u_k \neq 0$, then the third term will always be present and therefore the detector can detect replay attacks with a probability larger than the false alarm rate.

It is worth mentioning that the role of Δu_k is similar to an authentication signal on the measurements. When the system is under normal operation, it is expected that the measurements z_k will reflect the random disturbances Δu_k. On the other hand, when the replay begins, z_k and Δu_k become independent of each other. Therefore, the integrity and freshness of the measurements can be protected by checking the correlation between z_k and Δu_k. This technique is cyber–physical as it uses the physics of the system to authenticate data coming from the cyber portion.

We now wish to provide a numerical example to illustrate the performance of our detection algorithm. We impose the following parameters: $F = B = Q = R = \mathcal{P} = 1$, $K = 0.9161$, $L = -0.618$. One can verify that $\mathcal{A} = 0.0321 < 1$. The threshold of the filter is chosen such that the false alarm rate is 1%. We assume that the recording starts at time 1 and replay starts at time 11. Fig. 7 shows different detection rate over time as $\mathcal{Q}$ increases. It can be seen that the detection fails when there is no disturbance. Moreover, a larger disturbance can increase the performance of the detector.

B. Cyber Security Investment

In this example, we show how system theory can be used to expose the critical assets to protect and thus provide important insights toward the allocation of security investments. In particular, we consider how to deploy secure sensors to help detect corrupted measurements. We assume the true measurements of sensors follow a linearized model of (2), as discussed in Section III

$$z = Hx + v \tag{12}$$

where $z \in \mathbb{R}^m$ and $x \in \mathbb{R}^{2N-1}$ and $H \in \mathbb{R}^{m\times(2N-1)}$ is assumed to be of full column rank. For linearized models, (3) can be solved analytically as

$$\hat{x}(z) = (H'R^{-1}H)^{-1}H'R^{-1}z = Kz. \tag{13}$$

Therefore, the residue can be calculated explicitly as

$$r(z) = z - H\hat{x}(z) = (I - HK)z = Sz \tag{14}$$

where $S = I - HK$.

Suppose that an attacker is able to modify the readings of a subset of sensors. As a result, the corrupted measurements take the following form:

$$z^a = z + u = Hx + v + u \tag{15}$$

where $u = [u_1, \ldots, u_m]' \in \mathbb{R}^m$ indicates the error introduced by the attacker and $u_i \neq 0$ only if sensor i is compromised.

An attack is called stealthy if the residue r does not change during the attack. In mathematical terms, a stealthy attack u satisfies $r(z) = r(z + u)$. Since $r(z)$ is linear with respect to z, we can simplify the above equation to

$$r(u) = Su = 0 \tag{16}$$

without loss of generality.

As shown by Liu *et al.* [10], the χ^2 detectors fail to detect a stealthy input u. In fact, any detector based on r is ineffective against stealthy attacks as they do not change the residue r. On the other hand, the stealthy attack can introduce estimation error to $\hat{x}$.

To defend against such attacks, we deploy secure devices, such as tamper resistant devices, to protect the sensors. To this end, we define a sensor i to be secure if it cannot be compromised, i.e., the corresponding u_i is guaranteed to be 0. Let us also define the set of secure sensors to be $S_e \subseteq \{1, \ldots, m\}$. An attack u is feasible if and only if $u_i = 0$ for all $i \in S_e$.

Our security goal is to deploy the minimum number of secure sensors such that the system can detect the compromised nodes. In other words, we want to find the smallest set S_e such that there is no nonzero feasible and stealthy u.

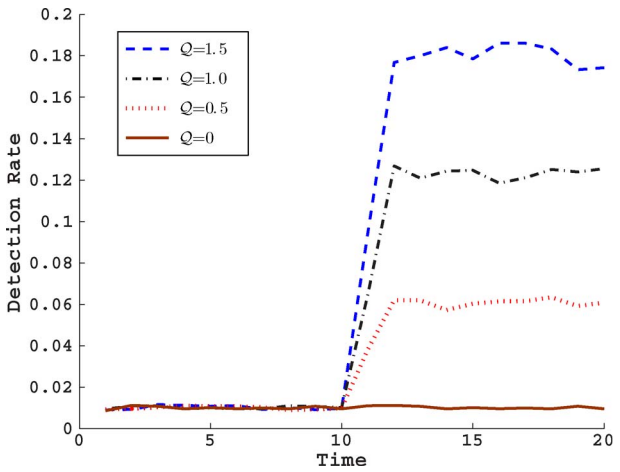


Fig. 7. ***Detection rate over time.***

This problem is of practical importance in the smart grid as the current insecure sensors can only be replaced gradually by secure sensors due to the scale of the grids. As a result, it is crucial to know which set of sensors to replace first to achieve better security.

Let us define $\Gamma(S_e) = \text{diag}(\gamma_1, \ldots, \gamma_m)$, where $\gamma_i = 1$ if and only if $i \in S_e$. A set S_e is called observable if and only if $\Gamma(S_e)H$ is of full column rank. In other words, if a vector $p \in \mathbb{R}^{2N-1} \neq 0$, then $\Gamma(S_e)Hp \neq 0$. The following theorem relates the observability of secure sensor set S_e with the existence of a feasible and stealthy attack u.

Theorem 1: The only feasible and stealthy attack is $u = 0$ if and only S_e is observable.

Proof: First suppose that S_e is observable and u is stealthy and feasible. As a result, $\Gamma(S_e)u = 0$. On the other hand, since u is stealthy, $Su = 0$, which implies that

$$HKu = (I - S)u = u.$$

Therefore

$$\Gamma(S_e)HKu = \Gamma(S_e)u = 0.$$

Since $\Gamma(S_e)H$ is full column rank, we know that $Ku = 0$, which implies that $HKu = 0$. Thus

$$u = (I - HK + HK)u = Su + HKu = 0.$$

On the other hand, suppose that S_e is not observable. Find $x \neq 0$ such that $\Gamma(S_e)Hx = 0$. Choose $u = Hx$. Since H is full column rank, $u \neq 0$. Moreover, $\Gamma(S_e)u = \Gamma(S_e)Hx = 0$. Hence, u is feasible. Finally

$$\begin{aligned} Su &= (I - HK)u = u - H(H'R^{-1}H)^{-1}H'R^{-1}u \\ &= Hx - H(H'R^{-1}H)^{-1}H'R^{-1}Hx = 0 \end{aligned}$$

which implies that u is stealthy. ■

Therefore, finding the smallest S_e such that there is no nonzero feasible and stealthy u is equivalent to finding the smallest observable S_e, which can be achieved using the following theorem:

Theorem 2: If S_e is observable and $\text{rank}(\Gamma(S_e)) > 2N - 1$, then there exists an observable S'_e, which is a proper subset of S_e.

Proof: Let $H' = [H_1, \ldots, H_m]$, where $H_i \in \mathbb{R}^{2N-1}$. Since S_e is observable, $\text{rank}(\gamma_1 H_1, \ldots, \gamma_m H_m) = 2N - 1$. Without loss of generality, let us assume that $S_e = \{1, \ldots, l\}$. Thus, $\gamma_1 = \ldots = \gamma_l = 1$ and $\gamma_{l+1} = \ldots = \gamma_m = 0$, where $l > 2N - 1$. Since $H_i \in \mathbb{R}^{2N-1}, H_1, \ldots, H_l$ are not linearly independent. Hence, there exist $\alpha_1, \ldots, \alpha_l \in \mathbb{R}$ that are not all zero such that $\alpha_1 H_1 + \ldots + \alpha_l H_l = 0$. Without loss of generality, let us assume that $\alpha_l \neq 0$. Therefore

$$\text{span}(H_1, \ldots, H_{l-1}) = \text{span}(H_1, \ldots, H_l) = \mathbb{R}^{2N-1}$$

which implies that $S'_e = \{1, \ldots, l-1\}$ is observable. ■

It is easy to see that $\text{rank}(\Gamma(S_e))$ must be no less than $2N - 1$ to make S_e observable. As a result, one can use the procedure described in the proof of Theorem 2 to find the smallest observable set. Analyses of this kind are essential to prioritize security investments.

Remark 1: It is worth noticing that the attacks we discussed in this section are cyber attacks which have physical consequences. The replay attack itself can render the system unstable if the original system is open-loop unstable or it can enable future attacks on the physical system, as in the case of Stuxnet. The stealthy integrity attack can cause large estimation error and potentially damage the system.

Furthermore, our approaches to security are hybrid in nature. In the first example, we use system-theoretic models and countermeasures to detect replay attacks, which is a cyber attack. Our detection algorithm complements the pure cyber security approaches and provides an additional layer of protection. In the second example, we use a system-theoretic model of the grid to develop an optimal cyber security countermeasure to integrity attacks. The results illustrate that combining cyber security and system theory can provide better level of security for the smart grid.

VI. CONCLUSION AND RESEARCH OUTLOOK

With the proliferation of remote management and control of cyber–physical systems, security plays a critically important role, because the convenience of remote management can be exploited by adversaries for nefarious purposes from the comfort of their homes.

Compared to current cyber infrastructures, the physical component of cyber–physical infrastructures adds significant complexity that greatly complicates security. On the one hand, the increased complexity will require more effort from the adversary to understand the system, but on the other hand, this increased complexity also introduces numerous opportunities for exploitation. From the perspective of the defender, more complex systems require dramatically more effort to analyze and defend, because of the state–space explosion when considering combinations of events.

Current approaches to secure cyber infrastructures are certainly applicable to securing cyber–physical systems: techniques for key management, secure communication (offering secrecy, authenticity, and availability), secure code execution, intrusion detection systems, etc. Unfortunately, these approaches are largely unaware of the physical aspects of cyber–physical systems.

System-theoretic approaches already consider physical aspects in more detail than the traditional security and cryptographic approaches. These approaches model the malicious behaviors as either components' failures, external inputs, or noises, analyze their effects on the system, and design detection algorithms or counter measures to the attacks. The strength of model-based approaches lies in a unified framework to model, analyze, detect, and counter various kinds of cyber and physical attacks. However, the physical world is modeled with approximations and is subject to noise, which can result in a deviation of any model to the reality. Therefore, system-theoretic approaches are nondeterministic as compared to information security.

As discussed in this paper, cyber–physical system security demands additional security requirements, such as continuity of power delivery and accuracy of dynamic pricing, introduced by the physical system. Such requirements are usually closely related to the models and states of the system, which are difficult to address by information security alone. Therefore, both information security and system-theory-based security are essential to securing cyber–physical systems, offering exciting research challenges for many years to come. ■

Acknowledgment

The authors would like to thank Prof. L. Xie for interesting discussions on topics discussed in this paper.

REFERENCES

[1] E. Marris, "Upgrading the grid," *Nature*, vol. 454, pp. 570–573, 2008.

[2] S. M. Amin, "For the good of the grid," *IEEE Power Energy Mag.*, vol. 6, no. 6, pp. 48–59, Nov./Dec. 2008.

[3] NIST, *Guidelines for Smart Grid Cyber Security*, Draft NISTIR 7628, Jul. 2010.

[4] NETL, *Understanding the Benefits of the Smart Grid*, Jun. 2010.

[5] US-DOE, NERCH, *High-Impact, Low-Frequency Event Risk to the North American Bulk Power System*, Jun. 2010.

[6] J. Vijayan, "Stuxnet renews power grid security concerns," *Computerworld*, Jul. 26, 2010. [Online]. Available: http://www.computerworld.com/s/article/9179689/Stuxnet_renews_power_grid_security_concerns

[7] F. Cleveland, "Cyber security issues for advanced metering infrastructure (AMI)," in *Proc. Power Energy Soc. Gen. Meeting—Conv. Delivery Electr. Energy 21st Century*, Apr. 2008, DOI: 10.1109/PES.2008.4596535.

[8] S. M. Amin, "Securing the electricity grid," *The Bridge*, vol. 40, pp. 13–20.Spring, 2010.

[9] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security Privacy*, vol. 7, no. 3, pp. 75–77, May/Jun. 2009.

[10] Y. Liu, M. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Security*, Nov. 2009, DOI: 10.1.1.148.1133.

[11] H. Khurana, M. Hadley, N. Lu, and D. A. Frincke, "Smart-grid security issues," *IEEE Security Privacy*, vol. 8, no. 1, pp. 81–85, Jan./Feb. 2010.

[12] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Oct. 2010, pp. 226–231.

[13] NIST, *NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 1.0*, NIST Special Publication 1108, Jan. 2010.

[14] T. Alpcan and T. Basar, *Network Security: A Decision and Game Theoretic Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[15] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, Jan. 2007.

[16] R. Goebel, R. Sanfelice, and A. Teel, "Hybrid dynamical systems," *IEEE Control Syst.*, vol. 29, no. 2, pp. 28–93, 2009.

[17] EPRI, *Advanced Metering Infrastructure (AMI)*, Feb. 2007.

[18] E. L. Quinn, *Smart Metering & Privacy: Existing Law and Competing Policies*, May 2009.

[19] A. Kerckhoffs, "La cryptographie militairie," *J. Sciences Militaires*, vol. IX, pp. 5–38, 1883.

[20] D. Kundur, X. Feng, S. Liu, T. Zourntos, and K. L. Butler-Purry, "Towards a framework for cyber attack impact analysis of the electric smart grid," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Oct. 2010, pp. 244–249.

[21] S. S. S. R. Depuru, L. Wang, V. Devabhaktuni, and N. Gudi, "Smart meters for power grid—Challenges, issues, advantages and status," in *Proc. IEEE/PES Power Syst. Conf. Expo.*, 2011, DOI: 10.1109/PSCE.2011.5772451.

[22] P. Huitsing, R. Chandia, M. Papa, and S. Shenoi, "Attack taxonomies for the Modbus protocols," *Int. J. Critical Infrastructure Protection*, vol. 1, pp. 37–44, Dec. 2008.

[23] E. Barker, D. Branstad, S. Chokhani, and M. Smid, *A Framework for Designing Cryptographic Key Management Systems*, NIST DRAFT Special Publication 800-130, Jun. 2010.

[24] H. Lee, J. Kim, and W. Lee, "Resiliency of network topologies under path-based attacks," *IEICE Trans. Commun.*, vol. E89-B, pp. 2878–2884, Oct. 2006.

[25] D. Seo, H. Lee, and A. Perrig, "Secure and efficient capability-based power management in the smart grid," in *Proc. Int. Workshop Smart Grid Security Commun.*, May 2011, pp. 119–126.

[26] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, "Theory of spread spectrum communications—A tutorial," *IEEE Trans. Commun.*, vol. 30, no. 5, pp. 855–884, May 1982.

[27] S. McLaughlin, D. Podkuiko, A. Delozier, S. Mizdzvezhanka, and P. McDaniel, "Embedded firmware diversity for smart electric meters," in *Proc. USENIX Workshop Hot Topics in Security*, 2010, DOI: 10.1.1.172.8354.

[28] M. LeMay, G. Gross, C. Gunter, and S. Garg, "Unified architecture for large-scale attested metering," in *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2007, DOI: 10.1.1.93.6432.

[29] M. LeMay and C. A. Gunter, "Cumulative attestation kernels for embedded systems," in *Proc. Eur. Symp. Res. Comput. Security*, Sep. 2009, pp. 655–670.

[30] A. Seshadri, A. Perrig, L. van Doorn, and P. Khosla, "SWATT: Software-based attestation for embedded devices," in *Proc. IEEE Symp. Security Privacy*, May 2004, pp. 272–282.

[31] A. Shah, A. Perrig, and B. Sinopoli, "Mechanisms to provide integrity in SCADA and PCS devices," in *Proc. Int. Workshop Cyber-Physical Syst. Challenges Appl.*, Jun. 2008, DOI: 10.1.1.168.4847.

[32] M. Shahidehpour, F. Tinney, and Y. Fu, "Impact of security on power systems operation," *Proc. IEEE*, vol. 93, no. 11, pp. 2013–2025, Nov. 2005.

[33] W. F. Tinney and C. E. Hart, "Power flow solution by newton's method," *IEEE Trans. Power Apparat. Syst.*, vol. PAS-86, no. 11, pp. 1449–1460, Nov. 1967.

[34] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. Boca Raton, FL: CRC Press, 2004.

[35] US-DOE, *Smart Grid System Report—Characteristics of the Smart Grid*, Jul. 2009.

[36] NETL, *Characteristics of the Modern Grid*, Jul. 2008.

[37] H. Sandberg, A. Teixeira, and K. H. Johansson, "On security indices for state estimators in power networks," in *Proc. 1st Workshop Secure Control Syst.*, 2010, DOI: 10.1.1.187.4332.

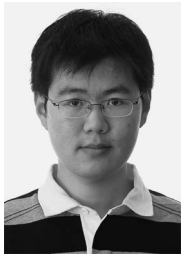
[38] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Limiting false data attacks on power system state estimation," in *Proc. 44th Annu. Conf. Inf. Sci. Syst.*, 2010, DOI: 10.1109/CISS.2010.5464816.

[39] NERC, "Technical analysis of the August 14, 2003, blackout: What happened, why, and what did we learn?" Tech. Rep., 2004.

[40] N. Falliere, L. O. Murchu, and E. Chien, "W32.stuxnet Dossier," Symantec Corporation, Tech. Rep., 2011.

[41] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *Proc. 47th Annu. Allerton Conf. Commun. Control Comput.*, 2009, pp. 911–918.

ABOUT THE AUTHORS



Yilin Mo received the Bachelor of Engineering degree from Department of Automation, Tsinghua University, Beijing, China, in 2007. He is currently working towards the Ph.D. degree at the Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA.

His research interests include secure control systems and networked control systems, with applications in sensor networks.

Tiffany Hyun-Jin Kim received the B.A. degree in computer science from University of California at Berkeley, Berkeley, in 2002 and the M.S. degree in computer science from Yale University, New Haven, CT, in 2004. She is currently working towards the Ph.D. degree at the Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA.

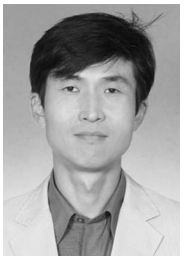
Her research covers trust management, usable security and privacy, and network security.

Kenneth Brancik completed a rigorous one year program in systems analysis at the former Grumman Data Information Systems in 1984 and an intensive two year program at Columbia University in the analysis and design of information systems in 1997. He received the M.S. degree in management and systems from New York University (NYU), New York, in 2002 and the Ph.D. degree in computing from Pace University, New York, in 2005.

He has been a leader in the technology and information assurance (IA) for over 30 years in both the public and private sectors. His professional work experience and leadership roles in IA and emerging technology have been focused primarily within the financial services sector and the federal government. His prior work affiliations have included working as the Managing Director of the Northrop Grumman Cybersecurity Research Consortium (NGCRC) and Cyber Architect, a Director and Trusted Security Advisor and consultant at VerizonBusiness security solutions group, a Manager within PricewaterhouseCoopers Advisory and Business Assurance Services sector supporting the federal government, Senior Technology Examiner at the Federal Reserve Bank of New York, a technology and safety and soundness National Bank Examiner for The United States Treasury Departments Office of the Comptroller of the Currency, a VP and Manager at CITIGROUP's Technology Project and Risk Review group and Corporate Technology Auditor at Merrill Lynch and companies World Headquarters in NYC. He is a published author with a 2008 Auerbach Publication entitled: "Insider computer fraud: An in-depth framework for detecting and defending against insider IT attacks," a coauthor of a 2010 white paper "The optimization of situational awareness for insider threat detection" in the Proceedings of the First ACM Conference on Application Security and Privacy (ACM CODASPY) in San Antonio, TX, 2011, a coauthor of a white paper in 2009 entitled "Cyber evaluation factors for full dimensional network management and control" and "The computer forensics and cyber security governance model" in *ISACA Information Systems and Control Journal* (vol. 2, 2003). He holds several well recognized professional security related certifications.

Dona Dickinson received the B.A. degree in industrial psychology from California State University, Long Beach, CA, in 1981.

She has worked in the information systems arena for over 20 years. Her experience spans the full systems life cycle. She currently provides design direction and technical oversight for information technology (IT) projects supporting commercial entities and large government agencies, with an emphasis on environment and energy. She leads the Northrop Grumman Corporation's (McLean, VA) Environment and Climate Working Group. Her past roles include Chief Technology Officer for the Postal Services and Health IT Solutions business units, and program chief architect for IRS Programs. She is a Certified Information Systems Security Professional and Northrop Grumman Information Systems Technical Fellow.



Heejo Lee received the B.S., M.S., and Ph.D. degrees in computer science and engineering from POSTECH, Pohang, Korea, in 1993, 1995, and 2000, respectively.

He is an Associate Professor at the Division of Computer and Communication Engineering, Korea University, Seoul, Korea. He was at AhnLab, Inc. as a CTO from 2001 to 2003. He was a Postdoctorate at Purdue University, West Lafayette, IN, in 2000. He was a Visiting Professor at CyLab, Carnegie Mellon University, Pittsburgh, PA, from January to December, 2010.

Adrian Perrig received the Ph.D. degree in computer science from Carnegie Mellon University, Pittsburgh, PA, in 2001.

Currently he is a Professor in Electrical and Computer Engineering, Engineering and Public Policy, and Computer Science at Carnegie Mellon University. He serves as the technical director for Carnegie Mellon's Cybersecurity Laboratory (CyLab).

Dr. Perrig is a recipient of the National Science Foundation (NSF) CAREER award in 2004, IBM faculty fellowships in 2004 and 2005, and the Sloan research fellowship in 2006.

Bruno Sinopoli received the Dr. Eng. degree from the University of Padova, Padova, Italy, in 1998 and the M.S. and Ph.D. degrees in electrical engineering from the University of California at Berkeley, Berkeley, in 2003 and 2005, respectively.

After a postdoctoral position at Stanford University, Stanford, CA, he joined the faculty at Carnegie Mellon University, Pittsburgh, PA, where he is an Assistant Professor in the Department of Electrical and Computer Engineering with courtesy appointments in Mechanical Engineering and in the Robotics Institute.

Dr. Sinopoli was awarded the 2006 Eli Jury Award for outstanding research achievement in the areas of systems, communications, control and signal processing at the University of California at Berkeley and the National Science Foundation (NSF) CAREER award in 2010. His research interests include networked embedded control systems, distributed estimation and control over wireless sensor-actuator networks, and cyber-physical systems security.