ACCEPTED FROM OPEN CALL

Detection of False Data Injection Attacks in Smart-Grid Systems

Po-Yu Chen, Shusen Yang, Julie A. McCann, Jie Lin, Xinyu Yang

ABSTRACT

Smart grids are essentially electric grids that use information and communication technology to provide reliable, efficient electricity transmission and distribution. Security and trust are of paramount importance. Among various emerging security issues, FDI attacks are one of the most substantial ones, which can significantly increase the cost of the energy distribution process. However, most current research focuses on countermeasures to FDIs for traditional power grids rather smart grid infrastructures. We propose an efficient and real-time scheme to detect FDI attacks in smart grids by exploiting spatial-temporal correlations between grid components. Through realistic simulations based on the US smart grid, we demonstrate that the proposed scheme provides an accurate and reliable solution.

INTRODUCTION

Cyber-physical systems (CPSs) embedded in the environment are used to monitor, understand behaviors of, and control, the physical world [1]. As a representative emerging CPS application, the proliferation of smart grids has been observed in our daily life. A smart grid is a relatively new type of energy distribution system that consists of the power lines as with traditional power grids, as well as information and communication technology (ICT) infrastructures connected to smart meters that may take the form of specialized devices, or laptops, cell phones, and so on. Some of these smart grid components allow information systems to perform prediction analysis, which balances the power production and consumption in the grid system. For example, real-time pricing gives both consumers and suppliers valuable indications to help manage their energy demands and supplies, respectively. Therefore, energy distribution (which controls the energy generation, consumption, and transmission processes) can be performed in a more dynamic and efficient manner [2, 3].

However, the heterogeneity, diversity, and complexity of smart-grids pose critical challenges in ensuring overall system integrity [2]. This is because in smart grids, grid-status inference and decision making may be performed on local smart devices rather than in well-protected control centers. Therefore, unlike traditional power grids where the majority of attacks and failures come from physical accesses to critical facilities [4], the ubiquity of smart-grid components further invite these abnormalities from cyber-infrastructures.

A typical attack found in smart grids is false data injection (FDI), which can be utilized to distort real energy demand and supply figures. Energy distribution may therefore be erroneous, which results in additional costs or more devastating hazards. One example was the Northeast blackout of 2003 in the USA caused by a lack of accurate real-time condition information (http://en.wikipedia.org/wiki/Northeast_blackout_of_2003).

It is imperative that we trust such systems and get the security right as national security, not just cyber-security, is at risk. However, recent countermeasures against FDIs focus more on traditional power-grid scenarios [2], in which FDI attacks are launched at physical meters rather than smart components [4]. Without considering the cyber-attacks and the distributed design of smart grid infrastructures, such approaches may not be able to provide comprehensive protection that is immediate to each local smart meter or device making decisions based on status information.

In response to this problem, we propose a lightweight approach that can be easily installed on any smart component for real-time FDI detection. Compared to traditional solutions, it identifies anomalies in *state estimations* (which are inferred from physical meter readings), as these *state estimations* are the ultimate targets for both traditional FDI attacks and cyber-attacks. Furthermore, our approach scales to large smart grid systems compared to other similar anomaly-detection mechanisms in cyber systems [5]. Results of simulations based on real-world data demonstrate the effectiveness of our proposed approach in terms of reducing extra energy transmission costs and user outage rates caused by FDIs.

SMART GRID AND ENERGY DISTRIBUTION

In a smart grid, entities are connected to the grid via ICT infrastructures and power lines. Energy is transmitted from energy-rich to ener-

Po-Yu. Chen, Shusen Yang (corresponding author), and Julie A. McCann are with Imperial College London.

Ji Lin and Xinyu Yang are with Xi'an Jiaotong University.

0163-6804/15/$25.00 © 2015 IEEE

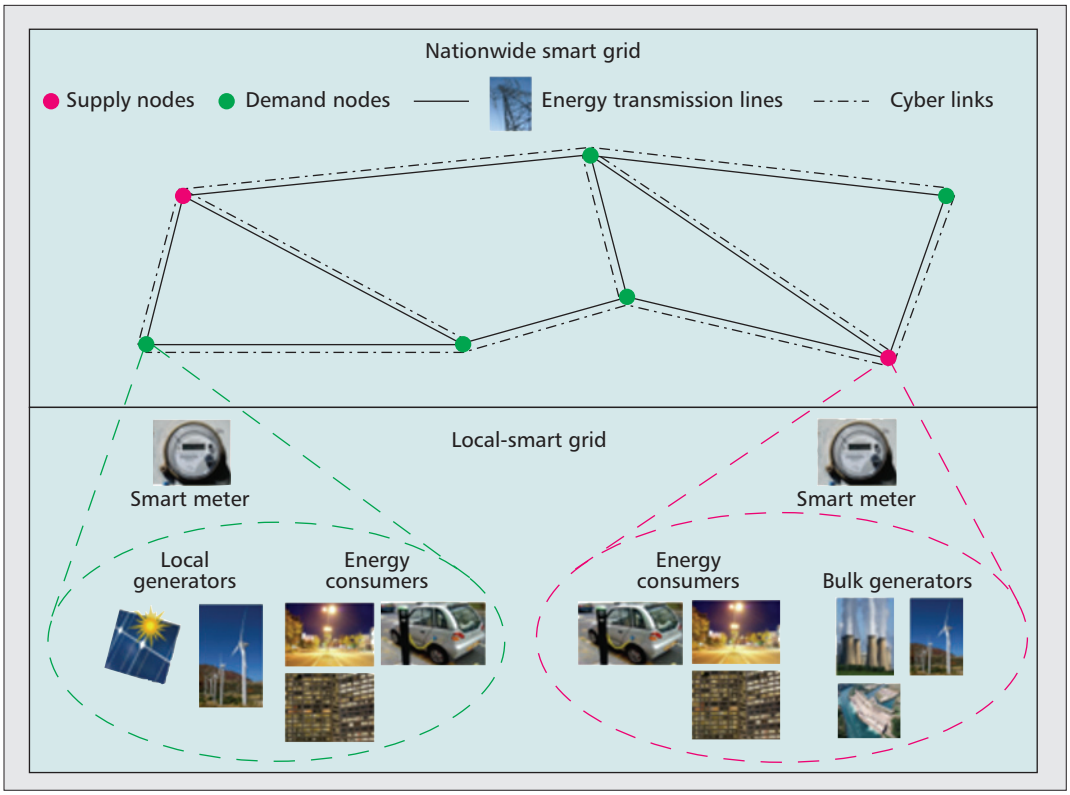

Figure 1. A conceptual illustration of cyber-physical topologies in a smart grid system.

In smart grids, these components are typically connections to the local networks or the Internet. Therefore, as the number of access points increase, compared with traditional power grids, smart grids are more prone to cyber-attacks.

gy-poor entities as per energy-distribution mechanisms. Figure 1 illustrates an example of a two-layer topology of a smart grid.

The upper layer represents the nationwide backbone power transmission grids and ICT infrastructures, where each *node* is an abstraction that represents the average energy demand/ supply of a local area (e.g. a town). Due to the difference in local energy needs and heterogeneous efficiencies of energy generation, each node in the grid will generate and consume different amounts of energy. When a node consumes more energy than it generates, the node is denoted as an *energy-demand node*, which needs to pull energy from the grid; similarly, when a node consumes less energy than it generates, the node is denoted as an *energy-supply node*, which pushes residual energy into the grid.

At the lower layer, smart meters are installed in every entity (e.g. house, factory, and school) that contains both power consuming facilities and energy producing equipment. Different from the traditional power meters that gather *raw measurements* (e.g. voltage and electrical current), smart meters are capable of further inferring *state estimations* (e.g. energy demands/ supplies) and make preliminary decisions (e.g. data fusion) before these estimations reach control centers. With the information obtained from smart meters, energy distribution can be optimized with regard to various power grid performance metrics. For example, they may maximize the network utility and energy usage efficiency, while minimizing energy transmission costs and user outage numbers [6].

FDI Attacks in Smart Grids

Although smart components such as smart meters play an important role in smart grids, they also increase the *vulnerability* of the grid system. In smart grids these components are typically connections to the local networks or the Internet. Therefore, as the number of access points increase, compared with traditional power grids, smart grids are more prone to cyber-attacks [2, 6]. Among these attacks, the most critical attack is false data injection (FDI) [2]. Adversaries can launch these attacks to compromise *raw measurements* or *state estimations* [4, 6].

These FDI attacks have been found to be launched by *consumers* and *adversaries* on purpose [2]. Consumers typically cheat the system by manipulating their energy consumption to reduce their energy bills. Adversaries, who can be opponent companies or countries, would aim to compromise *state estimation* (as described previously) to increase the cost of energy distribution and smart-grid operations. Considering the degree of these effects, the latter attacks are more critical for smart grids. False energy demand and supply can disrupt the energy distribution process, which may result in significant financial loss and even devastating outcomes.

FDI Attack Targets

The four most critical *state estimations* and the potential FDI attack targets related to these are listed as follows.

Energy Demand: The untruthful values of this state estimation are critical as they can dramatically increase financial costs to both the energy users and providers due to the extra cost

Detection is the most essential step in minimizing the damages resulting from the aforementioned FDI attacks. Therefore, the efficiency and effectiveness of FDI detection techniques have significant impacts on the overall performance of smart grid systems.

of power transmission and the waste of energy that can occur. Moreover, they can result in an even more devastating crisis, i.e. a power outage, where energy requests to the smart grid are less than the energy demand that nodes truly require. Since these nodes consist of energy consumers, such as personal users or companies, the adversary may launch such attacks either through infected personal devices such as laptops, or through poorly configured firewalls [2].

Energy Supply: The value of this state estimation is mainly provided by energy-supply nodes. FDIs can falsely decrease the advertised quantity of their supplied energy, which typically results in the starvation of energy-demand nodes who are not receiving what they request. Reversely, a falsely advertised increase in energy supply can incur an increase of wasted energy. These attacks can be achieved using malware to infect the servers of power suppliers, falsifying the quantity of energy it can truly provide.

Grid-Network States: The grid-network states are the configurations and conditions of the power grids, such as grid topologies and power-line capacities. For instance, adversaries can isolate nodes from the power grid by invalidating their power-line connection. Maliciously forging the network states can seriously mislead the energy distribution, again resulting in devastating power shortages or extra energy transmission costs.

Electricity Pricing: Dynamic electricity pricing can greatly reduce the electricity bill of consumers, such as energy-utility companies and houses, as well as help balance the power loads between peak and off-peak periods. Therefore, fake electricity pricing would incur significant damage to both the financial and physical subsystems that make up a smart grid, negating their advantages of optimum supply efficiencies. For instance, adversaries can falsely reduce prices during peak hours, which will ultimately result in grid system overload. Also, illegal users may use malware to modify their energy bills, leading to a loss of utility company revenue.

Detecting FDI Attacks

Detection is the most essential step in minimizing the damages resulting from the aforementioned FDI attacks. Therefore, the efficiency and effectiveness of FDI detection techniques have a significant impact on the overall performance of smart grid systems.

Traditional Solutions for FDI Detection

Since the smart grid is a relatively new topic compared to traditional power-grid systems, the main research on countermeasures against FDIs are mainly system-theoretic approaches [4, 7, 8]. In these approaches, the relationship between raw meter measurements and system states are simply described as $\vec{\mathbf{z}} = \mathbf{H}\vec{\mathbf{x}} + \vec{\mathbf{e}}$, where the matrix $\mathbf{H}$ represents system configuration, and vectors $\vec{\mathbf{z}}$, $\vec{\mathbf{x}}$, and $\vec{\mathbf{e}}$ represent the meter measurements, state estimations, and meter measurement errors, respectively. Therefore, for a predicted state $\hat{\mathbf{x}}$, abnormal values can be detected by computing the 2-Norm of its measurement residual $||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}||$. $\mathbf{z}$ will be considered to be abnormal, if $||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}||$ is larger than a given threshold τ.

Recent research [4] demonstrates that this traditional countermeasure approach can be easily bypassed if adversaries are capable of accessing enough meters and inject malicious data to the grid with specific values. Also, this traditional solution can only detect abnormalities at the group level (each of which contains several meters). Abnormalities in each individual meter are unobservable in this solution. Although some approaches try to solve this problem, such as [7], they are unable to guarantee the security of smart-grid systems, since they are limited to traditional power grid scenarios, where adversaries can only assess physical meters. Therefore, attacks through cyber infrastructures can directly compromise *state estimations* without being detected by the detection approaches in traditional power-grid systems [2].

Anomaly Detection for FDI Attacks

Since FDI attacks would result in abnormal state estimations, we can apply anomaly detection techniques [5] to smart grids, which has been studied in various areas such as cyber intrusion, sensor networks, and image processing [5]. In current research on intrusion detection, rule-based mechanisms are the simplest solutions. They typically exploit static thresholds to identify anomalies. When the value of a raw data input exceeds these thresholds, this value is regarded as anomalous. These schemes introduce very low overhead to the system, but they fail to adapt to valid changes in the environment. Furthermore, a high level of professionalism is required to manually maintain detection thresholds.

To cope with the dynamics and heterogeneities in environments, machine-learning-based solutions are more often exploited [5]. Statistical models (e.g. Bayesian, k-nearest neighbor) and artificial intelligence (e.g. neural networks) are typical approaches utilized to distinguish anomalies from norms. However, these solutions are usually complex, which limits their scalability when applied to large complicated networks, such as smart grids. Therefore, to achieve a more balanced solution, further assumptions are usually made to reduce computational complexity. What we exploit here the most is one simple general assumption: *spatiotemporal correlation*.

Spatio and Temporal Correlations in Smart Grids

Spatiotemporal correlation is a natural property found in various physical phenomena including human behaviors, since they are typically continuous over both the time and spatial domains. Since cyber states are mappings to these phenomena, the estimations of these states within a correlated sphere should also be spatiotemporally correlated. In general, the following correlations exist in smart grid systems.

Traditional power suppliers usually generate energy based on the current needs of power consumers: Typically, to reduce the extra overhead caused by energy storage, traditional power suppliers (such as fossil-fuel power stations) are inclined to dynamically adapt their power gener-

ation to the requirements of consumers. Therefore, these power suppliers, in the same area, can be spatially temporally correlated as they should all reflect the current state of power consumption nearby.

Renewable-energy suppliers are typically correlated to nearby suppliers: As renewable energy comes from natural resources (such as solar and wind power), energy generators, for example, wind farms and hydroelectricity plants, deployed in a nearby area should be able to simultaneously reveal the current state of the environment therein. That is, they are spatially correlated. Also, these spatial correlations should be similar to those that have occurred in the past, that is, these energy supplies are temporally correlated, as natural resources are continuous and should therefore result in interrelated effects to these power suppliers.

Energy consumers within the same area should behave in similar patterns: In general, power consumption in a related area should simultaneously reflect the current state of this region, such as weather, activities, and government policies. For example, in winter, power consumers of nearby cities tend to expend more power supplies than in summer due the use of heaters, even when each individual has its own patterns regarding power consumption. Therefore, the total node behavior can still be consistent and correlated to other consumer nodes at higher abstractions (such as districts of a city).

Real-Time Correlation-Based FDI Detection

In order to minimize the damage of FDIs in smart grid systems, we propose an effective real-time mechanism to detect FDI attacks against state estimation in smart grid systems, based on the spatiotemporal cyber-state correlations discussed previously. We call strongly spatiotemporally correlated smart components as neighbors. Potential anomalies can be detected by monitoring the temporal consistencies of the spatial correlations between state estimations. This detection mechanism can be divided into the following three phases, discussed below.

Spatial-Pattern Recognition and Temporal-Pattern-Consistencies Evaluation

This phase models the current states of the smart grid as its normal-state references (i.e. the state without FDI attacks). It is a semi-supervised process, similar to many other machine-learning mechanisms [5]. In our approach, correlation patterns are recognized between each pair of state estimations within the same correlation sphere defined by smart-grid operators, rather than the whole network. A correlation sphere typically contains several smart meters, while a smart meter may simultaneously belong to multiple correlation spheres. The defining of correlation spheres can guarantee that all the smart components are spatially and temporally correlated, which is required by our approach. This design also avoids the high computational complexity when analyzing data over high-dimensional spaces [5].

We assume that genuine changes and errors in state estimations follow Gaussian distributions, which is a common assumption [5]. Let $s_i(t)$ denote the state estimation of a smart component i (e.g. a photoelectric energy harvester) at current time t. The sequence of previous T state estimations for a smart component i before current time t can be represented as

$$S_i(t, T) = (s_i(t - T), s_i(t - T + 1), \cdots, s_i(t - 1)) \quad (1)$$

We consider each pair of smart components i and j in a correlation sphere G, such as a set of correlated photoelectric-energy harvesters in a *town*. Based on their previous state estimations $S_i(t, T)$ and $S_j(t, T)$, we can compute their spatial-correlation consistency region (SCCR), which represents the set of all possible potentially correct estimation pairs of $s_i(t)$ and $s_j(t)$ at current time t. If the current estimation pair $(s_i(t), s_j(t))$ belongs to the set defined by the SCCR, we say the current estimations $s_i(t)$ and $s_j(t)$ are *consistent*; otherwise, they are *inconsistent*.

Geometrically, SCCR can be approximated by a rotated ellipse, as illustrated in Fig. 2, Here, the center of the SCCR ellipse, $(\overline{s}_i(t), \overline{s}_j(t))$, is computed by using the exponential weighted moving average (EWMA) for the previous estimations $S_i(t, T)$ and $S_j(t, T)$. The major and minor axes of the SCCR ellipse are computed by using singular value decomposition (SVD), based on the previous estimations $S_i(t, T)$ and $S_j(t, T)$. In our approach, SVD is a mathematical procedure to convert observations of multiple observers into two orthogonal principal components $\vec{a}$ and $\vec{b}$ (which define the rotation angle θ) and their associated variances σ_a^2 and σ_b^2. We set the lengths of these axes as the three deviations $3\sigma_a$ and $3\sigma_b$, which cover 99.46 percent normal observations, respectively.

Trust-Based Voting

After the correlation mapping consistency between each pair of state estimations is obtained, trust-based voting is exploited to identify the *occurrences* of anomalies. This voting is divided into two different rounds. In the first round, reliable state estimations are selected with the votes from the state estimations of their correlation neighbors. For a smart component i in a correlation sphere G, define its correlation neighbor set $N_i \subset G$ as the set of all devices in G excluding i; and define its consistent neighbor set $N_i^c \subseteq N_i$ as a set of devices, where the state estimation $s_j(t)$ of each device j in N_i^c are *consistent* with the state estimation $s_i(t)$. Take Fig. 3 for instance: $N_A = \{B, C, D, E\}$, $N_E = \{A, B, C, D\}$, $N_A^c = \{B, C, D, E\}$, and $N_E^c = \{A\}$.

Let $|N_i|$ and $|N_i^c|$ represent the sizes of sets N_i and N_i^c, respectively. We say that component i and its estimation $s_i(t)$ are *likely* to be *reliable* (LR) at time t, if $|N_i^c|/|N_i| \geq 50$ percent. Otherwise, we say that the smart component i and its estimation $s_i(t)$ are *likely* to be *anomalous* (LA) at time t. After the first round voting, smart component E and its current state estimation are LA while others are LR, as shown in Fig. 3.

Spatiotemporal correlation is a natural property found in various physical phenomena including human behaviors, since they are typically continuous over both the time and spatial domains. Since cyber states are mappings to these phenomena, the estimations of these states within a correlated sphere should also be spatiotemporally correlated.

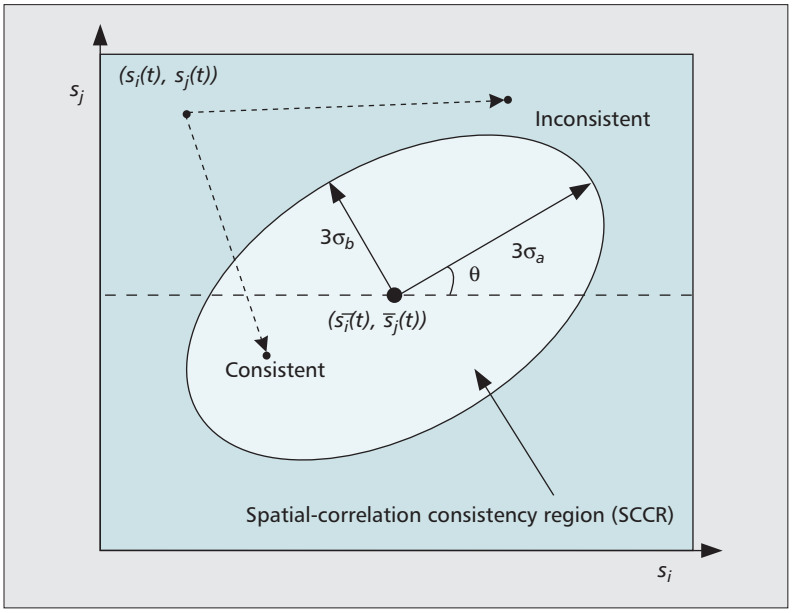

Figure 2. Illustration of a spatial correlation consistency region.

After the first round voting, only LR components and their state estimations are qualified to be involved in the second round voting, such as A, B, C, D in Fig. 3. For a smart component i, define $N_i^{LR} \subseteq N_i$ as the set of all LR components in the correlation sphere G excluding i; and $N_i^R \subseteq N_i^{LR}$ as the set of all LR components that are *consistent* with i. For instance, $N_E^{LR} = \{A, B, C, D\}$ and $N_E^R = \{A\}$, shown in Fig. 3. The second round voting for state estimation $s_i(t)$ has one of the following three outcomes:

- **Good**, if $|N_i^R|/|N_i^{LR}| \geq 50$ percent and $|N_i^{LR}|/|N_i| \geq 50$ percent.
- **Abnormal**, if $|N_i^R|/|N_i^{LR}| < 50$ percent and $|N_i^{LR}|/|N_i| \geq 50$ percent.
- **Unknown**, otherwise.

It can be seen that we can determine whether $s_i(t)$ is **Good** or **Abnormal**, if the majority of i's neighbors are labelled as LR at slot t; otherwise, the reliability of $s_i(t)$ is unable to be determined (i.e. **Unknown**), due to the lack of reliable references.

System Condition Inference

We further infer the phenomena that the three outcomes represent. First, **Good** indicates that current state estimation $s_i(t)$ is reliable, since it is highly correlated to other state estimations that are in the same correlation sphere G. This usually suggests that the state estimation is in a stable state and probably normal. On the contrary, **Abnormal** indicates the value of $s_i(t)$ deviates from the majority of state estimations in the same correlation group G. This typically suggests an abnormality in $s_i(t)$, which can either be due to failures or an FDI attack. Finally, the **Unknown** occurs when there are not enough state estimations that can be regarded as reliable references when we compute the reliabilities. This happens when there is an extensive change among the state estimations in G due to genuine occasions that dramatically change the behavior of the smart grid or large-scale attacks, which should be regarded as an indispensable event that requires further inspection from system administrators.

Practical Issues

Except for assuming spatiotemporal relationships and that errors follow Gaussian distributions, we do not require any further assumptions. Therefore, the proposed FDI detection mechanism is quite general and can be used for heterogeneous types of correlated components in the smart grid systems. When we apply the proposed detection mechanism to practical smart grid systems, the window size of previous state estimations, T, should be considered. This is a user defined parameter, which decides the computation of the SCCR (i.e. the ellipse shown in Fig. 2). Consequently, users should carefully assign T according to their practical scenarios to assure that the entire state estimations that lie in this interval are temporally correlated, while avoiding unnecessary extra storage and computational overheads.

A Case Study

To demonstrate how FDI attacks would impact the energy distribution of a smart grid system, we conducted an evaluation based on a simplified version of the US smart grid consisting of 48 states, as shown in Fig. 4 (http://www.oe.energy.gov/smartgrid.htm.). Furthermore, to simulate local energy generation and consumption of each state, we decomposed each state into 10 energy suppliers and 10 energy consumers. Each of these suppliers/consumers contains 365 daily energy generation/consumption profiles, which is correlated to other suppliers/consumers in the same state, to simulate the annual behavior of the smart-grid system. All these profiles are based on the 2009 US Energy Information Administration State Electricity Profiles (available on http://www.eia.gov/).

In this simulation, we study two different types of FDI attacks. The first FDI attack aims to increase the *total energy transmission cost* [4]. In this scenario, FDI attacks were randomly inserted to increase the advertised supply-demand values of suppliers and consumers by 15 percent and then 30 percent , that is, 0.15× false and 0.3× false shown in Fig. 5a. The total energy transmission cost $Cost_{total}$ is defined as

$$Cost_{total} = \Sigma_{L_{ij} \in \boldsymbol{L}} E_{ij} \cdot Cost_{ij}, \quad (2)$$

where $\boldsymbol{L}$ denotes the set of all power lines in a given smart-grid system; L_{ij} denotes the power line between supplier i and consumer j; E_{ij} denotes the amount of energy transmitted through power line L_{ij}; and $Cost_{ij}$ denotes the cost of power transmission per unit through power line L_{ij}.

The second FDI attack aims to promote significant power-supply outage. The demands of energy consumers were falsely increased by 1 and 2 times, that is, 1× false and 2× false shown in Fig. 5b, while the advertised supplies were falsely decreased by 10 percent and 20 percent. The user outage rate R_{out} is defined as:

$$R_{out} = \frac{total\ number\ of\ outage\ nodes}{total\ number\ of\ demanding\ nodes}$$

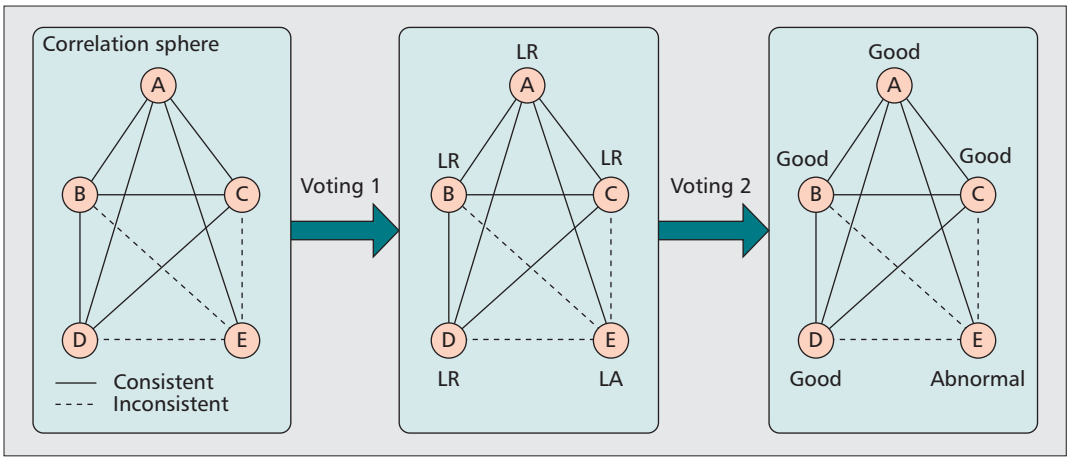

Figure 3. Visualization of a trust-based voting example.

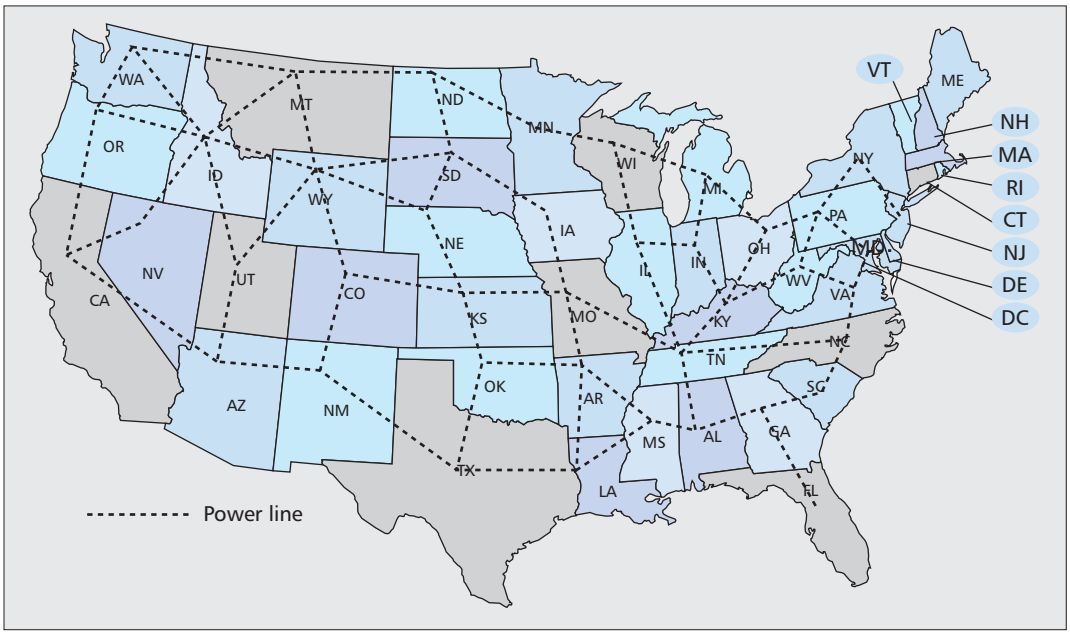

Figure 4. A simplified US smart grid topology.

Figure 5 shows the simulation results when applying our FDI-detection approach and the state-of-the-art distanced-based (DB) solution proposed in 2013 [9, 10]. Both approaches referenced the past 30 days data during evaluation (i.e. $T = 30$). Also, we set the user-defined parameters of the distanced-based solution as $r = 0.5$ and $D = 0.5$ [9]. Note that since there were no false-positive detections caused by our approach in the cost-increasing scenario, it cannot be plotted under log scale in Fig. 5c.

As can be seen in Figs. 5a and 5b, both FDI-detection mechanisms can significantly reduce the adverse impacts caused by FDIs. In the cost-increasing experiment, our solution reduced up to 1.14 million US dollars without any false-positive detection. Furthermore, our approach can achieve about 55.5 percent user outage reduction with only 0.43 percent false-positives.

In contrast, although the DB-detection scheme achieves comparable performance in detection FDIs, it suffers from extremely high false-positive detection rates, which means that it misreports correct state estimations when FDI attacks with a high probability. As can be observed in Fig. 5c, false-positive detections are increasing as the FDI-attack rate increases for both schemes; however, compared with our approach, DB detection incurs nearly two-orders of magnitude more false-positive results. Such a high false-positive detection rate of the DB-detection scheme would result in significant confusion in decision making and would incur additional costs, such as extra labour to understand the problem where the integrity of the information is required to be checked manually, that is, interpreted by human beings.

Furthermore, as shown in Fig. 5d, the computational overhead of our approach is extremely low, that is, nearly two-orders of magnitude less than that of DB-detection. This demonstrates that our approach has a great potential to be applied for real-time FDI attack detection in smart grid systems.

Except for assuming spatiotemporal relationships and that errors follow Gaussian distributions, we do not require any further assumptions. Therefore, the proposed FDI detection mechanism is quite general and can be used for heterogeneous types of correlated components in the smart grid systems.

Conclusion

Compared to traditional power grids, smart grids are predicted to be more reliable and effective energy-distribution solutions that can cope with complicated power supply and demand scenarios. However, they have the potential to be more vulnerable to cyber-attacks such as from malware or malicious cyber-intrusions. Among all these attacks, the most critical ones are false-data injection (FDI). Adversaries can launch these attacks to compromise *raw meter measurements* or *state estimations*. This can significantly disrupt the effectiveness of the energy distribution in smart grids, which results in additional costs or some hazards that would have larger impacts, such as blackouts or enormous costs.

Recent approaches to identifying FDIs are less

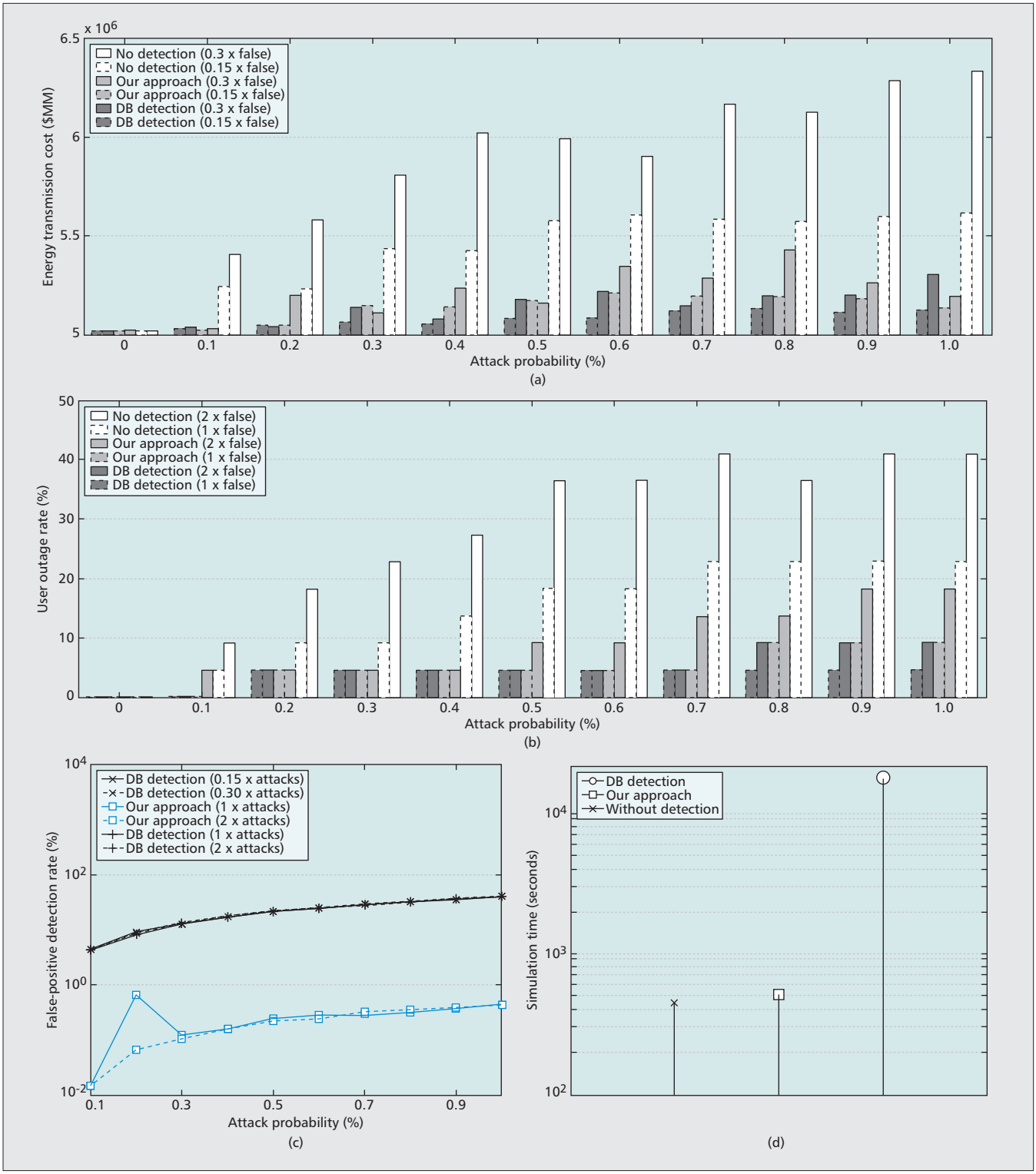

Figure 5. The correlation between (a) energy transmission cost (b) user outage rate (c) false-positive detection rate, and attack probability for different attack criteria; and (d) required simulation time.

effective, as they assume the technical architecture of traditional power grids. In this article we propose a simple and robust solution to detecting false data injection in smart grid systems that consist of many differing ICT components. This solution exploits spatiotemporal cyber-state correlations and trust-based voting to evaluate the reliabilities of *state estimations*. FDIs can be detected once those unreliable state estimations are identified. A case study is presented to demonstrate the impacts of FDIs with our approach versus the nearest state of the art approach. The simulation results show that the proposed solution can effectively and efficiently detect malicious FDIs and prevent potential extra costs and security threats.

Even with our solution, smart-grid systems remain vulnerable to some FDIs such as those that slowly evolve to prevent detection, as well as other sorts of attacks both physical and computational. Therefore, more powerful countermeasures are required as these attacks can become potential threats to national security and citizen wellbeing.

Acknowledgments

This work is sponsored by the Intel Collaborative Research Institute (ICRI) for Sustainable Connected Cites, and the Natural Science Foundation of China (NSFC), 61373115 and 61402356.

References

[1] E. A. Lee, "Cyber-Physical Systems—Are Computing Foundations Adequate," position paper for NSF Workshop on Cyber-Physical Systems: Research Motivation, Techniques and Roadmap, vol. 2, 2006.

[2] M. Yilin, *et al.*, "Cyber-Physical Security of a Smart Grid Infrastructure," *Proc. IEEE*, vol. 100, 2012, pp. 195–209.

[3] Y. Liu, P. Ning, and M. K. Reiter, "False Data Injection Attacks Against State Estimation in Electric Power Grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, 2011, pp. 1–33.

[4] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," *ACM Comput. Surv.*, vol. 41, 2009, pp. 1–58.

[5] M. Baghaie, S. Moeller, and B. Krishnamachari, "Energy Routing on the Future Grid: A Stochastic Network Optimization Approach," *Proc. IEEE POWERCON*, 2010, pp. 1–8.

[6] L. Jie *et al.*, "On False Data Injection Attacks Against Distributed Energy Routing in Smart Grid," *Proc. IEEE/ACM ICCPS*, 2012, pp. 183–92.

[7] R. B. Bobba *et al.*, "Detecting False Data Injection Attacks on DC State Estimation," preprints of the First Workshop on Secure Control Systems, CPSWEEK, 2010.

[8] O. Kosut *et al.*, "Malicious Data Attacks on the Smart Grid," *IEEE Trans. Smart Grid*, vol. 2, 2011, pp. 645–58.

[9] Y. Lei and L. Fengjun, "Detecting False Data Injection in Smart Grid In-Network Aggregation," *Proc. IEEE SmartGridComm*, 2013, pp. 408–13.

[10] S. Subramaniam *et al.*, "Online Outlier Detection in Sensor Data Using Non-Parametric Models," *Proc. VLDB*, 2006, pp. 187–98.

Biographies

Po-Yu Chen received a B.S. degree from National Cheng Kung University, Tainan, Taiwan, and an M.S. degree from National Taiwan University, Taipei, Taiwan. He is currently working toward a Ph.D. degree in the Department of Computing, Imperial College London, London, U.K. His research interest is sensor-data mining, including data-correlation discovery and anomaly detection, in wireless networks, networked sensing systems, and distributed systems.

Shusen Yang received a Ph.D. degree from Imperial College London, London, U.K., in 2013. He is currently a research associate with Imperial College London, where he is also with the Intel Collaborative Research Institute for Sustainable Connected Cities. His research interests are wireless networks, networked sensing systems, and distributed systems, including stochastic network optimization, energy harvesting networks, mobile crowd sensing, in-network processing, and network reliability. He is a member of the Association for Computing Machinery.

Julie A. McCann is a professor of computer systems at Imperial College London, where she leads the Adaptive Embedded Systems Engineering Research Group and the Intel Collaborative Research Institute for Sustainable Cities, and works with NEC and others on substantive smart city projects. Her research centers on highly decentralized and self-organizing scalable embedded systems. She is a Fellow of the British Computer Society.

Jie Lin received his Ph.D. degree from the Department of Computer Science and Technology at Xi'an Jiaotong University in 2013. He is currently a lecturer in the Department of Computer Science and Technology at the Xi'an Jiaotong University. His main research interests include wireless sensor networks security and cyber space security in cyber-physical systems.

Xinyu Yang received a Ph.D. degree in computer science from Xi'an Jiaotong University in 2001. He is currently a professor in the Department of Computer Science and Technology at the Xi'an Jiaotong University. His main research interests are in the areas of wireless communication, mobile ad hoc networks, network security, and cyber-physical systems security.

Even with our solution, smart-grid systems remain vulnerable to some FDIs such as those that slowly evolve to prevent detection, as well as other sorts of attacks both physical and computational. Therefore, more powerful countermeasures are required as these attacks can become potential threats to national security and citizen wellbeing.