Chapter 6

DISTRIBUTED BIAS DETECTION IN CYBER-PHYSICAL SYSTEMS

Simon Thougaard and Bruce McMillin

Abstract An attacker can effectively publish false measurements in distributed cyber-physical systems with noisy measurements. These biased false measurements can be impossible to distinguish from noise and enable the attacker to gain a small but persistent economic advantage. The residual sum, a fundamental measurement of bias in cyber-physical systems, is employed to develop a detection scheme for bias attacks. The scheme is highly efficient, privacy preserving and effectively detects bias attacks.

Keywords: Cyber-physical systems, security, privacy, bias attacks, smart grid

1. Introduction

False data injection attacks on power systems have been the subject of intense study since they were introduced by Liu et al. [4]. The attack model assumes an attacker who knows the power system configuration and has the ability to send corrupted measurements to a control entity (i.e., bad data injection). Liu and colleagues have also shown that such attacks can be undetectable by standard methods.

False data injection attacks pose a fundamental challenge to cyber-physical systems: if a node in a cyber-physical system is compromised by an attacker and the attacker knows what security measures are in place, the attacker can always inject bad data into a control system while avoiding detection. This chapter proves this result for any cyber-physical system that is tolerant to measurement error.

The effectiveness and limitations of false data injection attacks were discussed in the original paper by Liu et al. [4]. Subsequent papers have proposed defense schemes and variations of false data injection attacks. However, some proposed defense schemes suffer from a simple

lack of imagination. The question is, if an attacker knows the defense schemes, can the attacker still circumvent them? After all, Liu et al. [4] assumed that the attacker knows the system configuration and the bad data thresholds.

This research approaches the problem in a more general manner. Sound defense schemes result from specific criteria. The proposed scheme confronts economic attacks specifically and meets the relevant criteria. A novel queue-based approach to attack detection is employed to optimally trade-off the false positive and false negative rates. Attacks on the smart grid are considered. Conventional state estimation assumes the presence of a central system operator who may be able to counteract an attack if it is properly identified. Under a distributed electric grid architecture, a centralized entity may still exist, but it is relevant to consider privacy issues as well as the practical applicability of any defense scheme. The literature on false data injection attacks represents the systems as matrices of data, but for any individual node, only a slice of the data is available.

2. Related Work

Liu et al. [4] introduced the concept of false data injection attacks in power system state estimation and proved the existence of zero-residual attacks. This chapter does not propose a solution, but derives expressions for optimal attack vectors under different conditions. While the zero-residual attack is the most impressive version of a false data injection attack, it is not considered in this work. Zero-residual attacks can be considered to be unsolvable as they result from an attacker with complete power to arbitrarily inject bad data. However, good system design may make such attacks difficult to conduct. An attacker has to compromise every measurement related to a state variable – this is comparable to the attacker purchasing a bank in order to get access to the vault. It is theoretically possible, but perhaps not a practical security concern. Therefore, the focus is on attacks whose residuals are below a tolerable threshold.

Liang et al. [3] have conducted a thorough review of the literature on false data injection attacks on power systems. Much of the work focuses on variations of the attacks and systems under attack, as well as defense schemes. However, preference is given to attack scenarios that are easy to define mathematically instead of attack scenarios that capture the rational behavior of attackers. This stems no doubt from the academic norms of the control theory and power systems communities. As a result, an attacker who simply behaves in a sub-optimal or nonconforming

manner may go perfectly undetected by detection schemes that assume optimal or conforming behavior.

Xie et al. [8] have introduced economic attacks on electricity markets that leverage false data injection. Jia et al. [2] have proposed a solution for detecting such attacks. Economic attacks differ from most attacks in the literature by considering attackers with known and quantifiable goals. These attacks may be designed to go undetected at the expense of the magnitudes of the attacks. If an entity gains a small but reliable amount of value from a persistent subtle attack, the entity would want to conduct the attack for as long as possible. However, it is not critical to quickly uncover such attacks because they pose no security threats. It would be sufficient to guarantee eventual detection.

The smart grid presents many new opportunities and challenges given the ability of nodes to coordinate the production, consumption and distribution of electricity. Mengelkamp et al. [5] present a comprehensive approach for implementing such coordination in a decentralized manner. Molina-Markham et al. [6] discuss privacy attacks on smart meters. Specifically, how smart meter data may be used to infer private information about a home. The work is a reminder of how secondary data sources can reveal private sensitive information.

3. Problem Definition

This chapter considers false data injection attacks on power systems. In power system control, a central operator collects system measurements and decides whether or not to take actions. In a decentralized grid, these activities may be carried out in a distributed manner, where local nodes collaborate on decision making and control.

Since measurements may have errors, state estimation is employed to compute the most probable real state of the system. State estimation relies on the relations between states. This can be expressed using the model:

$$z = h(x) + e \tag{1}$$

where z is a measurement, x is a system state, $h(x)$ is a function relating states to measurements and e is the error.

State estimation is the problem of estimating the state $\hat{x}$ from z when there are multiple interrelated states. The state estimate SE is expressed as a function of the measurement z:

$$\hat{x} = SE(z) \tag{2}$$

The computation of state estimates is inconsequential to the rest of this chapter. It suffices to say, that given noisy measurements, estimates

of the system state can be obtained. This is especially useful in the smart grid where the coordination between nodes may be distributed and traditional state estimation does not apply.

Bad measurement detection is the problem of determining if measurements are abnormal or anomalous. It is computed using the residual r:

$$r = z - \hat{x} \tag{3}$$

which is simply the difference between the measurement and the estimate. The residual is tested against a threshold value $r \leq \tau$ to determine if the measurement is credible or not.

3.1 Attack Model

In a traditional false data injection attack, the attacker is assumed to have control over one or several measurements and have detailed knowledge of the system. The attack is modeled by:

$$z_a = h(x) + e + \alpha \tag{4}$$

where α is a non-zero value. The assumption is that an undetected value of α will yield a gain to the attacker whereas $-\alpha$ will yield a loss.

The goal of the attack has two parts: (i) avoid any detection scheme deployed by the controller; and (ii) effect some changes to the estimated states.

Avoiding detection is a matter of keeping the residual r below some threshold. Liu et al. [4] have proved the existence of zero-residual attacks that change $\hat{x}$ without changing r. These attacks are only possible if an attacker controls every measurement related to some state. Liu and colleagues have also identified another type of attack where r may change, but is kept under a threshold τ. These attacks have less impact on $\hat{x}$, but can be carried out with just one corrupt measurement. This work only considers the latter type of attack where there is some change to the residual.

Thus, the goal of the attacker is to maximize the change in $\hat{x}$ while satisfying $r \leq \tau$. Given that the attacker knows how $SE(z)$ is computed and how τ is set, it is not very difficult to determine the optimal α value to be injected.

This work assumes that an attacker attempts to inject a consistent, yet unobtrusive, bias. The bias may be relatively small compared with the measurement variance. The attacker attempts to "hide in the error," so to speak, by keeping the attack residual too small to be distinguishable from the measurement error, but consistently in a direction that benefits

the attacker. The benefit could be simple energy theft or overcharging for the amount of supplied energy.

3.2 Graph Approach

State estimation is traditionally considered to be an optimization problem where a power system is modeled as a system of linear equations. The collected data is represented as a vector of measurements whose relations are expressed by a matrix.

This work engages a graph model of the system instead of the standard form involving a system of linear equations. Because an optimization problem is not considered, there are no benefits to using the standard form. Additionally, in a distributed system such as a smart grid, control of the system may be distributed; long delays and response times in such a system render the collection of all the measurements for analysis problematic. Privacy may also be a concern.

Let $G = (V, E)$ be a graph representation of a smart grid where V is the set of nodes representing endpoints in the smart grid and E is the set of edges representing transmission lines between two nodes. Let $N(v)$ denote the set of neighbors of v.

The problem definition becomes simpler under the assumption that every node has an attached battery that stores or releases energy at will. This is not a restriction because a node without a battery is equivalent to a node that chooses never to use its battery. Each node can measure and report its incoming power P_i^+ from each neighbor, outgoing power P_j^- to each neighbor and its battery storage and discharge, P_b^+ and P_b^-, respectively. Each node operates under the conservation of power constraint:

$$\sum_{v_i \in N(v)} P_i^+ + \sum_{v_i \in N(v)} P_i^- + P_b^+ + P_b^- = 0 \tag{5}$$

Each node is required to report incoming and outgoing power readings to each neighbor. To preserve privacy, these readings are not reported to a central operator. Note that, even if a node reports all the measurements to an observer, it is trivial for the node to appear to be perfectly consistent internally even if it reports false measurements. If the node underreports its incoming power by a, it simply has to subtract a from P_b^+ to satisfy the equality.

In order to model economic attacks where an attacker seeks to gain some advantage from false data injection, it is useful to simplify the model further. Consider the case where two nodes, v_1 and v_2, share a transmission line, and v_1 intends to launch a false data attack against

v_2. There are two cases to consider. The first case is that the attack is strictly directed at v_2, which corresponds to a two-node problem. In second case, the attack is directed at multiple nodes, but it would still produce some attack residual between v_1 and v_2. In a situation involving faulty hardware, it may be reasonable to assume that a bad node provides bad data to all its neighbor nodes. However, an intelligent attacker may decide to provide bad data to only one neighbor or a select set of neighbors. Hence, it makes sense to reduce attacks to two-node problems. This may not be the optimal way to detect all attacks because some attacks may be detected faster by considering multiple residuals. But setting optimality aside, it makes for a much simpler problem definition:

- Given a set of edges E in a system graph $G = (V, E)$, determine the edges that are likely to be under attack.

Note the emphasis is on edges not nodes, which differs from the standard problem formulation that emphasizes nodes or state variables. This makes the problem explicitly about relations between nodes and not the nodes themselves. Residual sums represent relations between nodes; attacks, when they are detected, reside in the relations between nodes.

3.3 Distinguishing Victim from Attacker

The problem formulation only mentions determining an attacked edge, not the node that may be the attacker. Although it may be possible in some circumstances to determine which node is the attacker and which node is the victim, it is impossible to do so for all cases. Consider the case where an attacker at v_1 launches an attack strictly at v_2. That is, v_2 is the only node that can attest to the false data reported by v_1. In this case, a third-party observer would only be able to conclude that a disagreement exists between v_1 and v_2. In many practical situations, it would be required to know which node is acting falsely. However, for the purpose of this work, it is sufficient to determine the edge disagreement.

3.4 False Positives and False Negatives

A common approach to attack detection is to set a threshold that distinguishes normal data from corrupted data. The threshold could be derived statistically to have a desired property. Against a sophisticated attacker, the threshold would also define attacks that would be considered to be tolerable. If a threshold for divergence between two measurements is set to 10%, then an attacker who controls one of the measurements may design the attack vector to approach the 10% diver-

gence without exceeding it. If the threshold were to be reduced to 1%, it would impose a tighter limit on the attacks that are tolerated, but it would also increase the false positive rate.

The false positive rate can be adjusted by setting the threshold for attack detection. If the goal is to have fewer false positives, the threshold should be increased; if the goal is have fewer false negatives, the threshold should be reduced. The trade-off between the inversely correlated false positives and false negatives is adjustable, but a practical solution ought to include an effective way to balance them.

3.5 Smart Attacker

If an attacker knows the detection schemes that are employed, a solution must assume that the attacker would actively avoid detection. It is not sufficient to determine the most optimal attack strategy and then defend against it. This is because sub-optimal attacks would go completely unnoticed.

3.6 Smart Grid Example

The smart grid is used to demonstrate a distributed bias attack. Roth and McMillin [7] describe how power mitigation works in a distributed grid infrastructure. The individual nodes report how much power they consume and produce, and an observer then verifies that the reported values are consistent. The model does not take noise into account. However, noise can be handled by checking the difference between two related measurements. If the difference exceeds a certain threshold, then the observer detects the anomaly and takes the appropriate action. If the difference is within the threshold, then the arithmetic mean of the two measurements is agreed to be the true state.

Figure 1 shows how a bias attack is conducted on a smart grid power distribution line between two nodes. The attacker simply underreports the incoming power when consuming electricity and overreports the outgoing power when producing electricity. The amount of bias depends on the threshold for tolerance, which the attacker is assumed to know. In effect, the attack would look like noise to any observer.

The scenario is illustrated in Figure 1, where the attacker underreports the received power, which amounts to theft. The physical connection (1) shared between the nodes has an actual flow of power that both nodes measure (2). The nodes report the measurements to each other, but the attacker injects bias (3). Both the nodes compute the mean state value (4) and residuals (5). If the residuals are small enough, no malicious activity is detected.

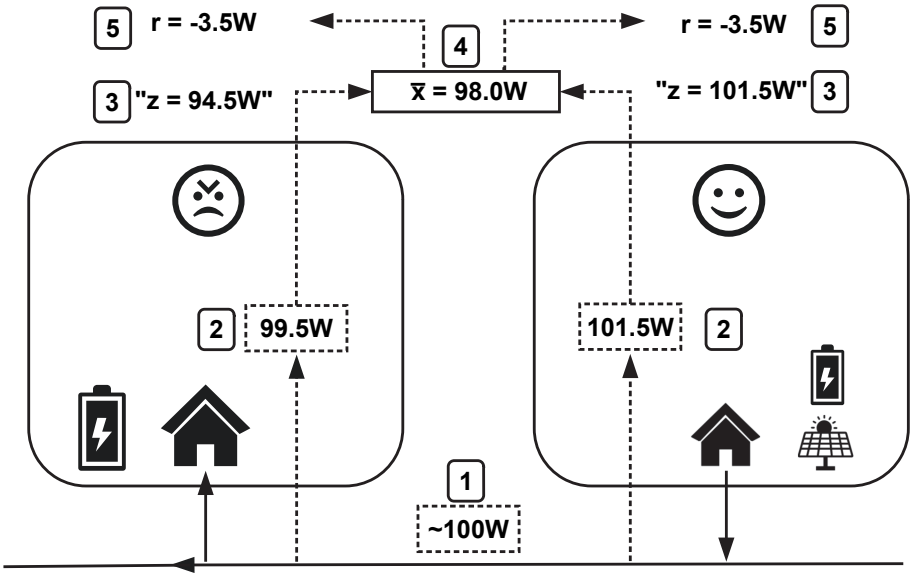

Figure 1. Bias attack on a smart grid power distribution line between two nodes.

In the example, the attacker has gained free power by abusing the noise tolerance. This works because the nodes use consensus to determine the real state of the system. In the absence of an attack, this would be the most reliable approach for many noisy distributed cyber-physical systems. The goal is therefore to demonstrate how bias attacks can be detected efficiently.

4. Proposed Solution

This section discusses describes the proposed solution for distributed bias detection.

4.1 Residual Sum

The residual sum is a central measurement of the integrity of a system. It is defined as:

$$r = z - \hat{x} \tag{6}$$

where $\hat{x}$ is an estimate of the system state x and z is a measurement of x. A large $|r|$ is obviously an indication of erroneous or false data. However, if $r \sim \mathcal{N}(0, \sigma^2)$ is assumed to hold when there is no attack, at any time step, r is expected to be non-zero and $|r| > \tau$ means that

an attack is detected at a steady rate where τ is a threshold for the residual. A certain rate of false positives is expected. If an attacker injects $a = \frac{1}{2}\sigma^2$, the false positives will increase, but it is difficult for an observer to determine that an attack is occurring because σ^2 is not known.

In a single time instant, it may be impossible to discriminate between an attack and a random event. If the residual is tracked over multiple time instants, a clearer picture can be obtained. Taking the sum of absolute residuals until time n would yield a monotonically increasing function. However, the following residual sum function $RSUM$ is obtained upon summing the signed values of r:

$$RSUM(n) = \sum_{i=0}^{n} r \tag{7}$$

The residual sum can be used in much the same way as the residual to determine if the reported measurements are within the expected bounds. A large $|RSUM|$ indicates erroneous or false data, but $RSUM$ has some properties that make it ideal for the problem at hand.

4.2 Statistical Behavior

The residual under no attack is assumed to be $r \sim \mathcal{N}(0, \sigma^2)$ and is $r \sim \mathcal{N}(c, \sigma^2)$ under a bias attack where c is a constant.

The residual sum $RSUM$, which is just the addition of residuals, has the distribution:

$$RSUM(n) \sim \mathcal{N}(0, n \cdot \sigma^2) \tag{8}$$

This follows from the fact that the sum of two normally distributed variables is a normally distributed variable with mean equal to the sum of the means and variance equal to the sum of the variances. Summing up n residuals yields the following distribution of $RSUM_a$ under biased attack:

$$RSUM_a(n) \sim \mathcal{N}(n \cdot c, n \cdot \sigma^2) \tag{9}$$

Over time, the mean of the biased $RSUM_a$ grows at a rate of c whereas the mean of the unbiased $RSUM$ is expected to stay at zero.

This leads to the important observation that too much of a good thing can be a bad thing.

Theorem 1. Let $RSUM_a(n)$ and $RSUM(n)$ be the biased and unbiased residual sums after n measurements, respectively. Let γ be a chosen

significance level for confidence intervals. For any γ, there exists some n for which the confidence interval of $RSUM_a(n)$ does not intersect the confidence interval of $RSUM(n)$.

Proof: Let $\pm z_\gamma$ be the confidence interval for $X \sim \mathcal{N}(0, 1)$ with confidence level γ. Let $\pm z_\gamma(n)$ be the confidence interval for $RSUM(n)$ with confidence level γ for some $n > 1$.

By definition:

$$P(X > z_\gamma) = \gamma \tag{10}$$

$$P(RSUM(n) > z_\gamma(n)) = \gamma \tag{11}$$

The distribution of $RSUM(n)$ is standardized as follows:

$$P(\frac{RSUM(n) - \mu}{\sigma\sqrt{n}} > \frac{z_\gamma(n) - \mu}{\sigma\sqrt{n}}) = \alpha \tag{12}$$

Since μ is zero and $RSUM(n)$ is normalized:

$$P(X > \frac{z_\gamma(n)}{\sigma\sqrt{n}}) = \gamma \tag{13}$$

Comparing Equations (11) and (14) yields:

$$z_\gamma = \frac{z_\gamma(n)}{\sigma\sqrt{n}} \tag{14}$$

And ultimately:

$$z_\gamma(n) = z_\gamma \cdot \sigma\sqrt{n} \tag{15}$$

Since z_γ is a constant, $z_\gamma(n)$ is $O(\sqrt{n})$. Let $\pm z_\gamma(n)'$ denote the confidence interval for $RSUM_a(n)$. Since the mean of $RSUM(n)$ is zero and the mean of $RSUM_a(n)$ is $n\mu$, $\pm z_\gamma(n)'$ will grow at a rate of $O(n\mu \pm z_\gamma \cdot \sqrt{(}n)) = O(n)$. Since $z_\gamma(n)'$ **REPHRASE: in a greater order than** $z_\gamma(n)$, there exists some n_i, such that for every n_j where $j > i$, $|z_\gamma(n)'| > |z_\gamma(n)|$. Hence the two confidence intervals do not intersect after n_i measurements. □

The interpretation of this result is that, no matter what confidence level γ is chosen and how small the bias residual value, eventually the biased $RSUM_a$ will distinguish itself from an unbiased $RSUM$. **DOES THIS GO HERE? Figure 2 shows the diverging confidence intervals for biased and unbiased** $RSUM$ **values**. The important point is that a small attack may go undetected for a long time, but eventually it will be detectable with any arbitrary level of confidence.

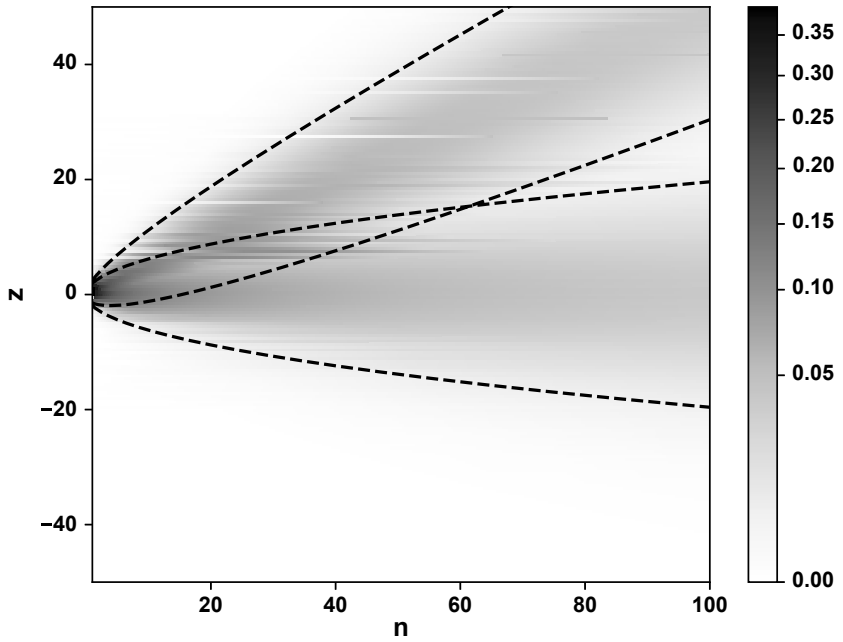

Figure 2. Diverging confidence intervals for biased and unbiased $RSUM$ values.

4.3 Resource Requirements

$RSUM$ is not only a useful tool for detecting bias attacks, but it also requires very limited computational resources for a distributed system. In terms of space, since $RSUM$ can be computed dynamically, each $RSUM$ value requires only a single (numeric) storage location. The solution requires each node to store $2 \cdot deg(v)$ values, where $deg(v)$ is the degree of node v.

In terms of computations, each $RSUM$ requires only one addition at each time step. However, the estimate $\hat{x}$ may require additional computations depending on the method.

In terms of bandwidth, at minimum, each RSUM requires a single value to be transmitted between two nodes. However, this value only has to be transmitted between adjacent nodes, so it will not affect the communications network significantly.

4.4 Published Residuals

A valuable property of the residual sum is its potential to protect the privacy of individual nodes. In a traditional cyber-physical system, where data is sent to a central controller, privacy depends on how well

the controller is trusted. In a fully distributed system, trust may not be assumed, but with the residual sum it may not be an issue.

Consider the case of a smart grid where the nodes may not wish to publish their energy consumption and production to a third-party controller. If the residual sum is published instead, an observer would not be able to determine much. The residual sum only expresses the disagreement between nodes, not the actual amounts of energy transferred.

The only observation that can be made from the residual sum is:

$$RSUM(n) < \sum_{i=0}^{n}(z) \tag{16}$$

This may be considered to be a privacy loss since a non-zero $RSUM$ would indicate activity, and vice versa. However, the concern can be alleviated by incorporating the following noise addition scheme.

Suppose two nodes share a physical connection and publish their shared residual. To obfuscate activity between them, they add a noise value e to every published value. The noise value e is drawn from a list of values L via the following steps:

1. A random seed value is exchanged between the nodes.
2. At coordinated intervals, n values are randomly generated by a linear congruential generator and added to a list L.
3. Every original element e in L is replaced with $\sigma \cdot (e\%100)/100$ and $-e$ is added to the end of L.
4. For every $2n$ time steps, an element e' is drawn randomly and removed from L. The element e' is then added to the current published $RSUM$.

This scheme makes it impossible for an observer to determine if published values between time 0 and $2n$ reflect activity or inactivity. Since the values in L sum to zero, it also guarantees that the published $RSUM$ is accurate at the end of each interval. Although σ may not be known, a suitable scalar may be used in its place.

4.5 Action and Queue-Based Inspection

This subsection demonstrates how the residual sum is used to discover bias attacks. The solution assumes that inspectors are tasked with finding and handling attacks and abnormalities. Such inspectors already exist in electric grids – they conduct routine inspections of meters for

possible tampering. It is assumed that inspectors can identify and handle tampering of any meter, have full access to the entire system and carry specialized equipment. While these assumptions are convenient, they are not far from reality. The main issues are resources and distribution – hiring inspectors and scheduling them effectively.

The solution is to use residual sums to prioritize the work of the inspectors. Simply put, the focus is on the edges in the system that have high residual sums. Since the residual sums are published, the inspectors do not have to travel to the sites and can operate from a central control facility to handle a large region.

The idea is to set a threshold τ for the residual sums and alert the inspectors to edges that exceed τ. Choosing a τ value effectively sets the rates of true and false positives. If τ is too low, more false positives would be generated than the inspectors could handle. If τ is too high, the wait times would be much longer before inspectors can act on any malfeasance. A good threshold should yield manageable true and false positive rates.

However, a simpler approach that does not rely on thresholds exists. This approach sorts the residual sums by magnitude and schedules inspections at the corresponding locations in descending order. Thus, the most likely attacked nodes are inspected first and at the exact rate that the inspection crews can handle.

5. Experimental Results

Simulation experiments were conducted to demonstrate the effectiveness of using residual sums to detect bias attacks. The simulations were conducted using the well-known IEEE 14-bus system [1] and the MATPOWER package in MatLab.

The variance of the measurements was artificially set. It was accomplished by producing a set of base measurements by iteratively running state estimation on an initial set of measurements and replacing them with the estimated values. This yielded a measurement set with a square sum residual close to zero. The measurement set served as the basis for the simulation experiments.

Simulations were executed over 10,000 time steps. At each step, some $\mathcal{N}(0, \sigma^2)$ distributed noise was added to the base measurement of a node with $\sigma = 0.1$. All the simulations used ten unbiased nodes and a varying number k of biased nodes. For the biased nodes, an additional 0.01 was added to each measurement, corresponding to exactly one-tenth of the standard deviation of the noise.

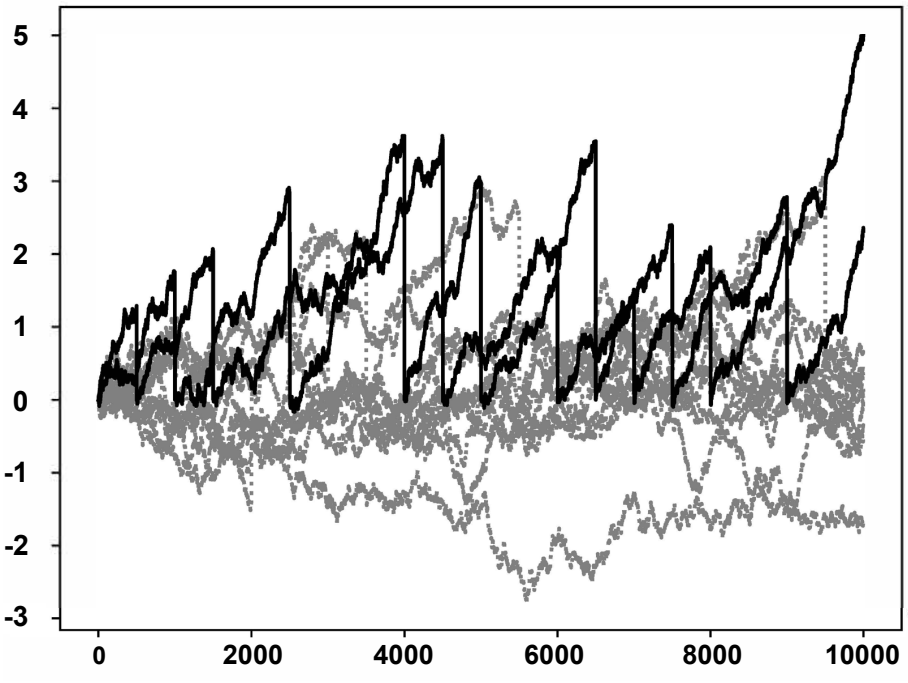

Figure 3. Simulations with ten unbiased (gray) and two biased (black) $RSUMs$.

State estimation of the power system was conducted using MATPOWER at each time step using the adjusted measurements. The residual of each variable was then calculated as the difference between the measurement and estimate, and each residual was added to its corresponding residual sum. At regular intervals h, the highest magnitude residual sum was identified and reset to zero.

Figure 3 shows the simulation results with ten unbiased $RSUMs$ (gray) and $k = 2$ biased $RSUMs$ (black) conducted at regular intervals of $h = 500$ time steps with the highest residual sums were reset to zero. In this particular simulation instance, 13 out of 19 inspections found a biased node to have the highest $RSUM$, corresponding to a true positive rate of 68%. The figure highlights the chaotic nature of the noisy measurements, where some of the unbiased $RSUMs$ become outliers while the biased nodes consistently grow in the positive direction.

Table 1 presents the true positive rates obtained for varying numbers of biased nodes k and inspection intervals with varying lengths (time steps) h. Note that when there are many biased nodes and the inspection intervals are long, the true positive rate effectively becomes 100%. The

Table 1. True positives for varying numbers of biased nodes and inspection intervals.

Biased Nodes (k)	**Inspection Intervals (h)**				
	100	**300**	**500**	**700**	**900**
1	0.27	0.45	0.47	0.57	0.73
2	0.42	0.64	0.68	0.64	0.91
3	0.55	0.67	0.74	0.86	0.91
4	0.7	0.82	0.95	1.0	1.0
5	0.74	0.88	0.95	1.0	1.0
6	0.78	0.91	1.0	1.0	1.0
7	0.86	0.94	1.0	1.0	1.0
8	0.86	1.0	1.0	1.0	1.0
9	0.83	0.94	1.0	1.0	1.0
10	0.93	1.0	1.0	1.0	1.0

table illustrates a practical benefit of the queue approach. Inspectors know the true positive rate and can adjust the inspection interval, but they do know have any information about the number of biased nodes. Nevertheless, they can achieve the desired true positive rate by adjusting the inspection interval.

The simulation results demonstrate the applicability of Theorem 1. Because the biased and unbiased nodes diverge, it becomes easy to prioritize the edges on which inspections should focus.

6. Conclusions

A sophisticated attacker can easily conduct bias attacks on a noisy cyber-physical system while evading conventional detection methods. The proposed scheme for detecting bias attacks leverages the residual sum, a fundamental measurement of bias in cyber-physical systems. The properties that render the residual sum optimal for detection are discussed and theoretical bounds are derived in the absence and presence of bias attacks. The theoretical treatment and the simulation results demonstrate that the detection scheme is highly efficient, privacy preserving and effectively identifies bias attacks.

Future research will attempt to conduct experiments on a physical testbed. While this work has focused on electricity theft, future research will investigate other economic attacks. Additionally, multiple colluding attackers will be considered, which may lead to new challenges and opportunities.

Acknowledgement

The research project was supported by the National Science Foundation under Grant No. CNS-1837472 and by the Missouri S&T Intelligent Systems Center.

References

[1] Illinois Center for a Smarter Electric Grid, IEEE 14-Bus System, Information Trust Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois (`icseg.iti.illinois.edu/ieee-14-bus-system`), 2020.

[2] L. Jia, R. Thomas and L. Tong, Malicious data attack on real-time electricity market, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5952–5955, 2011.

[3] G. Liang, J. Zhao, F. Luo, S. Weller and Z. Dong, A review of false data injection attacks against modern power systems, *IEEE Transactions on Smart Grid*, vol. 8(4), pp. 1630–1638, 2016.

[4] Y. Liu, P. Ning and M. Reiter, False data injection attacks against state estimation in electric power grids, *ACM Transactions on Information and System Security*, vol. 14(1), article no. 13, 2011.

[5] E. Mengelkamp, B. Notheisen, C. Beer, D. Dauer and C. Weinhardt, A blockchain-based smart grid: Towards sustainable local energy markets, *Computer Science – Research and Development*, vol. 33(1-2), pp. 207–214, 2018.

[6] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet and D. Irwin, Private memoirs of a smart meter, *Proceedings of the Second ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, pp. 61–66, 2010.

[7] T. Roth and B. McMillin, Physical attestation of cyber processes in the smart grid, in *Critical Information Infrastructures Security*, E. Luiijf and P. Hartel (Eds.), Springer, Cham, Switzerland, pp. 96–107, 2013.

[8] L. Xie, Y. Mo and B. Sinopoli, False data injection attacks in electricity markets, *Proceedings of the First IEEE International Conference on Smart Grid Communications*, pp. 226–231, 2010.