Correlative Monitoring for Detection of False Data Injection Attacks in Smart Grids

Michael G. Kallitsis [*], George Michailidis [†], Samir Tout [‡]
[*] Merit Network, Inc, University of Michigan, Ann Arbor, Michigan
[†] Department of Statistics, University of Florida, Gainesville, Florida
[‡] Information Assurance, Eastern Michigan University, Ypsilanti, Michigan
mgkallit@umich.edu, gmichail@ufl.edu, stout@emich.edu

***Abstract*—The overarching objective of the modernized electric grid, the *smart grid*, is to integrate two-way communication technologies across power generation, transmission and distribution to deliver electricity efficiently, securely and cost-effectively. However, real-time messaging exposes the entire grid to security threats ranging from attacks that disable information exchange between smart meters and data fusion centers to spurious payload content that would lead to incorrect assessment of actual demand. Such nefarious activities can compromise grid stability and efficiency. Hence, it is important to ensure secure communications and quickly detect malicious activity; this article proposes a framework for detection of *false data injection attacks* in smart grids. We present a *measurement-based* situation awareness framework that combines evidence from sensors at home-area networks, and aims to infer anomalies that signify a coordinated, well-orchestrated attack on residential smart meters at increasing spatial scales. By leveraging multi-view sensor readings, we present a *Bayesian-based correlative monitoring* approach that quickly detects power shifts to anomalous regimes. We evaluate our algorithms using real-world power traces.**

I. Introduction

The electric grid is a "system of systems" that has experienced an expansion of technological capabilities over the last years [1]. The continuously evolving modern grid, known as the *smart grid*, utilizes two-way communication technologies in an integrated fashion across electricity generation, transmission, distribution and consumption to achieve a system that is clean, safe, secure, reliable, resilient, efficient, and sustainable [1], [2]. A key development that is expected to experience significant growth in the next few years is broad-scale adoption of demand response schemes, supported by real-time signaling offered by advanced metering infrastructures (AMIs) [2]. Employed as a means for balancing supply and demand, it engages electricity users in an important role in the operation of the grid by adjusting their electricity usage during peak periods in response to time-based rates or other forms of financial incentives [3]. Demand response has been employed at coarse-time scales for many years; with new advances in smart metering, dynamic pricing information can be enabled in finer temporal and spatial scales to reduce consumption during peak hours and shift demand to off-peak periods.

As gleaned from the above discussion, real-time smart meter reporting would become an apt technology of the emerging smart grid and serve different constituencies. AMIs consist of a hierarchy of communication networks, such as wide-area, neighborhood and home-area networks, each of which is vulnerable to malicious acts [4], [5]. Berthier *et al.* [6] clustered attack techniques against electric grids into three main categories. *Network compromise* attacks could be the result of traffic modification, injection and replay [7], [8]. Further, a *system compromise* attack would trigger illegitimate network operations (e.g., sudden and unwanted power generation commands) by spoofing utility system nodes. Similarly, lower smart meter integrity could be the result of compromised nodes[1] or erroneous smart metering [7], [8], [10]–[13]. In addition, with *denial of service* (DoS) attacks unresponsive nodes could cause grid instability by leading to a faulty or stale system state. DoS attacks could be fabricated via resource exhaustion, wireless signal jamming, TCP injection, and others [7], [10], [11], [14].

Based on this vast array of security threats, one can envision different attack combinations that could result in smart meter *payload spoofing* and alteration of true power demand, which constitutes the primary *threat model* of our paper. Intercepted signaling between AMIs deployed at customer premises and data concentrators in neighborhood-area networks comprises the attack surface. *False data injection attacks* can mislead the grid's state estimation process when carefully crafted to avoid existing "bad data" detection techniques [13]. Numerous scenarios of adversaries who compromise meters and fabricate their readings are discussed in [15]–[18], including malware coordinating instantaneous demand drop, hacked data collector nodes programmed to send messages that reduce and then suddenly dramatically increase power demand, manipulation of electricity pricing, etc.

This paper proposes a framework for tackling the problem of detecting false data injection attacks. Motivated by recent advances in home and building monitoring (e.g., see [19], [20] and energy-harvesting metering [21]) we study a behavioral-based model that integrates sensor measurements from home-area networks and aims to "learn" normal electricity usage patterns[2]. Given the power state (i.e., the one reported by the smart-meter to the utility) and our forecasted usage, one can formulate a sequential hypothesis testing problem that reflects whether the system remains "in-control" or it is operating at

[1] Perhaps the most famous incident of this type is Stuxnet [9].

[2] Considering *user privacy* [22], such information is not sent to the utility.

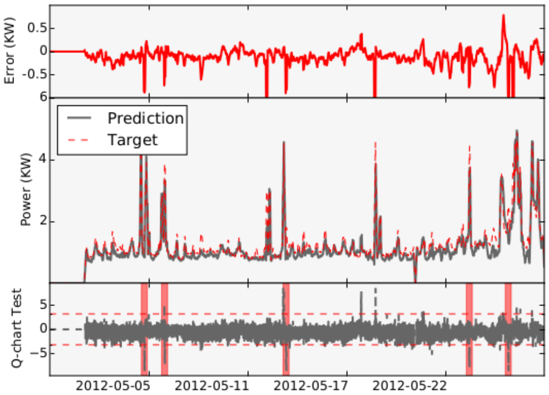


Fig. 1: Our detection framework in action on Smart* data [19]. As the *alert dashboard* shows (see Figure 2, right) all injection attacks are identified (see red vertical stripes).

anomalous regimes (see Figure 1). In addition, the proposed system can be employed as a home-health monitoring system that could identify intrinsic anomalies that originate from malware infecting smart appliances, energy theft or other failures in Internet of Things devices [22].

Our main contributions include: i) the *formulation* of false data detection as an anomaly detection problem based on correlative home-area monitoring, complementing signature and specification-based methods (see next section); ii) a systematic, *modular* approach based on integrating sensor measurements and fusing them into a forecasting and hypothesis testing framework. For example, alternative forecasting modules could be employed, *mutatis mutandis*, assuming they offer a predictive distribution (e.g., Gaussian processes regression) and computational restrictions are not in play; iii) a *lightweight* (see computing times in Figure 2), elegant and adaptive solution to the problem at hand that can be easily transitioned to practice and implemented using inexpensive sensor devices (see, e.g., [19], [21]) and computing nodes.

The rest of the paper is organized as follows. Section II gives an overview of related work in anomaly detection for power grids. Our Bayesian-based framework is introduced in III, followed by performance evaluation with real-world datasets.

II. Related Work

Discovery of nefarious activities in the electric grid can be performed using a combination of signature-based, specification-based and anomaly-based detection methods [4], [6], [7]. The first is suitable for identifying malware that has already appeared in a smart grid environment, and its behavior has been recorded in black-list databases with known malware signatures. It examines packets as they arrive to the utility's control center and looks for patterns of malicious activity (e.g., *Snort* [23]). Specification-based detection is accomplished by measuring deviations from a normal operational profile that is predefined. Examples include finite state machine monitors, data validation with range checks, authentication monitor and physical health inquiries for catching unresponsive nodes, and verification of system state [24], [25].

One shortcoming of relying on prior knowledge recorded in black-lists is that new malware activities will not be uncovered. Similarly, specification-based methods can be cumbersome to fine-tune; finding a valid range for the AMI power demand and supply is not easily determined. Further, subtle attacks might exist that involve modifying control parameters in a way that appears to be within a normal range, but still being capable of inflicting system damage. Instead, anomaly-based methods try to identify anomalies by checking for significant deviations from normal traffic patterns; epigrammatically, one monitors the signal of interest to "learn" its normal behavior through a training period, and detects outliers when a statistic exceeds a predefined threshold. Our framework falls in this category; via a model-based probabilistic technique, we build the normal operating regime and seek for outliers.

Existing anomaly-based defenses against adversaries that inject spurious data measurements into the power grid follow a "network-view" perspective. Such countermeasures for detecting false data injection appear in [18], [26]–[28]. [26] proposes an adaptive cumulative sum test combined with a multivariate hypothesis testing problem to prevent an erroneous grid-state estimate. [27] studies a graph theoretic method for securing an optimal set of meter measurements so that state estimation is not compromised. [18] couples anomaly-based methods with a data integrity check to combat stealth attacks, while [28] looks for inconsistent grid behavior using clustering techniques. Instead, we tackle the problem from a different vantage point. The "home-area view" we suggest aims to detect arbitrary data injection attempts at their origin, i.e., compromised residential smart meters. Our framework complements the above-mentioned work since the alert output signal generated by our methods could serve as an additional input[3] to [18], [26]–[28].

III. Proposed Framework

The data injection threat highlighted above can be tackled through the lens of a decision problem on non-stationary time-series measurements. We present an adaptive system for home-area situation awareness. Our system employs sensor measurements that can be readily collected in home-area networks (e.g., motion, temperature, appliance usage, total electricity; see [19], [20]) to train a model that keeps track of the *expected* electricity pattern. Deviations between predicted versus realized power consumption are fed to a sequential hypothesis testing module that decides on the presence of abnormalities. Framework 1 provides a sketch of the proposed approach.

We first introduce the *forecasting module* and continue with the *hypothesis testing* one. We base our predictions on Bayesian linear regression. By following a self-tuned Bayesian approach, we avoid the need for model regularization and cross-validation that can be computationally expensive when

[3] These signals should be integrity protected by encryption schemes.

Framework 1 Measurement-based False Data Detection

```
Require: For each forecasting period: new training set $\mathbf{X}$ and $\mathbf{t}$.
Require: Control chart parameters $\lambda$ and $L$.
Require: Robust threshold $\theta_r$ and period $\nu$.
[Start] Fit the model and begin data monitoring.
[Forecast] Upon observing $(t_n, \mathbf{x}_n)$, compute $y(\mathbf{x}_n, \mathbf{w})$.
[Update] Compute error $e_n = t_n - y(\mathbf{x}_n, \mathbf{w})$.
[Control Chart] Compute $S_n = f(\lambda, L, e_n)$.
[Robust EWMA] Apply two-in-a-row rule on $S_n$ (see section III-B).
[Robust Filter] Update $A = \{k : \ |S_k| > L\sigma_\lambda, k = n - \nu, \ldots, n\}$.
[Decision] Raise alarm if $|A| > \theta_r$, else system is in-control.
```

online monitoring is required. Further, a linear model is well-suited when individual power circuits and appliances are monitored, as in our case. However, the system designer is not restricted to these choices, and alternative prediction models can be considered. Similarly, one can opt for a different decision module. We decided to work with an exponentially weighted moving average control chart due to its simplicity and robustness, but other stopping rules also apply (e.g., see Wald-based detection [29]).

A. Forecasting Power Utilization

We provide a basic overview of Bayesian regression, and the reader is referred to [30] for extended discussion. Throughout the paper, the smart meter power state is denoted as t and sensor observations are denoted by vector $\mathbf{x} = (x_1, \ldots, x_M)^T$. In other words, variables t would play the role of *target* values in our prediction scheme, and input $\mathbf{x}$ would be the vector of independent variables known as *features*. For each new forecasting epoch (see Framework 1), a training set of size N is available; $\mathbf{t} := (t_1, \ldots, t_N)^T$ represents the target values in the training set, and $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ are the corresponding target values. We construct the $N \times M$ *measurement matrix* $\mathbf{X}$ by stacking the input variables of each data point. Our linear regression model involves a linear combination of inputs, i.e., $y(\mathbf{x}, \mathbf{w}) = \mathbf{w}^T\mathbf{x}$, where $\mathbf{w}$ are the *parameters/weights* that need to be determined. We further assume that given the value of $\mathbf{x}$, the corresponding value of t has a Gaussian distribution with mean equal to $y(\mathbf{x}, \mathbf{w})$ and variance β^{-1}. Thus,

$$p(t|\mathbf{x}, \mathbf{w}, \beta) = \mathcal{N}(t|y(\mathbf{x}, \mathbf{w}), \beta^{-1}). \tag{1}$$

Assuming the data is drawn independently from (1), the *likelihood* is $p(\mathbf{t}|\mathbf{X}, \mathbf{w}, \beta) = \prod_{n=1}^{N} \mathcal{N}(t_n|y(\mathbf{x}_n, \mathbf{w}), \beta^{-1})$. In a Bayesian setting, a prior of the model parameters $\mathbf{w}$ is introduced. We consider a *conjugate prior*, zero mean isotropic Gaussian governed by a single parameter α, i.e., $p(\mathbf{w}|\alpha) = \mathcal{N}(\mathbf{0}, \alpha^{-1}\mathbf{I})$, where $\mathbf{I}$ is the identity matrix of appropriate dimension. The *posterior distribution*, which is proportional to the product of the likelihood function and the prior, takes the form of another Gaussian distribution

$$p(\mathbf{w}|\mathbf{t}) = \mathcal{N}(\mathbf{w}|\mathbf{m}_N, S_N), \tag{2}$$

with $\mathbf{m}_N = \beta S_N \mathbf{X}^T\mathbf{t}$ and $S_N^{-1} = \alpha\mathbf{I} + \beta\mathbf{X}^T\mathbf{X}$. The optimal parameter vector $\mathbf{w}^*$ in $y(\mathbf{x}, \mathbf{w})$ is obtained by maximizing the posterior distribution. Since this is a Gaussian distribution, its mode coincides with the mean, and thus the maximizing vector is $\mathbf{w}^* = \mathbf{m}_N$.

Further, our framework requires knowledge of the *predictive distribution*. For a new data point $(t, \mathbf{x})$ (we omit time indexing to keep notation uncluttered) this is defined by $p(t|\mathbf{x}, \mathbf{t}, \alpha, \beta) = \int p(t|\mathbf{x}, \mathbf{w}, \beta)p(\mathbf{w}|\mathbf{t}, \alpha, \beta)d\mathbf{w}$. The conditional distribution $p(t|\mathbf{x}, \mathbf{w}, \beta)$ is given by (1) and the weight posterior distribution is given by (2). The predictive distribution is hence the result of the convolution of two Gaussians [30], and takes the form

$$p(t|\mathbf{x}, \mathbf{t}, \alpha, \beta) = \mathcal{N}(t|\mathbf{m}_N^T\mathbf{x}, \sigma_N^2(\mathbf{x})), \tag{3}$$

where the variance of the predictive distribution is given by $\sigma_N^2(\mathbf{x}) = \beta^{-1} + \mathbf{x}^T S_N \mathbf{x}$. The first term represents the noise in the data, and the second term reflects the uncertainty in making predictions associated with the parameter vector $\mathbf{w}^*$. Thus, our model is adaptively learning the variance of the predictive distribution, something that would be proven very important for tracking the "reference distribution" in our hypothesis testing module, described shortly.

So far we have assumed that *hyperparameters* α and β are known. In a fully Bayesian treatment, one introduces prior distributions for them. Predictions are then made by marginalizing with respect to these hyperparameters as well as with respect to parameters $\mathbf{w}$. Instead of performing a complete marginalization over all these variables (which is analytically intractable for some choices of prior hyperparameters, can be computationally intensive if done numerically or can lead to poor results [30]) we follow a technique called *evidence approximation* [30], [31]. Using this approximation, the hyperparameters are determined by just looking at the training data. The technique amounts to an iterative approach, similar in spirit to Expectation-Maximization algorithms. In the evidence approximation, the values of α and β are obtained by maximizing the *marginal likelihood* $p(\mathbf{t}|\alpha, \beta) = \int p(\mathbf{t}|\mathbf{w}, \beta)p(\mathbf{w}|\alpha)d\mathbf{w}$, that represents the "evidence" for a particular choice of the hyperparameters given the observed data. The iterative procedure starts with initial values of α and β and uses them to compute $\mathbf{m}_N$. It then derives the eigenvalues λ_i of the eigenvector equation

$$(\beta\mathbf{X}^T\mathbf{X})\mathbf{u}_i = \lambda_i\mathbf{u}_i, \text{ for } i = 1, \ldots, M. \tag{4}$$

Then quantity γ is computed as $\gamma = \sum_{i=1}^{M} \frac{\lambda_i}{\alpha + \lambda_i}$, which is used to obtain the updated value of α that maximizes the marginal likelihood

$$\alpha = \frac{\gamma}{\mathbf{m}_N^T\mathbf{m}_N}. \tag{5}$$

Following similar steps, one can maximize the marginal likelihood with respect to β as well, and obtain a new value for β as

$$\frac{1}{\beta} = \frac{1}{N - \gamma}\sum_{n=1}^{N}\{t_n - \mathbf{m}_N^T\mathbf{x}_n\}^2. \tag{6}$$

The iterative cycle of finding $\mathbf{m}_N$, γ and using them to update α, β repeats until α and β reach a stationary point.

B. Online Detection: Hypothesis Testing

The forecasting module provides predictions about electricity usage based on house sensor measurements. The next step computes the difference of the predicted value with the actual smart meter reading, and formulates a *sequential hypothesis testing* problem to decide whether the sequence of values observed comes from a system operating at the normal regime (i.e., values obey the *Null Hypothesis* or, equivalently, *reference* distribution).

The reference distribution, denoted as F_n, for the differences (referred as *errors* henceforth) comes from the predictive distribution described earlier. Following [32], for each new observation $(t_n, \mathbf{x}_n)$ we calculate the error $e_n := t_n - \mathbf{m}_N^T \mathbf{x}_n$, and then find the p value corresponding to that error using the fact that

$$p(e_n|\mathbf{x}_n, \mathbf{t}, \alpha, \beta) = \mathcal{N}(e_n|0, \sigma_N^2(\mathbf{x}_n)). \tag{7}$$

The p value p_n for negative errors e_n with reference distribution F_n is set to be the lower-tail probability, $F_n(e_n)$. If the error is positive, then $p_n = 1 - F_n(e_n)$. We are interested in employing a hypothesis testing criterion for detecting sequences of "abnormally" small p values. We monitor for anomalies by utilizing an Exponentially Weighted Moving Average (EWMA) control scheme [33], known as *Q-charting* in quality control. We first take the normal score $\Phi^{-1}(p_t)$ of the p value, where Φ^{-1} is the standard normal cumulative distribution inverse function. This allows application of standard control chart methods for detecting "out-of-control" values.

In short, event detection is based on thresholding

$$S_n = (1 - \lambda)S_{n-1} + \lambda Z_n, \quad \text{where } Z_n = \Phi^{-1}(p_n),$$

for a weight λ in $(0, 1]$. Both the magnitude and duration of the anomalous event can drive the value of S_n to a level where an alert is triggered. For example, abrupt power shifts (e.g., elevating power by 6KW, see Table I) would be almost instantaneously detected with high probability. On the other hand, "stealthy" power shifts (e.g., 1KW) could be unnoticeable for awhile, but as their duration persists detection probability elevates.

The sensitivity of EWMA is tuned by the weight λ and the threshold parameter L. When the process is under control and the reference distribution is suitable, Z_n is distributed approximately as normal $\mathcal{N}(0, 1)$. Assuming independent Z_n's, the severity test S_n is approximately normal $\mathcal{N}(0, \sigma_\lambda^2)$ with $\sigma_\lambda^2 = \lambda/(2 - \lambda)$. [33] provides guidelines on calibrating the control chart by choosing appropriate values of λ and L that balance the time between false alarms (named as average run-length in [33]) and the ability to determine whether the process under control has "shifted" to anomalous regimes of certain magnitude. Extensive experimentation suggests that (λ, L) pairs $(.53, 3.714)$, $(.84, 3.719)$, $(1, 3.719)$ are sensible options for monitoring real-world data (see Table I).

To tame the false alarm rate we engage a *robust EWMA* technique by exercising the two-in-a-row rule [33]. When a single outlier is observed, i.e., $|S_n| > \sigma_\lambda L$, the control statistic S_n remains unchanged and a counter is set. If the next observation (the new normal score) makes the updated control statistic to lie within the outlier limits, the counter resets; otherwise, an out-of-control signal is given. The final step of our framework is the *robust filter*, and checks for persistent out-of-control signals. As Framework 1 depicts, we keep a history window of ν observations and maintain a counter of the out-of-control signals seen in that window. An alert is raised whenever the counter exceeds the user-defined threshold θ_r. The severity of alert level can be visualized with a lively updated dashboard, as illustrated in Figure 2 (right panel).

IV. Performance Evaluation

We evaluate our framework using the Smart* (SmartStar) dataset [19], and data from a colleagues's residence (with added, synthetic noise). The Smart* projects provides real-world fine granularity power data from several households; we use "Home A". Aggregate power along with the power of each individual *circuit* in the house (25 in total) are recorded. In addition, measurements from several switches, meters, environmental factors (e.g., indoor/outdoor temperature, etc.), motion and others are available. Overall, we use 26 features as independent inputs for prediction, taken at 1-minute snapshots.

Figure 2 (left) evaluates the prediction accuracy and computational performance of our forecasting algorithm. Computationally expensive operations include the matrix inversion, the singular value decomposition (both entail a $O(M^3)$ complexity), and the matrix multiplication of order $O(NM^2)$ needed to fit the regression model (i.e., training phase, see Eqs. (2) and (4)). As Figure 2 illustrates though, training time is of the order of milliseconds, and predictions take negligible time. We also compute and show the *relative mean squared error* (ReMSE). ReMSE is defined as, $\mathrm{ReMSE} = \sum_n (y(\mathbf{x}_n, \mathbf{w}) - t_n)^2 / \sum_n t_n^2$, for all n in the forecasting period. We demonstrate performance for various sizes of training and forecasting intervals (note that re-training is needed after the end of each forecasting interval). As expected, training times monotonically increase linearly (left panel of Figure 2) as the dimension N of matrix $\mathbf{X}$ increases (in our case, $M \ll N$). Figure 2 (left) suggests that good prediction accuracy is achieved by balancing the number of observations N (training size) with the forecasting period. Extensive experimentation advocates training sizes ranging from 24 to 96hrs worth of measurements and a look-ahead forecasting period of 30 to 60 time points.

Table I tabulates results on detection accuracy. We focus on four different scenarios in which power is altered from its true value by the corresponding *shift* value (in KWs). For each scenario, we inject five anomalies at random times with *duration* of 30 observation intervals, and we repeat the given experiment for a total of 50 runs. For example, Figure 1 (lower panel) shows five anomalies occurring at the time instances with red vertical stripes. It also displays time-series of errors (upper panel) and predictions aside the actual targets for three months (middle). We mitigate false positives with a robust filter of length $\nu = 10$ minutes and $\theta_r = 3$, as shown in Figure 2 that depicts a dashboard-like visualization that can be lively updated. We report results on the *mean delay* elapsing before EWMA notices the first out-of-control point after a data injection, and on the *precision* and *recall* of the events classification. Let Tp, Fp and Fn denote the number of *true positives*, *false positives* and *false negatives*, respectively. Precision is defined as the ratio $Tp/(Tp + Fp)$, recall as $Tp/(Tp + Fn)$ and both lie in $[0, 1]$. We also include the *F1-score*, the harmonic mean thereof. Briefly, a hypothesis test that is too sensitive gives a higher number of false alarms and this would lower the detection precision. On the contrary, a test that misses anomalies (i.e., false negatives) will have a low recall score. We tabulate results on 3 pairs of EWMA parameters λ and L.

It is intuitively appealing that accuracy (as reflected by precision, recall and detection delay) elevates as the power shift increases. Indeed, to defend against false data injection attacks, these sudden power spikes should be rapidly diagnosed. In addition, Figure 2 (middle panel) confirms that the model chosen for the reference distribution is adequate. The figure shows the distribution of the p values under the Null hypothesis. When the system is in control, the p values of the reference distribution (see Eq. (7)) are *uniformly distributed*. This visualization serves as a *model validation* for the user, and the system can be restarted or stopped for re-calibration when the distribution of p values deviates from uniformity.

Next, we check identification accuracy on stealth data injections (i.e., spurious data of low intensity). Even though sporadic, low magnitude injections are unlikely to threaten grid stability (they fall within the normal load tolerance of the grid [34]; instead, detecting dramatic shifts is the primary concern), we delve more into it. Table II sheds more light on performance for power shifts around 10% of the peak demand ("Home A" of [19] peaks around 10KW). As shown, for the selected $(.53, 3.714)$ EWMA parameters, probability of exposure increases as the duration of attack persists. We notice, however, that some small injections may remain undetected and false positives may surface. Although this needs further examination as part of ongoing work, we conjecture that the network-wide detection accuracy would not be compromised. We believe that at a data concentrator center that collectively monitors alerts generated by our system, **false positives would be uniformly spread in time and would not trigger a network-wide alert. On the contrary, coordinated attacks (stealth or intense injections) would be correctly identified.**

Further, we performed experiments on simulated data; i.e., data for which the linear relationship between independent inputs and the target variable actually exists, given the value of an additive Gaussian noise (zero mean, standard deviation 300W). Results for EWMA pairs $(.29, 3.686)$ and $(.53, 3.714)$ are tabulated in Table III. The high recall and precision scores clearly emphasize that under the correct model assumptions, our framework is highly suitable for the problem at hand.

Finally, we employed the naïve approach of an EWMA-based change-point detection by just looking at the target electricity values. This method is severely inaccurate due to a plethora of false positive alerts (around 4 false alarms every 100 observations).

TABLE I: Evaluation of detection performance on the Smart* dataset. Values in parenthesis signify standard deviations.

Shift (KW)	Weight λ	Delay (in mins)	Precision	Recall	F1-score
-1	1	$9.7_{(7.2)}$	$.29_{(.45)}$	$.07_{(.11)}$	.11
-1	.53	$8.1_{(4.6)}$	$.76_{(.41)}$	$.29_{(.21)}$	.42
-1	.84	$10.4_{(5.5)}$	$.48_{(.50)}$	$.12_{(.14)}$	.19
1	1	$8.0_{(4.5)}$	$.75_{(.43)}$	$.26_{(.19)}$	.38
1	.53	$3.4_{(1.7)}$	$.95_{(.17)}$	$.50_{(.22)}$	.66
1	.84	$6.6_{(3.4)}$	$.86_{(.31)}$	$.31_{(.19)}$	.46
3	1	$1.1_{(.5)}$	$.98_{(.05)}$	$1.00_{(.03)}$	.99
3	.53	$1.0_{(.0)}$	$.98_{(.06)}$	$1.00_{(.03)}$	.99
3	.84	$1.0_{(.2)}$	$.98_{(.06)}$	$1.00_{(.03)}$	.99
6	1	$1.2_{(.9)}$	$.96_{(.11)}$	$.99_{(.05)}$	.97
6	.53	$1.0_{(.0)}$	$.97_{(.08)}$	$1.00_{(.05)}$	.98
6	.84	$1.0_{(0.0)}$	$.96_{(.09)}$	$.99_{(.05)}$	.98

TABLE II: Detection performance on stealth injections.

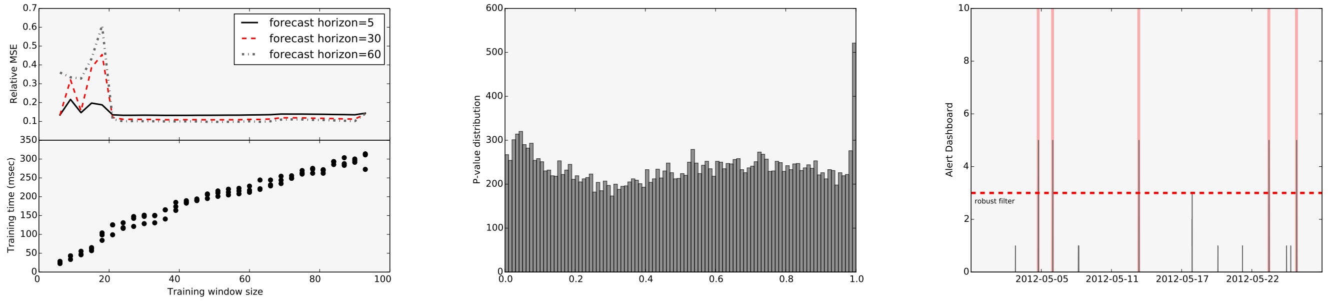
Duration (in mins)	Shift (KW)	Weight λ	Delay (in mins)	Precision	Recall
10	1	.53	$2.7_{(0.7)}$	$.05_{(.11)}$	$.03_{(.07)}$
20	1	.53	$2.9_{(1.0)}$	$.61_{(.19)}$	$.40_{(.20)}$
30	1	.53	$3.8_{(1.8)}$	$.65_{(.20)}$	$.49_{(.23)}$
40	1	.53	$3.8_{(1.8)}$	$.70_{(.09)}$	$.52_{(.19)}$
50	1	.53	$4.0_{(2.2)}$	$.73_{(.14)}$	$.61_{(.23)}$
60	1	.53	$4.2_{(2.5)}$	$.72_{(.09)}$	$.58_{(.21)}$
120	1	.53	$4.8_{(4.7)}$	$.74_{(.09)}$	$.60_{(.20)}$

V. Discussion

Electric grids are verging on large technological changes due to the introduction of modern devices with two-way communication capabilities. Hence, it is important to ensure secure communications together with developing the necessary infrastructure to timely detect nefarious activities. A major security threat is false data injection attacks. Previous work tackles the problem from a network-view perspective (i.e., looking at data from multiple grid nodes). Here, we consider a different vantage point and aim at detecting attacks directly at their origin; the home-area level. We follow a measurement-based approach, and use data collected from sensors deployed in a home network to learn and be able to forecast electricity consumption. Large deviations from expected consumption are flagged anomalies by our online Q-chart scheme.

We use a light-weight algorithm for forecasting, such as Bayesian linear regression, but our framework can be extended with more elaborate models. One could employ non-linear regression (e.g., Gaussian processes, neural networks) if non-linearities between the input variables and the targets exist. In addition to spoofing detection, our system offers a health-assessment for other threats, such as energy theft or other power leaks. We believe that the proposed approach can work in harmony with signature-based anomaly detection methods, and can offer a complementary "grid health" signal to existing methods that take a network-view approach.

Acknowledgements: We thank Meghan Clark and Prabal



(a) Forecasting Evaluation.

(b) Model Validation.

(c) Alert Dashboard.

Fig. 2: (a) Evaluation of forecasting with different training sizes (in hrs). For these experiments, the average prediction time / iteration was 23.8μsec with standard deviation of 1.8μsec. Experiments were run on a 3.07GHz CPU. (b) Model validation: histogram of p values under the Null. (c) Detection performance on 4 weeks of Smart* dataset (see Fig. 1). Injected anomalies shown with red vertical stripes (as on Fig. 1). As shown in the right panel, all power shifts of 3KW were detected.

TABLE III: Detection performance on simulated data.

Duration (in obs)	**Shift (KW)**	**Weight λ**	**Delay (in obs)**	**Precision**	**Recall**
10	1	.29	$1.3_{(.2)}$	$1.00_{(.00)}$	$.98_{(.06)}$
	1	.53	$1.2_{(.5)}$	$1.00_{(.00)}$	$.82_{(.19)}$
	3	.29	$0.0_{(.0)}$	$1.00_{(.00)}$	$1.00_{(.00)}$
	3	.53	$0.0_{(.0)}$	$1.00_{(.00)}$	$1.00_{(.00)}$

Dutta at Lab11, UM for valuable discussions. Work supported by NSF CNS-1422078.

References

[1] H. Gharavi and R. Ghafurian, "Smart Grid: The Electric Energy System of the Future," *Proceedings of the IEEE*, June 2011.

[2] Massachusetts Institute of Technology, *The Future of the Electric Grid: An Interdisciplinary MIT study*. MIT Energy Initiative, 2001.

[3] S. Caron and G. Kesidis, "Incentive-based energy consumption scheduling algorithms for the smart grid," in *IEEE SmartGridComm*, 2010, pp. 391–396.

[4] W. Wang and Z. Lu, "Survey cyber security in the smart grid: Survey and challenges," *Comput. Netw.*, Apr. 2013.

[5] D. Wei, Y. Lu, M. Jafari, P. Skare, and K. Rohde, "Protecting smart grid automation systems against cyberattacks," *Smart Grid, IEEE Transactions on*, vol. 2, no. 4, pp. 782–795, Dec 2011.

[6] R. Berthier, W. Sanders, and H. Khurana, "Intrusion detection for advanced metering infrastructures: Requirements and architectural directions," in *IEEE SmartGridComm*, 2010, pp. 350–355.

[7] F. Cleveland, "Cyber security issues for advanced metering infrastructure," in *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, 2008, pp. 1–5.

[8] S. McLaughlin, D. Podkuiko, and P. McDaniel, "Energy theft in the advanced metering infrastructure," in *Proceedings of CRITIS'09*, 2010, pp. 176–187.

[9] N. Falliere, L. Murch, and E. Chien, "W32.stuxnet dossier," 2011.

[10] W. Sikora, M. Carpenter, and J. Wright, "Smart Grid and AMI Security Concerns," inGuardians and Industrial Defender, 2009.

[11] T. Goodspeed, D. R. Highfill, and B. A. Singletary, "Low-level Design Vulnerabilities in Wireless Control System Hardware," proceedings of the Scada Security Scientific Symposium (S4), 2009.

[12] M. Davis, "SmartGrid Device Security," presentation at BlackHat 2009.

[13] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proceedings of CCS '09*, ser. CCS '09. New York, NY, USA: ACM, 2009, pp. 21–32.

[14] D. Formby, S. S. Jung, J. Copeland, and R. Beyah, "An empirical study of TCP vulnerabilities in critical power system devices," in *Proceedings of SEGS '14*. New York, NY, USA: ACM, 2014, pp. 39–44.

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

[16] A. Metke and R. Ekl, "Security technology for smart grid networks," *Smart Grid, IEEE Transactions on*, vol. 1, no. 1, pp. 99–107, 2010.

[17] N. S. T. Bed and U. D. of Energy Office of Electricity Delivery Energy Reliability, "Study of Security Attributes of Smart Grid Systems - Current Cyber Security Issues," April 2009.

[18] W. Yu, D. Griffith, L. Ge, S. Bhattarai, and N. Golmie, "An integrated detection system against false data injection attacks in the smart grid," *Security and Communication Networks*, vol. 8, no. 2, pp. 91–109, 2015.

[19] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht, "Smart*: An open data set and tools for enabling research in sustainable homes," 2012.

[20] D. J. Cook, A. S. Crandall, B. L. Thomas, and N. C. Krishnan, "Casas: A smart home in a box," *Computer*, vol. 46, no. 7, pp. 62–69, 2013.

[21] B. Campbell and P. Dutta, "An energy-harvesting sensor architecture and toolkit for building monitoring and event detection," in *Proceedings of BuildSys '14*. ACM, 2014, pp. 100–109.

[22] C. Medaglia and A. Serbanati, "An overview of privacy and security issues in the internet of things," in *The Internet of Things*, 2010, pp. 389–395.

[23] Sourcefire, "Snort," http://www.snort.org/snort.

[24] H. M. Hassan, M. Mahmoud, and S. El-Kassas, "Securing the aodv protocol using specification-based intrusion detection," in *Proceedings of Q2SWinet '06*. New York, NY, USA: ACM, 2006, pp. 33–36.

[25] A. Carcano, I. N. Fovino, M. Masera, and A. Trombetta, "State-based network intrusion detection systems for SCADA protocols: A proof of concept," in *Proceedings of CRITIS'09*, 2010, pp. 138–150.

[26] Y. Huang, M. Esmalifalak, H. Nguyen, R. Zheng, Z. Han, H. Li, and L. Song, "Bad data injection in smart grid: attack and defense mechanisms," *Communications Magazine, IEEE*, Jan. 2013.

[27] S. Bi and Y. J. Zhang, "Graphical methods for defense against false-data injection attacks on power system state estimation," *Smart Grid, IEEE Transactions on*, vol. 5, no. 3, pp. 1216–1227, May 2014.

[28] A. Almalawi, X. Yu, Z. Tari, A. Fahad, and I. Khalil, "An unsupervised anomaly-based detection approach for integrity attacks on SCADA systems," *Computers and Security*, vol. 46, no. 0, pp. 94 – 110, 2014.

[29] J. Jung, V. Paxson, A. Berger, and H. Balakrishnan, "Fast portscan detection using sequential hypothesis testing," in *Security and Privacy, 2004. Proceedings. 2004 IEEE Symposium on*, May 2004, pp. 211–225.

[30] C. M. Bishop, *Pattern Recognition and Machine Learning*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[31] D. MacKay, "Bayesian interpolation," *Neural Computation*, 1991.

[32] D. Lambert and C. Liu, "Adaptive thresholds: Monitoring streams of network counts," *online, J. Am. Stat. Assoc*, pp. 78–89, 2006.

[33] J. M. Lucas and M. S. Saccucci, "Exponentially weighted moving average control schemes: Properties and enhancements," *Technometrics*, vol. 32, no. 1, pp. 1–29, Jan. 1990.

[34] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Phys. Rev. E*, vol. 69, p. 045104, Apr 2004.