ADAPTIVE STATISTICAL DETECTION OF FALSE DATA INJECTION ATTACKS IN SMART GRIDS

Michael G. Kallitsis[⋆] *Shrijita Bhattacharya*[§] *Stilian Stoev*[§] *George Michailidis*[‡]

[⋆] Merit Network, Inc., Ann Arbor, MI
[§] Statistics, University of Michigan, Ann Arbor, MI
[‡] Statistics, University of Florida, Gainesville, FL
mgkallit@merit.edu, shrijita@umich.edu, sstoev@umich.edu, gmichail@ufl.edu

ABSTRACT

The smart power grid is a synergistic system that integrates diverse network components for power generation, transmission, and distribution. Its advanced metering infrastructure (AMI) enables the grid's efficient and reliable operation. Nevertheless, it is amenable to advanced cyber threats; malicious actors can compromise vulnerable meters and arbitrarily alter their readings. These orchestrated "false data injection attacks" can lead to power outages and service interruption. We propose a framework that uses measurements from trusted (secure) nodes in order to detect abnormal "spoofing" activity of other nodes, possibly tampered. Our model considers the structural similarities in the electricity consumption of AMI nodes, and exploits the spatial correlation amongst meters. To alleviate the problem's large-dimensionality aspect, the meters are *clustered* into classes of similar energy patterns. We evaluate our algorithms using *real-world* building data obtained from a large university campus.

Index Terms— Network kriging, anomaly detection, false data injection attacks, AMI power data

1. INTRODUCTION

The next generation electric grid, the smart grid, is expected to tackle some of the fundamental limitations of the traditional electric grid. Key functionality includes the provisioning of accurate situation awareness of the entire grid, capabilities of fine-grained asset control, incorporation of renewable energy sources, etc. A critical grid component for addressing these new requirements is the smart grid's advanced metering infrastructure (AMI) that provides two-way communication capabilities. This leads to better network monitoring and problem mitigation, and paves the way for a network that self-heals from outages or other anomalies [1].

Notwithstanding the foregoing, this ever-increasing connectivity of the electric grid has ended its isolation with "external" communication networks, such as the Internet. Adversaries are nowadays capable of inflicting *physical* damage into critical smart grid infrastructure. A plethora of vulnerable industrial control or smart grid devices can easily be enlisted with scanning tools [2, 3]. In fact, several high-impact attacks have already been documented; the list includes the Stuxnet worm and the attacks against Iranian nuclear facilities [4], the compromise of a steel mill in Germany [5], and the cyber attacks on the Ukrainian power gird [6].

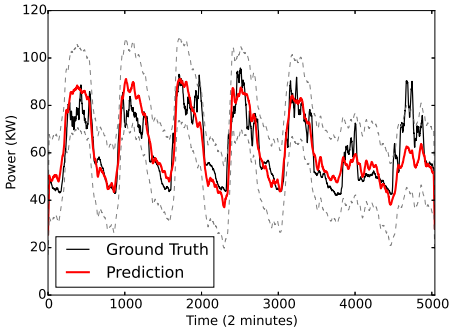

Fig. 1: Power prediction (with 95-percentile bounds).

Compromising an AMI meter can allow nefarious actors to spoof messages that carry power demand/supply values. Such coordinated *false data injection attacks* can endanger demand response mechanisms and compromise the grid's stability by misleading its state estimation process [7, 8]. Numerous scenarios of adversaries who compromise meters and fabricate their readings are discussed in [7, 9–11, 22].

In this article, we propose a statistical-based approach for tackling the problem of "bad data" injection in wide-area smart grid networks. Our *threat model* considers attackers that are restricted to accessing only specific sensors [12]. We focus on *AMI data* that convey information for electricity usage consumption (e.g., time-series of building power demand). Power usage is modeled via a linear factor model that aims to capture structural similarities in energy consumption between buildings (e.g., type of business, user behavior). Our model also captures spatial correlations amongst buildings that are "close" in space (e.g., due to similar weather conditions within a neighborhood, university campus, town, etc.). Therefore, we "borrow prediction strength" from a subset of metering nodes within the same area to forecast / predict the electricity usage in other locations.

The main contributions of this work are: a) the develop-

ment of an *adaptive* "*network kriging*" model[1] for predicting the energy usage of metering nodes based on observations from other, "trusted" nodes, within the same area, and b) a detection technique for identifying metering nodes that might be victims of bad data injection attacks. We assume that trusted readings involve nodes that transmit encrypted data and whose identity is authenticated [16].

2. DETECTION METHODOLOGY

Next, we describe our detection methodology. We start with data *clustering*, a step necessary for making our techniques scalable in situations when a large network needs to be monitored. The proposed *factor model* and *adaptive kriging* for anomaly detection are described in the sequel[2].

2.1. Building Grouping

We consider time-series of electricity usage of the form $Y(t) = (Y_i(t))_{i\in\mathcal{B}}$, where $\mathcal{B} = \{1, \cdots, B\}$ denotes the set of all buildings, and $t = 1, 2, \ldots$. For a monitoring window of size m, we define the $m \times B$ matrix of observations

$$\mathbf{D}(t_0, m) = [Y_i(t)]_{t_0-m\le t<t_0,\ i\in\mathcal{B}}. \tag{1}$$

To obtain a "signature" for each building, we partition our data into M windows of size m, and define the following matrix that corresponds to the empirical averages of (1), $\mu(t_0, m) = \frac{1}{M}\sum_{w=0}^{M-1}\mathbf{D}(t_0 - wm, m)$. We view the *columns* of $\mu(t_0, m)$ as points in a high-dimensional Euclidean space and apply standard clustering techniques, such as *K-means* [17]. Similar power utilization patterns in the columns of $\mu(t_0, m)$ suggest that the corresponding buildings are expected to lie within the same class.

2.2. Modeling Power Consumption via Linear Factors

We posit a parsimonious, *linear factor* model for modeling electricity consumption. Consider N samples of measurements $Y(t) \in \mathbb{R}^B$, and let $\mathbf{Q} = \sum_{n=1}^{N} Y(n)Y(n)^\top$ be a positive semidefinite $B \times B$ matrix. Performing *principal component analysis* (PCA) on this matrix, we obtain its *spectral decomposition* $\mathbf{Q} = \sum_{j=1}^{B}\lambda_j v_j v_j^\top$, where vectors $v_j \in \mathbb{R}^B$, $j = \{1, \ldots, B\}$, are orthonormal and $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_B \ge 0$. Using *real-world* AMI building data (see Section 3), we observed that much of the variability in the $Y(t)$'s can be explained by only few principal components of $\mathbf{Q}$. We therefore suggest the model,

$$Y(t) = \mu_Y(t) + Z(t) := \mathbf{F}\beta(t) + Z(t), \tag{2}$$

where $\mathbf{F}$ is a matrix $B \times k$ of factors, $\beta(t) \in \mathbb{R}^k$ is a parameter that can be estimated from available metering data (see next section), and $Z(t)$ is the measurement noise modeled as a multivariate normal distribution with zero mean and variance-covariance matrix $\mathbf{\Sigma}$. The k columns of the factor matrix $\mathbf{F}$ correspond to the $k \le B$ eigenvectors of $\mathbf{Q}$ with largest eigenvalues. Such choice of factor matrix yields the *best linear model* for capturing the temporal variability of the $Y(t)$'s (see [14], Prop. 1). To dynamically track temporal changes in electricity consumption, one can *adaptively* obtain the factors $\mathbf{F}$ and the variance-covariance matrix $\mathbf{\Sigma}$ over moving time windows (see Algorithm 1).

2.3. Kriging-based Prediction and Detection

With the modeling choice of (2), we now propose our anomaly detection methodology. Let $Y(t) = (Y_{i_1}(t), \cdots, Y_{i_b}(t))$, $\{i_1, \cdots, i_b\} \subset \mathcal{B}$ denote the electricity consumption of buildings within the same cluster. We partition meters into observed (trusted) nodes, $\mathcal{O} \subset \{i_1, \cdots, i_b\}$, and unobserved (untrusted) $\mathcal{U} = \{i_1, \cdots, i_b\}\backslash O$. Let $Y_o = (Y_j)_{j\in\mathcal{O}}$ and $Y_u = (Y_j)_{j\in\mathcal{U}}$ denote the partitioned vector Y (we drop time t to keep the notation unclattered). Thus, from (2),

$$\begin{pmatrix} Y_u \\ Y_o \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{F}_u\beta \\ \mathbf{F}_o\beta \end{pmatrix}, \begin{pmatrix} \mathbf{\Sigma}_{uu} & \mathbf{\Sigma}_{uo} \\ \mathbf{\Sigma}_{ou} & \mathbf{\Sigma}_{oo} \end{pmatrix}\right). \tag{3}$$

Given the limited set of observed nodes $\mathcal{O}$, and if the true parameter β is known, the minimum variance unbiased predictor of Y_u is the kriging estimate [14, 18]:

$$\hat{Y}_u(Y_o, \beta) := \mathbf{F}_u\beta + \mathbf{\Sigma}_{uo}\mathbf{\Sigma}_{oo}^{-1}(Y_o - \mathbf{F}_o\beta). \tag{4}$$

In practice, the parameter β is unknown, but can be estimated using linear regression from data on the observed nodes (see Eq. (2)). The *generalized least squares estimate*, $\hat{\beta}$, is

$$\hat{\beta} = (\mathbf{F}_o^\top\mathbf{\Sigma}_{oo}^{-1}\mathbf{F}_o)^{-1}\mathbf{F}_o^\top\mathbf{\Sigma}_{oo}^{-1}Y_o =: \mathbf{P}Y_o. \tag{5}$$

The *ordinary least squares estimator* $\hat{\beta} = (\mathbf{F}_o^\top\mathbf{F}_o)^{-1}\mathbf{F}_o^\top Y_o$ may alternatively be used. Henceforth, in our predictions for the unobserved nodes $\mathcal{U}$, the "plug-in" estimator $\hat{Y}_u(Y_o, \hat{\beta})$ is used. Using the expression for $\hat{\beta}$ from (5), $\hat{Y}_u$ simplifies to $\hat{Y}_u = \mathbf{F}_u\mathbf{P}Y_o + \mathbf{\Sigma}_{uo}\mathbf{\Sigma}_{oo}^{-1}(\mathbf{I} - \mathbf{F}_o\mathbf{P})Y_o$.

For detecting anomalies in the set of unobserved meters, we need the distribution of the prediction errors (residuals) between the *actual* meter readings and their *predictions*, i.e., $Y_e = Y_u - \hat{Y}_u$.

Proposition 2.1. *Under the Null hypothesis of no anomalies and the model of* (2)*, the prediction residuals Y_e follow a multivariate normal distribution $Y_e \sim N(0, \mathbf{\Sigma}_{err})$, with*

$$\mathbf{\Sigma}_{err} = \mathbf{\Sigma}_{uu} - \mathbf{C}\mathbf{\Sigma}_{ou} - \mathbf{\Sigma}_{uo}\mathbf{C}^\top + \mathbf{C}\mathbf{\Sigma}_{oo}\mathbf{C}^\top \tag{6}$$

and $\mathbf{C} = \mathbf{F}_u\mathbf{P} + \mathbf{\Sigma}_{uo}\mathbf{\Sigma}_{oo}^{-1}(\mathbf{I} - \mathbf{F}_o\mathbf{P})$.

Proof. Observe that $Y_e = Y_u - \mathbf{C}Y_o$ is a linear transformation of Y, and therefore Y_e has a multivariate normal distribution.

[1] In our context, *network kriging* refers to statistical prediction of smart meter electricity consumption, based on observations from other meters within the network. The term was first introduced in [13]. See also [14, 15].

[2] Throughout, matrices are denoted with bold uppercase symbols, vectors with uppercase, and scalars with lowercase, except if noted otherwise.

Algorithm 1 Kriging for Detection of Data Injection Attacks

Input: Training data $\mathcal{D}(t_0) := \{Y_i(t), i \in \mathcal{B}, t_0 - N \leq t \leq t_0\}$;
Input: Set of "observed" nodes $\mathcal{O}$;
Input: Set of "unobserved" nodes $\mathcal{U} = \{1, \ldots, b\} \backslash \mathcal{O}$;
Output: Sequence of p-values for prediction errors.

1: Obtain $b \times k$ factor matrix $\mathbf{F}$ using PCA on data $\mathcal{D}(t_0)$
2: Estimate covariance matrix $\mathbf{\Sigma}$ using data $\mathcal{D}(t_0)$
3: **for** each new observation $Y = Y(t), t = t_0 + 1, \ldots$ **do**
4: Partition vector Y into Y_o and Y_u
5: Estimation of $\hat{\beta} = (\mathbf{F}_o^\top \mathbf{\Sigma}_{oo}^{-1} \mathbf{F}_o)^{-1} \mathbf{F}_o^\top \mathbf{\Sigma}_{oo}^{-1} Y_o = \mathbf{P} Y_o$.
6: *Prediction:* $\hat{Y}_u = \mathbf{F}_u \hat{\beta} + \mathbf{\Sigma_{uo}} \mathbf{\Sigma_{oo}}^{-1} (Y_o - \mathbf{F}_o \hat{\beta})$
7: Calculate the error covariance matrix $\mathbf{\Sigma}_{\text{err}}$ (see Eq. (6))
8: With prediction error $Y_e := Y_u - \hat{Y}_u$, get test statistic

$$r^2 = Y_e^\top \mathbf{\Sigma}_{\text{err}}^{-1} Y_e \text{ (Mahalanobis distance)}$$

9: **output** p=1-$F(r^2)$, $F(x)$ is a chi-squared cdf (d.f.= $|\mathcal{U}|$).
10: **end for**

The expected error, $\mu_{\text{err}} = \mathbb{E}[Y_u - \mathbf{C}Y_o]$, becomes $\mu_{\text{err}} = \mathbf{F}_u \mathbb{E}[\beta] - \mathbf{C}\mathbf{F}_o \mathbb{E}[\beta]$ from (2),(3). $\mathbf{P}\mathbf{F}_o = \mathbf{I}$, which implies $\mathbf{C}\mathbf{F}_o = \mathbf{F}_u \mathbf{P}\mathbf{F}_o + \mathbf{\Sigma}_{uo}\mathbf{\Sigma}_{oo}^{-1}(\mathbf{I} - \mathbf{F}_o\mathbf{P})\mathbf{F}_o = \mathbf{F}_u$, and, thus, $\mu_{err} = 0$. For the error variance, $\text{Var}(Y_u - \mathbf{C}Y_o) = \mathbb{E}\big[(Y_u - \mathbf{C}Y_o)(Y_u - \mathbf{C}Y_o)^\top\big]$, and the result follows using (3). □

The vector of prediction errors Y_e is employed to diagnose anomalies on unobserved meters. We perform a *single* test (rather than running individual tests on each error component) using the *statistic* $r^2 = Y_e^\top \mathbf{\Sigma}_{\text{err}}^{-1} Y_e$, which corresponds to the Mahalanobis distance. In particular, we obtain the p-value, namely $p = 1 - F(r^2)$ where $F(x)$ is the chi-squared cumulative distribution function with degrees of freedom equal to rank($\mathbf{\Sigma}_{\text{err}}$). Algorithm 1 summarizes our methodology. To tame the false alarm rate, we apply an *exponential weighted moving average* (EWMA) control chart to the standardized z-scores $z = \Phi^{-1}(1-p)$, where $\Phi(x)$ is the cumulative distribution for standard normal. With EWMA, the sequence of z-scores $\{z(t)\}$ is smoothed, and an alert is raised when $s(t) = wz(t) + (1-w)s(t-1), 0 < w \leq 1$ gets out of control (i.e., $|s(t)| > L\sigma_s, \sigma_s = \sqrt{w/(2-w)}$ [19,20]).

3. PERFORMANCE EVALUATION

We use *real-world* smart meter data from a large university campus to evaluate our methods. Our dataset includes time-series for the power usage of 154 buildings (we exclude 9 buildings with unreliable data) at a granularity of 2 minutes. The dataset includes buildings for student housing, lecture halls, laboratories and offices, parking structures, buildings for health services, etc.

Model validation using the distribution of p-values is depicted in Fig. 2. We employ our detection algorithm on data of an AMI meter during an *anomaly-free* measurement period, and obtain a sequence of p-values. Fig. 2 plots the distribution of those values, and the *uniformity* thereof suggests that the selected model should be adequate for the task at hand (a similar conclusion was drawn with simulated data too).

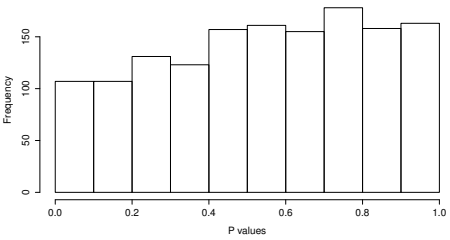

Fig. 2: Model validation (real-world data)

Table 1: Detection performance (meter 1) based on average precision and recall (standard deviation in parentheses).

EWMA (w, L)	Shift ($\times\sigma_1$) (Watts)	Precision	Recall
(1, 3.719)	1	$.00_{(.00)}$	$.00_{(.00)}$
(1, 3.719)	2	$.74_{(.44)}$	$.56_{(.39)}$
(1, 3.719)	3	$.96_{(.20)}$	$.92_{(.19)}$
(.53, 3.714)	1	$.15_{(.25)}$	$.13_{(.23)}$
(.53, 3.714)	2	$.69_{(.07)}$	$.92_{(.08)}$
(.53, 3.714)	3	$.69_{(.06)}$	$.98_{(.01)}$
(.84, 3.719)	1	$.06_{(.24)}$	$.02_{(.09)}$
(.84, 3.719)	2	$.84_{(.37)}$	$.70_{(.35)}$
(.84, 3.719)	3	$.96_{(.14)}$	$.95_{(.14)}$

Performance in the presence of injected anomalies is tabulated in Table 1. We obtain the factors and the variance-covariance matrix of Algorithm 1 using a two-week long learning period (weekend days excluded), and apply our detection method for the next 48 hours. We employ EWMA with (w, L) as shown (design option (1, 3.719) avoids smoothing, (.53, 3.714) is more sensitive to outliers and (.84, 3.719) is less amenable to false positives). To tame the false alarm rate, we leverage the *two-in-a-row* rule [21]. Further, an alarm is raised only if the EWMA statistic exceeds its control limits at least 15 times over the past hour. We inject one simulated attack on a randomly chosen epoch (lasting one hour), and perform 50 independent experiments. We assess the detection performance in terms of average *precision* and *recall* (see [20]). The attack magnitude is a usage shift (in Watts) proportional to the standard deviation of the building under study. In Table 1, one building is considered "unobserved" (a student dormitory, see Fig. 1), and the rest are considered trusted. We observe that the proposed method is not sensitive enough to detect the low-volume attack of σ_1 Watts. On the other hand, precision and recall are high for the larger data attacks.

To obtain further insights into detection accuracy we can examine Fig. 3 (left). All buildings are examined. When meter-i is considered unobserved, we tamper its reading with an hour-long data attack of σ_i Watts, injected around time 1300. In the majority of cases, our method detects this low-volume event, and false positives are relatively low. However, the fact that for few buildings (see also Table 1) factor-kriging is unable to detect small attacks or raises excessive false alerts indicates that such buildings should be studied in isolation using alternative models, especially when detection of stealth attacks is necessary (such as energy theft [22, 23]).

Fig. 3 (middle) presents the prediction performance when varying the number of trusted buildings. We display the *root*

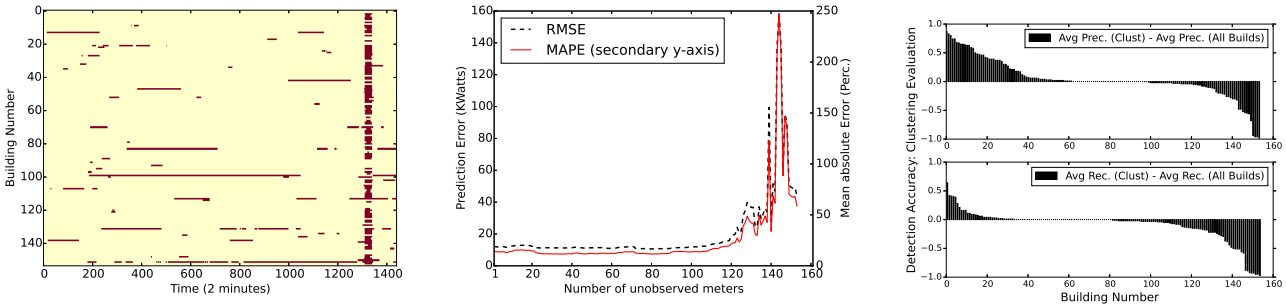

Fig. 3: *Left:* Detection alerts (red) over a two-day period. The vertical red stripe denotes a 60min period of injected anomalies. We study the behavior of each building; the building under study is consider as unobserved (unsecured) and we use observations from the remaining ones. EWMA pair $(1, 3.719)$ is used. *Middle:* Prediction performance (for building 1) as the number of observed nodes decreases. *Right:* The effect of clustering in detection performance. (Due to sorting, the building orderings in the top and bottom panels differ.).

Table 2: Silhouette values for cluster number selection.

Cluster number	**2**	3	4	5	6	7	8	9
Silhouette score	**.81**	.59	.60	.56	.52	.47	.43	.45

mean square error RMSE $= \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_i(t) - \hat{Y}_i(t))^2}$, and *mean absolute percentage error* MAPE $= \frac{1}{T}\sum_{t=1}^{T}|(Y_i(t) - \hat{Y}_i(t))/Y_i(t)|$, where i is the unobserved building and T the prediction period. We observe that even with few trusted buildings (e.g., about 1/3 of total) the prediction accuracy in terms of RMSE / MAPE is not deteriorating substantially.

To obtain a comparison baseline, we also study a detection method based on autoregressive modeling. In particular, the power consumption of each building is modeled via an AR(1) process. The AR(1) model performs better than kriging in forecasting consumption, but its detection performance (precision/recall) is inferior. AR(1) is able to detect the onset of the attack, but then "adapts" to bad data and missing most of the remaining ones. On the contrary, kriging has the advantage of testing for "structural" changes using only trusted nodes and is, thus, immune to contaminated readings.

Finally, we assess the performance of our methods when clustering is employed. Fig. 3 (right) examines the detection accuracy with and without clustering. The number of K-means clusters is selected by looking at the silhouette [24] scores (see Table 2); two main classes are identified. Even though, for some buildings, information from *all* meters provides better detection results, a good clustering algorithm can improve detection performance and scalability.

4. DISCUSSION AND FUTURE DIRECTIONS

We present a behavioral-based detection method for tracking false data injection attacks in the smart grid. We consider data for power consumption from meters within a wide-area network. Our system detects nefarious meter activity via a factor-based kriging model; the electricity usage of the monitored meters is forecasted using measurements from ones whose integrity and identity is "trusted".

Methodologies for anomaly detection in the electric grid can in general be categorized into signature-, specification- and anomaly-based methods [25–27]. The proposed approach falls within the latter category and complements the former ones. For example, new attacks with signatures agnostic to signature-based intrusion detection system (e.g., *Snort*) would always evade detection. At the same time, specification-based systems [28] can be cumbersome to fine-tune (e.g., finding a valid range for the AMI demand / supply is not easily determined). Existing defenses against smart grid data attacks to impede its state estimation appear in [11, 16, 20, 29]. In [20], the problem of detecting aberrant behavior of residential smart meters is tackled from the *home-area network* perspective. [29] proposes an adaptive cumulative sum test combined with a multivariate hypothesis testing problem to prevent an erroneous grid-state estimate. [16] studies a graph theoretic method for securing an optimal set of meter measurements so that state estimation is not compromised. [11] couples anomaly-based methods with a data integrity check to combat stealth attacks. [30] sheds light into situations of *multiple adversaries* performing injection attacks, and discusses optimal defense strategies from game theory.

Contrary to the related work in [8, 11, 12, 16, 29], we emphasize that our method does not require parameter knowledge of the system's *DC power flow model* [8, 12]; instead, the power utilization of the AMI meters is the sole data input. This implies that having an anomaly-free learning period is of paramount importance. However, since some nodes are considered unsecured, data contamination even in the training period is likely. In order to alleviate this we plan to study robust methods for obtaining the factors (e.g., robust subspace learning [31]) and performing the kriging step [32]. Ongoing works includes the incorporation of temporal information into our model, and investigation of alternative clustering techniques and features for building grouping.

Acknowledgements: Work supported by NSF CNS-1422078.

5. REFERENCES

[1] H. Farhangi, "The path of the smart grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, January 2010.

[2] Z. Durumeric, E. Wustrow, and J. A. Halderman, "Zmap: Fast internet-wide scanning and its security applications," in *USENIX SEC'13*, 2013, pp. 605–620.

[3] Z. Durumeric et al., "A search engine backed by internet-wide scanning," in *ACM CCS '15*, 2015.

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

[5] R. Lee, M. Assante, and T. Conway, "German steel mill cyber attack," 2014.

[6] R. Lee, M. Assante, and T. Conway, "Analysis of the cyber attack on the ukrainian power grid," 2016.

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

[8] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *ACM CCS '09*, 2009.

[9] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE S & P*, vol. 7, 2009.

[10] National SCADA Test Bed and US DOE, "Study of Security Attributes of Smart Grid Systems - Current Cyber Security Issues," April 2009.

[11] W. Yu et al., "An integrated detection system against false data injection attacks in the smart grid," *Security and Communication Networks*, vol. 8, no. 2, 2015.

[12] R. B. Bobba et al., "Detecting false data injection attacks on DC state estimation," in *SCS Workshop*, 2010.

[13] D. B. Chua, E. D. Kolaczyk, and M. Crovella, "Network kriging," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 12, pp. 2263–2272, Dec 2006.

[14] J. Vaughan, S. Stoev, and G. Michailidis, "Network-wide statistical modeling, prediction, and monitoring of computer traffic," *Technometrics*, vol. 55, no. 1, 2013.

[15] M.G. Kallitsis, S.A. Stoev, and G. Michailidis, "Fast algorithms for optimal link selection in large-scale network monitoring," *Signal Processing, IEEE Transactions on*, vol. 61, no. 8, pp. 2088–2103, 2013.

[16] 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.

[17] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis,*, John Wiley and Sons, Inc., Hoboken, NJ, USA, 1990.

[18] N. Cressie, *Statistics for Spatial Data*, John Wiley, New York, USA, 1993.

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

[20] M. G. Kallitsis, G. Michailidis, and S. Tout, "Correlative monitoring for detection of false data injection attacks in smart grids," in *IEEE SmartGridComm*, Nov. 2015.

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

[22] S. McLaughlin et al., "A multi-sensor energy theft detection framework for advanced metering infrastructures," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1319–1330, July 2013.

[23] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Electricity theft detection in ami using customers's consumption patterns," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 216–226, Jan 2016.

[24] P. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, no. 1, pp. 53–65, Nov. 1987.

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

[26] F.M. 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.

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

[28] A. Carcano et al., "State-based network intrusion detection systems for SCADA protocols: A proof of concept," in *Proceedings of CRITIS'09*, 2010, pp. 138–150.

[29] Yi Huang et al., "Bad data injection in smart grid: attack and defense mechanisms," *Communications Magazine, IEEE*, Jan. 2013.

[30] A. Sanjab and W. Saad, "Smart grid data injection attacks: To defend or not?," in *2015 IEEE SmartGridComm*, Nov 2015.

[31] F. De La Torre and M. J. Black, "A framework for robust subspace learning," *Int. J. Comput. Vision*, vol. 54, no. 1-3, pp. 117–142, Aug. 2003.

[32] B. Baingana et al., "Robust kriged kalman filtering," in *49th Asilomar Conference on Signals, Systems and Computers*, Nov 2015, pp. 1525–1529.