QUICKEST DETECTION OF TIME-VARYING FALSE DATA INJECTION ATTACKS IN DYNAMIC SMART GRIDS

Jiangfan Zhang

Dept. of ECE, Missouri University of Science and Technology, Rolla, MO 65409 USA
Email: jiangfanzhang@mst.edu

ABSTRACT

Quickest detection of false data injection attacks (FDIAs) in dynamic smart grids is considered in this paper. The unknown time-varying state variables of the smart grid and the FDIAs impose a significant challenge for designing a computationally efficient detector. To address this challenge, we propose new Cumulative-Sum-type algorithms with computational complex scaling linearly with the number of meters. Moreover, for any constraint on the expected false alarm period, a lower bound on the threshold employed in the proposed algorithm is provided. For any given threshold employed in the proposed algorithm, an upper bound on the worst-case expected detection delay is also derived. The proposed algorithm is numerically investigated in the context of an IEEE standard power system under FDIAs, and is shown to outperform some representative algorithm in the test case.

***Index Terms*—** Cybersecurity, cumulative sum, false data injection attacks, dynamic smart grid systems.

1. INTRODUCTION

In a smart grid, the meter measurements are collected and employed at a control center to estimate state variables of the smart grid, such as bus voltages and phase angles, and then the operation of the smart grid is performed and controlled based on these estimated states. If any adversary can falsify the meter measurements, the control center may produce erroneous state estimates, which give rise to wrong decisions on billing, power dispatch, and even blackout. In light of this, it is of paramount importance for modern smart grid systems to have the capability of detecting malicious attacks as quickly as possible. The sequential change detection (also known as quickest detection), which minimizes the expected detection delay subject to certain constraint on the average false alarm period, enables online monitoring for smart grid systems, and therefore suits well to attack detection in such systems.

We assume the measurements from a set of meters are corrupted by additive malicious data. Such attacks are referred to as false data injection attacks (FDIAs), which are considered as one of most detrimental attacks to smart grid systems [1]. The set of attacked meters and the injected malicious data are time-varying and unknown to the control center. In addition, we assume that the state variables of the smart grid system are dynamic, and also unknown to the control center. It is worth mentioning that we don't make any assumption on how the state variables of the system evolve over time. The control center aims at detecting any FDIAs as soon as possible when they are launched, and the quickest detection scheme is the focus of this paper.

In [2], an adaptive Cumulative Sum (CUSUM) algorithm is proposed which builds on the assumption that the state variables of the smart grid follow a Gaussian prior and the FDIA at any time is always positive and small. Another sequential algorithm based on the Rao test statistic is proposed for static smart grid systems in [3]. The quickest detection of FDIA in smart grids is investigated in [4], where the set of effective attacked meters is assumed to be fixed over time. More recently, in [5], a CUSUM-type algorithm based on the Kalman filter is proposed, which assumes that the state variables of the smart grid evolve over time by following a fixed linear model. In this paper, we consider a more general model, and none of these approaches can be applied here, since no such assumptions made in [2–5] are made in this paper. Moreover, no performance analysis is provided for the approaches in [2–5], while we pay more attention to the performance characterization of the proposed approach with the aim of providing provable detection performance guarantee in this paper. The main contributions are summarized as follows.

(1) We propose new CUSUM-type algorithms which are robust to arbitrarily time-varying state variables and arbitrary FDIAs. Moreover, the computational complexity of the proposed algorithms just scales linearly with the number of meters in the power system.

(2) For any constraint on the expected false alarm period, a lower bound on the threshold employed in the proposed algorithm is derived, which provides a guideline for the design of the proposed algorithm to achieve the prescribed performance requirement..

(3) For any given threshold employed in the proposed algorithm, an upper bound on the worst-case expected detection delay is provided.

2. PROBLEM STATEMENT

Consider M meters in an $(N+1)$-bus smart grid system. Let $\boldsymbol{\theta}^{(t)} \in \mathbb{R}^N$ denote the time-varying N phase angles (one reference angle) at time t, and let $\mathbf{x}^{(t)} \in \mathbb{R}^M$ denote the measurements of the power flows and power injections at the M meters at time instant t. Then the dynamic direct current (DC) power flow model of the system can be formulated as

$$\mathbf{x}^{(t)} = \mathbf{H}\boldsymbol{\theta}^{(t)} + \mathbf{n}^{(t)} \tag{1}$$

where $\mathbf{H} \in \mathbb{R}^{M\times N}$ is the measurement matrix which depends on the topology of the smart grid, the placement of the meters, and the susceptance of each transmission line. Typically, the number of measurements is greater than that of the unknown parameters in order to provide necessary redundancy against the noise effect, i.e., $M > N$. In addition, we assume that $\{\mathbf{n}^{(t)}\}$ is a sequence of independent and identically distributed (i.i.d.) noise vectors obeying Gaussian distribution with $\mathbf{0}$ mean and covariance $\sigma^2\mathbf{I}_M$.

Suppose that at time t_a, a malicious attacker intentionally manipulates the observation vector $\mathbf{x}^{(t)}$ by injecting a sequence of unknown false data $\{\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \mathbf{b}^{(3)}, ...\}$ to the smart grid. Accordingly, we write the attack-incurred observation change as

$$\begin{cases} \mathbf{x}^{(t)} = \mathbf{H}\boldsymbol{\theta}^{(t)} + \mathbf{n}^{(t)}, & \text{if } t < t_a, \\ \mathbf{x}^{(t)} = \mathbf{H}\boldsymbol{\theta}^{(t)} + \mathbf{a}^{(t)} + \mathbf{n}^{(t)} & \text{if } t \geq t_a, \end{cases} \tag{2}$$

where $\mathbf{a}^{(t_a+t-1)} = \mathbf{b}^{(t)}$ for any t_a and $t \geq 1$. Note that the injected false data $\mathbf{a}^{(t)}$ can be decomposed into two parts $\mathbf{a}^{(t)} = \mathbf{H}\mathbf{c}^{(t)} + \boldsymbol{\mu}^{(t)}$ where $\mathbf{H}\mathbf{c}^{(t)}$ denotes the component of $\mathbf{a}^{(t)}$ that lies in the column space of $\mathbf{H}$, while $\boldsymbol{\mu}^{(t)}$ represents the component of $\mathbf{a}^{(t)}$ that lies in the complementary space $\mathcal{R}^{\perp}(\mathbf{H})$ of the column space of $\mathbf{H}$, that is

$$\boldsymbol{\mu}^{(t)} = \mathbf{P}_{\mathbf{H}}^{\perp}\mathbf{a}^{(t)} \in \mathcal{R}^{\perp}(\mathbf{H}) \tag{3}$$

where $\mathbf{P}_{\mathbf{H}}^{\perp} \triangleq \mathbf{I} - \mathbf{H}\big(\mathbf{H}^T\mathbf{H}\big)^{-1}\mathbf{H}^T$. As demonstrated in [6], $\boldsymbol{\mu}^{(t)}$ is the only informative part of the injected false data that is detectable. The reason is that since the parameter vector $\boldsymbol{\theta}^{(t)}$ is unknown, the other part $\mathbf{H}\mathbf{c}^{(t)}$ of $\mathbf{a}^{(t)}$ is not distinguishable from $\mathbf{H}\boldsymbol{\theta}^{(t)}$, and hence can bypass any monitoring system.

Let ρ_L and ρ_U denote the lower and upper bounds on the magnitudes of the nonzero elements of $\boldsymbol{\mu}^{(t)}$, respectively. The constant ρ_L indicates the minimal magnitude of $\boldsymbol{\mu}^{(t)}$ that draws security concerns, and the constant ρ_U represents the limited power of the adversaries. Let $\mathcal{A}^{(t)}$ represent the set of nonzero elements of $\boldsymbol{\mu}^{(t)}$ at time instant t. As such, we can write the attack-incurred change event of interest as

$$\begin{aligned} t < t_a: &\quad \mu_m^{(t)} = 0, m = 1, 2, ..., M, \\ t \geq t_a: &\quad \begin{cases} \rho_L \leq |\mu_m^{(t)}| \leq \rho_U, m \in \mathcal{A}^{(t)}, \\ \mu_m^{(t)} = 0, m \notin \mathcal{A}^{(t)}, \end{cases} \end{aligned} \tag{4}$$

where $\mu_m^{(t)}$ is the m-th element of $\boldsymbol{\mu}^{(t)}$. It is worth mention that since $\boldsymbol{\mu}^{(t)}$ can be time-varying, the set $\mathcal{A}^{(t)}$ can also be distinct over time. The quickest detection technique, that exploits the statistical difference before and after t_a, provides a suitable framework to achieve this goal. The commonly used performance measure, proposed by Lorden, is the worst-case expected detection delay which is defined as [7]

$$J(T) \triangleq \sup_{t_a} \operatorname*{ess\,sup}_{\mathcal{F}_{t_a-1}} \mathbb{E}_{t_a}\left\{ (T_R - t_a + 1)^+ \Big| \mathcal{F}_{t_a-1} \right\}, \tag{5}$$

where the random variable T is a stopping time corresponding to a certain sequential detection scheme and $\mathcal{F}_{t_a}$ is the filtration generated by all the observations up to time t_a. The expectation $\mathbb{E}_{t_a}$ is evaluated with respect to the true distribution of $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$, $\mathbf{x}^{(3)}$, ... when the attack occurs at time instant t_a. The quickest detection problem is formulated as follows:

$$\inf_T J(T) \text{ subject to } \mathbb{E}_\infty\{T\} \geq \gamma. \tag{6}$$

Note that the expectation $\mathbb{E}_\infty$ is evaluated with respect to the probability measure where $t_a = \infty$, i.e., no attack occurs, and γ is a prescribed constant which specifies the required lower bound on the expected false alarm period. Before proceeding, we denote the pre-attack and post-attack probability density functions of the observation $\mathbf{x}^{(t)}$ as $f_u(\mathbf{x}^{(t)} \big| \boldsymbol{\theta}^{(t)})$ and $f_a(\mathbf{x}^{(t)} \big| \boldsymbol{\theta}^{(t)}, \mathbf{a}^{(t)})$, respectively. If all the parameters $\{\boldsymbol{\theta}^{(t)}\}$ and $\{\mathbf{a}^{(t)}\}$ are known, the quickest detection problem in (6) is optimally solved by the well-known CUSUM test [8]

$$T_{\mathrm{C}} = \min\left\{ K : \max_{1\leq k\leq K} \sum_{t=k}^{K} \ln \frac{f_a\left(\mathbf{x}^{(t)} \big| \boldsymbol{\theta}^{(t)}, \mathbf{a}^{(t)}\right)}{f_u\left(\mathbf{x}^{(t)} \big| \boldsymbol{\theta}^{(t)}\right)} \geq h \right\} \tag{7}$$

where the threshold h is determined by the constraint in (6).

However, in the problem considered in this paper, the parameters $\{\boldsymbol{\theta}^{(t)}\}$ and $\{\mathbf{a}^{(t)}\}$ are unknown, causing the CUSUM test infeasible. To address this, this paper resort to the generalized likelihood ratio (GLR) method by replacing the unknown parameters with their maximum likelihood estimates (MLE) [9, 10].

3. RELAXED GENERALIZED CUSUM TEST

3.1. Generalized CUSUM Test

Based on the model of the attack-incurred change event in (4), by replacing the unknown parameters $\{\boldsymbol{\theta}^{(t)}\}$ and $\{\mathbf{a}^{(t)}\}$ with their MLEs in (7), the generalized CUSUM (GCUSUM) test can be written as

$$T_G = \min\left\{ K : \max_{1\leq k\leq K} \sup_{\{\mathcal{A}^{(t)}\}} \Lambda_k^{(K)} \geq h \right\}, \tag{8}$$

where the statistic $\Lambda_k^{(K)}$ is defined in (9) on the top of the next page. Considering that $\{\mathbf{n}^{(t)}\}$ is i.i.d. white Gaussian

$$\Lambda_k^{(K)} \triangleq \ln \frac{\sup\limits_{\boldsymbol{\theta}^{(t)},\mathbf{a}^{(t)}:\{\rho_L\le|\mu_m^{(t)}|\le\rho_U\}_{m\in\mathcal{A}^{(t)}},\boldsymbol{\mu}^{(t)}\in\mathcal{R}^{\perp}(\mathbf{H})} \prod\limits_{t=1}^{k-1} f_u\left(\mathbf{x}^{(t)}\left|\boldsymbol{\theta}^{(t)}\right.\right) \prod\limits_{t=k}^{K} f_a\left(\mathbf{x}^{(t)}\left|\boldsymbol{\theta}^{(t)},\mathbf{a}^{(t)}\right.\right)}{\sup\limits_{\boldsymbol{\theta}^{(t)}} \prod\limits_{t=1}^{K} f_u\left(\mathbf{x}^{(t)}\left|\boldsymbol{\theta}^{(t)}\right.\right)} = \sum_{i=k}^{K} \Lambda_{k,t}^{(K)}. \tag{9}$$

$$\Lambda_{k,t}^{(K)} = \sup_{\boldsymbol{\mu}^{(t)}:\{\rho_L\le|\mu_m^{(t)}|\le\rho_U\}_{m\in\mathcal{A}^{(t)}},\boldsymbol{\mu}^{(t)}\in\mathcal{R}^{\perp}(\mathbf{H})} \frac{1}{2\sigma^2} \sum_{m\in\mathcal{A}^{(t)}} \left[2\mu_m^{(t)}\tilde{x}_m^{(t)} - \left(\mu_m^{(t)}\right)^2\right]. \tag{10}$$

noise with zero mean and covariance $\sigma^2\mathbf{I}_M$, the $\Lambda_{k,t}^{(K)}$ in (9) can be simplified to (10) on the top of this page, where $\tilde{x}_m^{(t)}$ is the m-th elements of $\tilde{\mathbf{x}}^{(t)}$, and $\tilde{\mathbf{x}}^{(t)}$ is the component of $\mathbf{x}^{(t)}$ in the complementary space of the column space of $\mathbf{H}$, i.e., $\tilde{\mathbf{x}}^{(t)} \triangleq \mathbf{P}_{\mathbf{H}}^{\perp}\mathbf{x}^{(t)}$.

It is seen from (10) that there is no closed-form expression for $\Lambda_{k,t}^{(K)}$ in general. Hence, to obtain the decision statistic in (8), we need to numerically obtain $\Lambda_{k,t}^{(K)}$ for each $\mathcal{A}^{(t)}$, and then maximize $\Lambda_k^{(K)}$ over all possible $\mathcal{A}^{(t)}$ as illustrated in (8). Since the number of possible $\mathcal{A}^{(t)}$ is on the order of 2^M, it is not feasible to implement the GCUSUM test in (8) in practice especially when M is large, which motives us to pursue more computationally efficient algorithms.

3.2. Relaxed GCUSUM Test

In order to facilitate the computation of $\Lambda_{k,t}^{(K)}$, we relax the constraint $\boldsymbol{\mu}^{(t)} \in \mathcal{R}^{\perp}(\mathbf{H})$ in (10), and correspondingly, $\Lambda_{k,t}^{(K)}$ can be bounded from above as per

$$\begin{aligned}
&\Lambda_{k,t}^{(K)} \\
&\le \sup_{\boldsymbol{\mu}^{(t)}:\{\rho_L\le|\mu_m^{(t)}|\le\rho_U\}_{m\in\mathcal{A}^{(t)}}} \frac{1}{2\sigma^2} \sum_{m\in\mathcal{A}^{(t)}} \left[2\mu_m^{(t)}\tilde{x}_m^{(t)} - \left(\mu_m^{(t)}\right)^2\right] \\
&= \sum_{m\in\mathcal{A}^{(t)}} \zeta_m^{(t)} \triangleq \widetilde{\Lambda}_{k,t}^{(K)},
\end{aligned} \tag{11}$$

where

$$\zeta_m^{(t)} \triangleq \begin{cases} \frac{1}{2\sigma^2}\left(\tilde{x}_m^{(t)}\right)^2 & \text{if } \rho_L \le \left|\tilde{x}_m^{(t)}\right| \le \rho_U, \\ \frac{1}{2\sigma^2}\left(2\left|\tilde{x}_m^{(t)}\right|\rho_L - \rho_L^2\right) & \text{if } \left|\tilde{x}_m^{(t)}\right| < \rho_L, \\ \frac{1}{2\sigma^2}\left(2\left|\tilde{x}_m^{(t)}\right|\rho_U - \rho_U^2\right) & \text{if } \left|\tilde{x}_m^{(t)}\right| > \rho_U. \end{cases} \tag{12}$$

As a result, by replacing $\Lambda_{k,t}^{(K)}$ with $\widetilde{\Lambda}_{k,t}^{(K)}$ in (9), a relaxed generalized CUSUM (RGCUSUM) test can be expressed as

$$\begin{aligned}
T_R &= \min\left\{K : \max_{1\le k\le K} \sup_{\{\mathcal{A}^{(t)}\}} \sum_{t=k}^{K} \widetilde{\Lambda}_{k,t}^{(K)} \ge h\right\} \\
&= \min\left\{K : \sum_{t=1}^{K}\sum_{m=1}^{M} \max\left\{\zeta_m^{(t)}, 0\right\} \ge h\right\}.
\end{aligned} \tag{13}$$

It is seen from (12) and (13) that the RGCUSUM is more amenable than the GCUSUM to implementation in practice, since the computational complexity of the RGCUSUM just scales linearly with the number M of meters.

4. PERFORMANCE ANALYSIS OF THE RGCUSUM

In this section, we provide the performance analysis of our proposed RGCUSUM test. In particular, we provide a sufficient condition under which the constraint in (6) can be guaranteed, which sheds insight into the design of the proposed RGCUSUM test to achieve the prescribed performance requirement. Moreover, an upper bound on the worst-case expected detection delay defined in (5) is derived for any h.

Let $\mathbf{p}_i^T$ denote the i-th row of the projection matrix $\mathbf{P}_{\mathbf{H}}^{\perp}$, i.e., $\mathbf{P}_{\mathbf{H}}^{\perp} = [\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_M]^T$. As demonstrated in (6), the expected running length of the RGCUSUM under no attack needs to be guaranteed to be larger than the required lower bound γ to avoid frequent false alarms. In general, we can set the threshold h in (13) to be sufficiently large so that the constraint in (6) is satisfied. In the following, we provide a sufficient condition on h which can guarantee that the expected false alarm period of the RGCUSUM is larger than the prescribed γ.

Theorem 1 *The constraint on the expected false alarm period in (6), i.e., $\mathbb{E}_\infty\{T\} \ge \gamma$, is satisfied provided that*

$$h \ge \gamma \sum_{m=1}^{M} \left(\frac{1}{2}\|\mathbf{p}_m\|_2^2 + \frac{\rho_L+\rho_U}{\sigma}\|\mathbf{p}_m\|_2\sqrt{\frac{2}{\pi}}\right). \tag{14}$$

The proof of Theorem 1 is omitted due to space constraints. It is worth mentioning that the lower bound on h in (14) is only determined by the projection operator $\mathbf{P}_{\mathbf{H}}^{\perp}$, the variance of the noise, and the prescribed lower and upper bounds on the magnitude of the nonzero elements of $\boldsymbol{\mu}^{(t)}$. Therefore, when γ is given, the lower bound on h in (14) can be calculated beforehand, and then employed in the RGCUSUM. Hence, Theorem 1 provides a guideline for the design of the proposed RGCUSUM to achieve the prescribed performance requirement.

Besides the expected false alarm period, another key performance measure for quickest detection is the worst-case expected detection delay defined in (5). We have the following

theorem regarding the worst-case expected detection delay of the proposed RGCUSUM.

Theorem 2 *By employing Wald's approximations [10], i.e., ignoring the expectation of the overshoots in the presence of attacks, for any given h, the worst-case expected detection delay of the RGCUSUM can be bounded from above as per*

$$J\left(T_R\right) \leq h\left\{\sum_{m=1}^{M} \frac{\rho_L^2}{2\sigma^2}\left[\operatorname{erf}\left(\frac{2\rho_U}{\sqrt{2}\sigma\|\mathbf{p}_m\|_2}\right) - \operatorname{erf}\left(\frac{\rho_L+\rho_U}{\sqrt{2}\sigma\|\mathbf{p}_m\|_2}\right)\right]\right\}^{-1}, \quad (15)$$

where the function $\operatorname{erf}(x) \triangleq \frac{2}{\sqrt{\pi}}\int_0^x e^{-s^2} ds$.

The proof of Theorem 2 is omitted due to space constraints, which is based on the fact that the RGCUSUM can be shown to have the equalizer rule. As demonstrated in Theorem 2, the worst-case expected detection delay of the RGCUSUM can be bounded by a term which is proportional to the threshold h. It is worth mentioning that Wald's approximations are employed in Theorem 2, which implicitly assumes that the expectation of the overshoot should be negligibly small when compared to the threshold h. We have numerically examined the validity of the Wald's approximations in some practical cases, and the numerical results show that the Wald's approximations are valid when h is large.

5. NUMERICAL RESULTS

In this section, we consider the standard IEEE-14 bus power system to test the performance of the proposed RGCUSUM, and the measurement matrix $\mathbf{H}$ in (1) is determined accordingly for the DC model of the power system. The initial state of the power system is defined in the MATPOWER "case14" [11]. We assume that the resistive load at one bus decreases by 100 watts per time instant, while the resistive loads at other two buses increase by 100 watts per time instant. As such, the state variables of the smart grid evolve accordingly over time. Although the proposed RGCUSUM can be applied to the cases where the FDIA $\mathbf{a}^{(t)}$ is time-varying, for simplicity, $\mathbf{a}^{(t)}$ is set to be a constant vector in the simulations, which is

$$\mathbf{a}^{(t)} = [-2.629, -2.704, 2.781, 2.923, 0.516, -0.936, 1.969, -3.938, -0.033, 0, -0.483, -0.033, -1.934, 1.934, -1.934, 4.259, 2.842, 0.110, 1.314, -0.520, 2.195, -0.046, 1.778]^T.$$

In Fig. 1, we scrutinize the performance of the proposed RGCUSUM, and compare it with that of a representative approach, called adaptive CUSUM algorithm proposed in [2]. In the simulation, the variance of noise $\sigma^2 = 0.005$, and ρ_L and ρ_U are set to be 0.025 and 100, respectively. The number of Monte Carlo runs is 300. It is seen from Fig. 1 that for a given average false alarm period, the average detection delay of the proposed RGCUSUM is shorter than that of the adaptive CUSUM algorithm, which implies that the proposed RGCUSUM can detect the FDIA more efficiently than the adaptive CUSUM test. This is expected since the adaptive CUSUM test builds on some assumptions on the model as mentioned in Section 1. However, these assumptions do not hold for the simulation setup. The efficiency loss of the adaptive CUSUM test may be brought about by the model mismatch. Moreover, the adaptive CUSUM test requires that the FDIA is positive and small, and hence is prone to efficiency loss for large and negative FDIA, which is the case in the simulation.

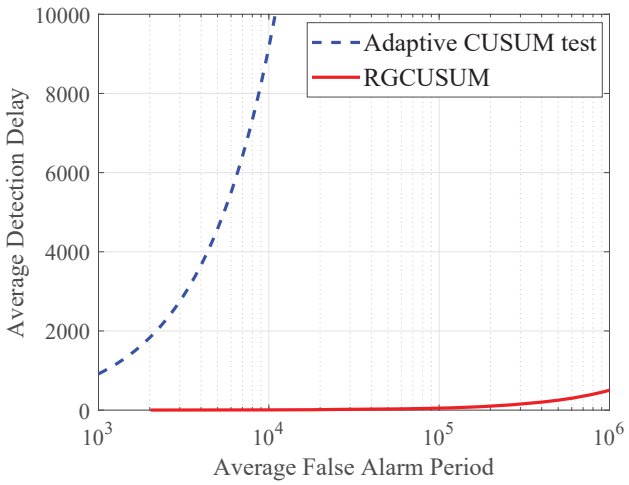

Fig. 1: Performance Comparison between the RGCUSUM and the Adaptive CUSUM Test in [2].

6. CONCLUSIONS

In this paper, we have considered the problem of sequentially detecting time-varying FDIAs in dynamic smart grids. New CUSUM-type algorithms have been proposed to address this problem, and we have shown that the computational complexity of the proposed algorithm scales linearly with the number of meters in the smart grid. Furthermore, we also have provided performance analysis for the proposed algorithm. To be specific, considering Lordon's setup, for any given constraint on the expected false alarm period, a lower bound on the threshold employed in the proposed algorithm has been derived. Furthermore, for any given threshold employed in the proposed algorithm, we have provided an upper bound on the worst-case expected detection delay. In the end, the performance of the proposed algorithm has been numerically studied based on an IEEE standard power system under FDIAs.

7. REFERENCES

[1] S. Cui, Z. Han, S. Kar, T. T. Kim, H. V. Poor, and A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 106–115, 2012.

[2] Y. Huang, H. Li, K. A. Campbell, and Z. Han, "Defending false data injection attack on smart grid network using adaptive CUSUM test," in *Proc. 45th Annu. Conf. Inf. Sci. Syst.* Baltimore, MD, USA, March 2011, pp. 1–6.

[3] Y. Huang, J. Tang, Y. Cheng, H. Li, K. A. Campbell, and Z. Han, "Real-time detection of false data injection in smart grid networks: an adaptive cusum method and analysis," *IEEE Syst. J.*, vol. 10, no. 2, pp. 532–543, 2016.

[4] S. Li, Y. Yılmaz, and X. Wang, "Quickest detection of false data injection attack in wide-area smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2725–2735, Nov 2015.

[5] M. N. Kurt, Y. Yılmaz, and X. Wang, "Distributed quickest detection of cyber-attacks in smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 8, pp. 2015–2030, 2018.

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

[7] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, pp. 1897–1908, 1971.

[8] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *The Annals of Statistics*, pp. 1379–1387, 1986.

[9] M. Basseville, I. V. Nikiforov *et al.*, *Detection of abrupt changes: theory and application*. Prentice Hall Englewood Cliffs, 1993, vol. 104.

[10] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. CRC Press, 2014.

[11] R. D. Zimmerman and C. E. Murillo-Sánchez, *Matpower 6.0 User's Manual*. [Online]. Available: http://www.pserc.cornell.edu/matpower/MATPOWER-manual.pdf, Dec. 2016.