This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

Secure Distributed Dynamic State Estimation in Wide-Area Smart Grids

Mehmet Necip Kurt, *Student Member, IEEE*, Yasin Yılmaz, *Member, IEEE*, and Xiaodong Wang, *Fellow, IEEE*

***Abstract*— Smart grid is a large complex network with a myriad of vulnerabilities, usually operated in adversarial settings and regulated based on estimated system states. In this study, we propose a novel highly secure distributed dynamic state estimation mechanism for wide-area (multi-area) smart grids, composed of geographically separated subregions, each supervised by a local control center. We firstly propose a distributed state estimator assuming regular system operation, that achieves near-optimal performance based on the local Kalman filters and with the exchange of necessary information between local centers. To enhance the security, we further propose to (i) protect the network database and the network communication channels against attacks and data manipulations via a blockchain (BC)-based system design, where the BC operates on the peer-to-peer network of local centers, (ii) locally detect the measurement anomalies in real-time to eliminate their effects on the state estimation process, and (iii) detect misbehaving (hacked/faulty) local centers in real-time via a distributed trust management scheme over the network. We provide theoretical guarantees regarding the false alarm rates of the proposed detection schemes, where the false alarms can be easily controlled. Numerical studies illustrate that the proposed mechanism offers reliable state estimation under regular system operation, timely and accurate detection of anomalies, and good state recovery performance in case of anomalies.**

***Index Terms*— Smart grid, distributed, secure, dynamic state estimation, Kalman filter, blockchain, real-time anomaly detection, trust management.**

I. Introduction

The next-generation electrical power grid, i.e., the smart grid, is vulnerable to a variety of cyber/physical system faults and hostile cyber threats [1]–[4]. In particular, random anomalies such as node and topology failures might occur all over the network. Moreover, malicious attackers can deliberately manipulate the network operation and tamper with the network data, from sources such as sensors, control centers, network database, and network communication channels (see Fig. 1). False data injection (FDI), jamming, and denial of service (DoS) attacks are well-known attack types [2], [5]–[8]. Further, Internet of Things (IoT) botnets can be used to target against critical infrastructures such as the smart grid [9]–[11].

In the smart grid, online state estimates are utilized to make timely decisions in critical tasks such as load-frequency control and economic dispatch [12]. Hence, a fundamental task in the grid is reliable state estimation based on online measurements. On the other hand, the main objective of the adversaries is to damage/mislead the state estimation mechanism in order to cause wrong/manipulated decisions, resulting in power blackouts or manipulated electricity prices [13]. Additionally, random system faults may degrade the state estimation performance. Our objective in this study is to design a highly secure and resilient state estimation mechanism for wide-area smart grids, that provides reliable state estimates in a fully-distributed manner, even in the case of cyber-attacks and other network anomalies.

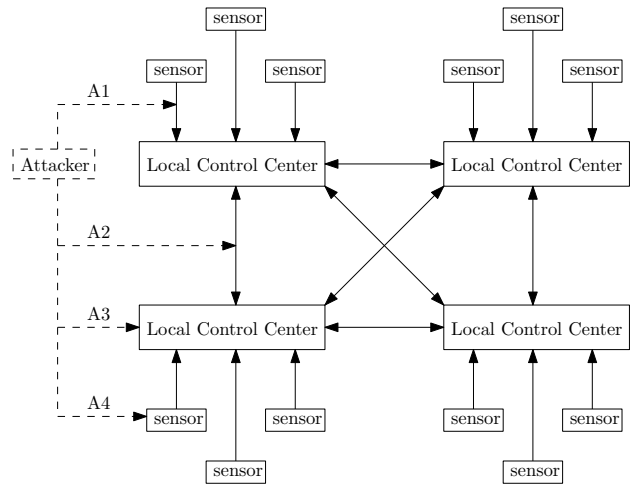

Fig. 1: An illustration of the vulnerabilities of wide-area smart grids against malicious adversaries. The solid lines illustrate the network communication channels and the dashed lines illustrate possible attacks. The attacker can (i) manipulate, jam or block the communication channels (A1, A2), (ii) hack a local control center to manipulate its operation or to tamper with the network database (A3), and (iii) hack a sensor to corrupt its online measurements (A4).

A. Background and Related Work

1) Secure Dynamic State Estimation: Feasibility of dynamic modeling and efficiency of dynamic state estimation have been widely discussed and various dynamic models have been proposed for the power grids [14]–[18]. The general consensus is that the dynamic modeling better captures the time-varying characteristics of the power grid and dynamic state estimators are more effective to track the system state compared to the conventional static least squares (LS) estimators. Moreover, state forecasting capability achieved with dynamic estimators is quite useful in real-time operation and security of the grid [5], [14], [15].

In the literature, various techniques have been proposed to make the dynamic state estimation mechanism secure/robust

This work was supported in part by the U.S. National Science Foundation (NSF) under Grant ECCS-1405327, and in part by the U.S. Office of Naval Research under Grant N000141410667. The work of Y. Yılmaz was supported in part by the NSF under Grant CNS-1737598.

M. N. Kurt and X. Wang are with the Department of Electrical Engineering, Columbia University, New York, NY 10027, USA (e-mail: m.n.kurt@columbia.edu; wangx@ee.columbia.edu).

Y. Yılmaz is with the Department of Electrical Engineering, University of South Florida, Tampa, FL 33620, USA (e-mail: yasiny@usf.edu).

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

against outliers, heavy-tailed noise processes, model uncertainties including unknown noise statistics, rank-deficient observation models, and cyber-attacks, etc., [19]–[34]. For instance, robust statistics-based approaches [19]–[22] aim to suppress the effects of outliers by assigning less weights to more significant outliers. As the outliers are still incorporated into the state estimation process, their effects are not completely eliminated, and due to the recursive nature of dynamic state estimators, errors are accumulated over time and the corresponding state estimator breaks down at some point, i.e., fails to keep track of the system operating point. The estimator also breaks down in case of gross outliers. Furthermore, this approach is based on the solution of an iterative weighted LS problem, repeated at each time, that might be prohibitive for real-time processing.

To deal with outliers, another method is modeling the system noise as a heavy-tailed, e.g., Student's t or Laplace, distribution [23]–[25]. This method can handle (nominal) outliers observed during the regular system operation, however, it is expected to be ineffective against attacks and faults that behave significantly different from nominal outliers. Further, some studies are based on the assumption that attacks are sparse over the network or with bounded magnitudes [26]–[30], which significantly limits the robustness of the corresponding state estimators. This is because imposing restrictions on the capabilities or strategies of attackers make the corresponding state estimation mechanism robust against only a certain subset of attacks. Yet another approach is to completely ignore the outliers in the state estimation process [31]–[33], [35]. In the literature, this approach is usually based on the sample-by-sample classification of observations as either outlier or non-outlier. Although it might be useful to detect and eliminate the effects of gross outliers, the corresponding state estimator is expected to fail if small/moderate-level (difficult-to-detect) outliers are observed in a persistent manner. Moreover, it might fail even in the nominal case due to the rejection of nominal outliers. Finally, several dynamic state estimation techniques are evaluated in [34] for the nonlinear power system model and it is shown that the cubature Kalman filter is more robust to model uncertainties and attacks/anomalies compared to the extended Kalman filter and the unscented Kalman filter.

Distributed dynamic state estimation has been also studied extensively, see e.g., [36] for a review of the distributed Kalman filtering techniques. Particularly, two main architectures are considered: hierarchical and fully-distributed. In the former, there exists a central controller coordinating multiple local controllers, while in the latter, no central controller exists. In the hierarchical schemes, the information filter, an algebraic equivalence to the centralized Kalman filter, can be useful to fuse the information processed across the local controllers in a simple manner [37], [38]. The fully-distributed schemes usually require an iterative consensus mechanism to reduce the disagreement of the local controllers on the estimates of common state variables [36], [39]. More recently, a distributed extended Kalman filter is proposed in [40] based on the internodal transformation theory, not requiring iterative data exchanges. A recent survey of model-based distributed filtering and control including the distributed Kalman filtering and secure control algorithms for large-scale cyber-physical systems is presented in [41]. For instance, [42] studies the secure control of stochastic linear systems against deception attacks, that manipulate sensor measurements or control inputs.

2) Secure Distributed System Design: Traditionally, through the supervisory control and data acquisition (SCADA) system, the power grid is controlled in a centralized manner. In particular, the system-wide data are collected, stored, and processed at a single node. Considering the increasing speed and size of the measurement data, collecting and processing such large volume of data in real-time at a single center seem practically infeasible in the modern power grids [40]. Moreover, the traditional implementation is based on the assumption that the centralized node is completely trustable. In practice, however, the centralized node can be the weakest point of the network in terms of security. This is because by hacking only the centralized node, adversaries can arbitrarily modify the control decisions and the network database. On the contrary, hacking a distributed system is usually more difficult for attackers, especially for the smart grid that is distributed over a geographically wide region. Therefore, distributing the computation and the trust over the network can be useful to achieve a more feasible and a more secure grid.

Blockchain (BC) is an emerging secure distributed database technology, operating on a peer-to-peer (P2P) network with the following key components [43]–[45]: (i) a distributed ledger (chronologically ordered sequence of blocks) that is shared and synchronized over the network where each block is cryptographically linked to the previous blocks and the ledger is resistant/immune to modifications, (ii) advanced cryptography enabling secure data exchanges and secure data storage, and (iii) a mutual consensus mechanism, that enables collective verification/validation on the integrity of exchanged and stored data, and thereby distributes trust over the network instead of relying on a single node/entity. BC technology was firstly used in financial applications [46], [47] but due to its security-by-design and the distributed nature without needing any trusted third party, it has been applied to diverse fields such as vehicular networks, supply chain management, cognitive radio, and insurance [48]–[50].

The smart grid critically relies on a database and network communication channels, both are quite vulnerable to attacks and manipulations. As a countermeasure, an effective approach is the detection and then mitigation of such threats. On the other hand, a significantly better approach is the prevention of the threats as much as possible. In this direction, the BC technology has a great potential due to its advanced data protection/attack prevention capabilities. Hence, we aim to integrate some salient features of the BC technology to the smart grid in order to improve the resilience of the system, particularly to protect the system database and the communication channels.

Research on the integration of BC to smart grids has mainly focused on secure energy transactions/trade so far [4], [51]–[53], and only a few studies have examined the integration of BC to make the power supply system more secure, see [54] and [55], where in both studies, the grid is protected by securely storing all the system-wide measurements at every

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

meter. However, this seems infeasible in many aspects, e.g., meters are small-size devices with limited memory, power, and processing capabilities [52] and hence not suitable to perform advanced computations such as encryption/decryption, to store the distributed ledger, and to constantly communicate with all other meters in the network. Moreover, since the measurements are collected mainly to estimate the system state, the BC-based system can be designed to protect the state estimation mechanism in a more direct way, rather than to protect the entire history of raw measurement data, that then enables secure state estimation.

Although BC can be useful to secure the network database and the communication channels, the online sensor measurements are still vulnerable to attacks and faults. We would like to design a state estimation mechanism that is secure against all types of anomalies. Towards this goal, as a complement to the BC-based data protection, robust bad data detection and mitigation, i.e., state recovery, schemes need to be integrated into the state estimation mechanism. We have recently proposed in [7] a robust dynamic state estimation scheme for the smart grid, in case the attack models are known (with some unknown parameters). In practice, unknown attacks/anomalies may occur in the smart grid as it has many vulnerabilities and attackers might have arbitrary strategies. Hence, in general, anomaly/attack models need to be assumed unknown and the state estimation mechanism should be designed accordingly. Furthermore, in a BC-based distributed system, there is no centralized trusted node to check and recover a node that is faulty or hacked by a malicious entity. Hence, a distributed trust management mechanism needs to be employed over the network to evaluate the trustability of each node against the possibility of misbehaving nodes.

B. Contributions

In this study, we propose a novel BC-based resilient system design to achieve secure distributed dynamic state estimation in wide-area smart grids. Our aim is to reduce the risks at each part of this highly complex network, specifically, the network database, sensors, local control centers, and network communication channels (see Fig. 1).

Firstly, assuming regular system operation (no anomaly), we propose a novel fully-distributed dynamic state estimation scheme that achieves near-optimal performance thanks to the local Kalman filters and with the exchange of necessary information between local centers. Then, to improve the resilience of the proposed mechanism, we further propose to

- use salient features of the emerging BC technology to secure both the network database and the network communication channels against attacks and manipulations,
- embed novel online anomaly detection schemes into the state estimation mechanism to make it secure against measurement anomalies, and
- detect and eliminate the effects of misbehaving nodes in real-time via a novel distributed trust management mechanism over the network.

We provide theoretical guarantees regarding the false alarm rates of the proposed online detection schemes, where the false alarms can be easily controlled by the system designer. Finally, since the proposed anomaly detection, trust management, and state recovery schemes do not assume any knowledge about the type/cause of the anomaly/misbehavior, they are all robust to unknown attacks/anomalies.

C. Organization and Notations

The remainder of the paper is organized as follows. Sec. II presents the system model. Sec. III describes the proposed BC-based secure system design. Sec. IV discusses the proposed distributed state estimation mechanism under regular (non-anomalous) network operation. Sec. V explains the proposed online anomaly detection scheme against measurement anomalies and the corresponding state recovery scheme. Sec. VI discusses the proposed distributed trust management scheme against misbehaving nodes. Sec. VII then summarizes the proposed mechanism. Sec. VIII illustrates the advantages of the proposed mechanism over a simulation setup. Finally, Sec. IX concludes the paper.

Notations: Boldface letters denote vectors and matrices. $\mathbb{R}$ denotes the set of real numbers. $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the Gaussian probability density function (pdf) with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. $\mathbf{I}_m$ denotes an $m \times m$ identity matrix. $\mathrm{P}(\cdot)$ and $\mathrm{E}(\cdot)$ denote the probability and the expectation operators, respectively. $\mathbb{I}(\cdot)$ denotes the indicator function. $\log(\cdot)$ denotes the natural logarithm and e denotes the Euler's number. $|\cdot|$ denotes the cardinality of a set. $\emptyset$ denotes an empty set. $\mathcal{S}_1 \backslash \mathcal{S}_2$ denotes the set of elements belonging to $\mathcal{S}_1$ but not belonging to $\mathcal{S}_2$. $\inf$, $\sup$, and $\max$ denote the infimum, supremum, and maximum operators, respectively, and $(\cdot)^+ \triangleq \max\{\cdot, 0\}$. Finally, $\cdot^{\mathrm{T}}$ denotes the transpose operator.

II. System Model

We consider a smart power grid with N buses and K sensors. System state at time t, $\mathbf{x}_t = [x_{1,t}, \ldots, x_{2N-1,t}]^{\mathrm{T}}$, represents voltage magnitudes and phase angles of the buses, where a bus is chosen as the reference for the phase angles. The measurements are real and reactive power flow, power injection, and voltage magnitude measurements. The measurement taken at sensor $k \in \{1, \ldots, K\}$ at time t is denoted with $y_{k,t}$ and the measurement vector is denoted with $\mathbf{y}_t = [y_{1,t}, \ldots, y_{K,t}]^{\mathrm{T}}$. We model the grid as a discrete-time linear dynamic system as follows [14], [15], [32], [56], [57]:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{v}_t, \tag{1}$$

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \mathbf{w}_t, \tag{2}$$

where $\mathbf{A} \in \mathbb{R}^{2N-1 \times 2N-1}$ is the state transition matrix, $\mathbf{H} \in \mathbb{R}^{K \times 2N-1}$ is the measurement matrix, $\mathbf{v}_t = [v_{1,t}, \ldots, v_{2N-1,t}]^{\mathrm{T}}$ is the process noise vector, and $\mathbf{w}_t = [w_{1,t}, \ldots, w_{K,t}]^{\mathrm{T}}$ is the measurement noise vector. We assume that $\mathbf{v}_t$ and $\mathbf{w}_t$ are independent AWGN processes where $\mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_{2N-1})$ and $\mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \sigma_w^2 \mathbf{I}_K)$.

A. State-Space Model

To model the time behavior of the power system state, several attempts have been made. Firstly, considering small

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

time intervals between consecutive measurements, a linear state transition model is assumed [15], [32]. Further, a quasi-static state model is commonly employed where the system state is periodic over a day since the system loads typically have a daily cycle [14], [32]. Moreover, a linear exponential smoothing model is proposed in [17], where the effects of past measurements on the state estimates are reduced over time. In the simplest case, the system state is assumed to be the same with the previous time step, up to a random uncertainty [14], i.e.,

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{v}_t,$$

where the process noise $\mathbf{v}_t$ corresponds to the random uncertainty. In a more general form, we can write

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{v}_t.$$

In the actual power system operation, the relationship between the measurements and the system state is nonlinear. Hence, we have

$$\mathbf{y}_t = h(\mathbf{x}_t) + \mathbf{w}_t,$$

where $h(\cdot)$ is a nonlinear function. Linearizing the system around an operating point $\mathbf{x}_0$, we obtain

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \mathbf{w}_t,$$

where

$$\mathbf{H} = \left.\frac{\partial h(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}$$

is the Jacobian matrix.

The parameters of the state-space model, particularly the state transition matrix and the process and measurement noise variances can be determined from power system data obtained under regular operating conditions [14], [15].

B. Wide-Area Monitoring Model

The wide-area smart grid is composed of geographically separated subregions (see Fig. 5). Each subregion contains a set of sensors, supervised by a local (control) center. Since the sensors are distributed over the network, each local center partially observes the measurement vector $\mathbf{y}_t$. Assuming that the grid is composed of L subregions and the subset of sensors in the ℓth subregion is denoted by $\mathcal{R}^\ell$, the measurement vector $\mathbf{y}_t$ is decoupled into L sub-vectors $\mathbf{y}_t^\ell \in \mathbb{R}^{K^\ell}$, $\ell = 1, \ldots, L$, where $\mathbf{y}_t^\ell \triangleq \{y_{k,t} \,|\, k \in \mathcal{R}^\ell\}$ denotes (with an abuse of notation) the measurement vector of the ℓth local center at time t and $K^\ell \triangleq |\mathcal{R}^\ell|$ is the number of sensors in the ℓth subregion. Since each sensor belongs to only one subregion, for any two subregions ℓ and j, $\mathbf{y}_t^\ell$ and $\mathbf{y}_t^j$ do not overlap and we have $\sum_{\ell=1}^{L} K^\ell = K$.

The smart grid is an interconnected system, where there exist tie-lines between neighboring subregions (see Fig. 5) that leads to some common (shared) state variables between neighboring local centers. Hence, denoting the state vector of the ℓth local center at time t by $\mathbf{x}_t^\ell \in \mathbb{R}^{N^\ell}$, for any two neighboring local centers ℓ and j, $\mathbf{x}_t^\ell$ and $\mathbf{x}_t^j$ might overlap. This implies $\sum_{\ell=1}^{L} N^\ell \geq 2N-1$. In general, if the state transition matrix $\mathbf{A}$ is non-diagonal, additional state variables might be shared between neighboring or non-neighboring local centers due to dependencies between state variables over time through the state transition matrix. In this study, for the simplicity of the presentation, we assume $\mathbf{A}$ is diagonal. Under this assumption, we next determine the state vector of a local center, say the ℓth one. For the derivation of local state vectors in the case of non-diagonal $\mathbf{A}$, please see [5, Sec. IV-A]. Note that in the non-diagonal case, the proposed system design directly extends, where the only difference is that the size of the local state vectors might be larger.

Let $\mathbf{h}_k^{\mathrm{T}} = [h_{k,1}, \ldots, h_{k,2N-1}]$ be the kth row of the measurement matrix $\mathbf{H}$, i.e., $\mathbf{H}^{\mathrm{T}} = [\mathbf{h}_1, \ldots, \mathbf{h}_K]$. Then, using (2), each measurement $y_{k,t}, k \in \{1, \ldots, K\}$ can be written as follows:

$$y_{k,t} = \mathbf{h}_k^{\mathrm{T}}\mathbf{x}_t + w_{k,t}. \tag{3}$$

Based on (3), we can argue that $y_{k,t}$ depends on, equivalently bears information about, the following state variables:

$$\mathcal{X}_{y_k} \triangleq \{x_{n,t} \,|\, h_{k,n} \neq 0, n = 1, \ldots, 2N-1\}.$$

Then, the local state vector $\mathbf{x}_t^\ell$ consists of the union of all such state variables for all the sensors k in the subregion ℓ:

$$\mathbf{x}_t^\ell = \bigcup_{k \in \mathcal{R}^\ell} \mathcal{X}_{y_k}.$$

For each local center $\ell \in \{1, \ldots, L\}$, we then have the following local state transition model:

$$\mathbf{x}_t^\ell = \mathbf{A}^\ell \mathbf{x}_{t-1}^\ell + \mathbf{v}_t^\ell, \tag{4}$$

and the following local measurement model:

$$\mathbf{y}_t^\ell = \mathbf{H}^\ell \mathbf{x}_t^\ell + \mathbf{w}_t^\ell, \tag{5}$$

where the local state transition matrix, $\mathbf{A}^\ell \in \mathbb{R}^{N^\ell \times N^\ell}$, and the local measurement matrix, $\mathbf{H}^\ell \in \mathbb{R}^{K^\ell \times N^\ell}$, can be easily obtained from $\mathbf{A}$ and $\mathbf{H}$, respectively. Moreover, the local process noise vector, $\mathbf{v}_t^\ell \in \mathbb{R}^{N^\ell}$, is a sub-vector of $\mathbf{v}_t$ corresponding to $\mathbf{x}_t^\ell$. Similarly, the local measurement noise vector, $\mathbf{w}_t^\ell \in \mathbb{R}^{K^\ell}$, is a sub-vector of $\mathbf{w}_t$ corresponding to $\mathbf{y}_t^\ell$.

III. Blockchain-Based Secure System Design

A. Overview of the Proposed System

We consider a distributed P2P network of local centers where each node (local center) can communicate with all other nodes (see Fig. 1). We aim to design a system in which the nodes collaborate with each other to perform the state estimation task in a safe and reliable manner. For a reliable distributed dynamic state estimation, particularly the Kalman filter, we need safe updates and hence the following three items must be secure/reliable at each time t:

- state estimates of the previous time $t-1$,
- sensor measurements acquired at the current time t, and
- the nodes functioning in the state estimation process, i.e., the local centers.

In other words, at each time, we need to make sure that the previous state estimates are not modified, the online sensor measurements are not anomalous, and the nodes are working according to predesigned network rules. Furthermore, in case

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

of an anomaly over the network, the state estimates can be recovered using the previous reliable state estimates, that requires us to also protect the previous state estimates against tampering. Considering these requirements, our proposed system is composed of the following three main components:

- *BC-based data protection/attack prevention:* BC enhances the security of the grid by reducing the risk of manipulations at the network database and the network communication channels. In particular, to protect the previous state estimates against tampering and make them widely available and accessible over the network against the possibility of node failures and hacking, we record them in a shared distributed ledger that is resistant to alterations. Moreover, we secure the inter-node data exchanges via cryptography against attacks and manipulations.
- *Secure state estimation against measurement anomalies:* Each local center quickly and reliably detects local measurement anomalies and then employs an online state recovery mechanism.
- *Distributed trust management:* All nodes collectively (via voting-consensus) evaluate the trustability of each node, specifically whether the local state estimates provided by a node exhibit an anomalous pattern over time.

Firstly, the following subsection explains how we use the BC technology to enhance the security of the state estimation mechanism.

B. Blockchain Mechanism

The BC operates on the P2P network of local centers. Since each node is pre-specified and pre-authenticated, we have a permissioned (private) BC mechanism [52], [54]. In BC-based systems, duties of each node and interactions between nodes are determined via a smart contract, which is a software code that specifies the predefined rules of network operation. In our proposed mechanism, each node collects and analyzes the sensor measurements in its subregion, estimates its local state vector, exchanges information with other nodes, performs encryption/decyrption, participates in voting-consensus procedures, and stores the distributed ledger in its memory. The details regarding the duties of the nodes will be more clear in the subsequent sections. Next, we explain the proposed BC mechanism in more detail.

1) Data Exchanges: In all inter-node data exchanges, we use asymmetric encryption based on the public key infrastructure (see Fig. 2). In this mechanism, each node owns a public-private key pair that forms the digital identity of the node. The public key is available at every node. On the other hand, the private key is available only at its owner. Moreover, a secure hash algorithm (SHA), e.g., SHA-256, SHA-512, etc. [58], [59], is used in the data encryption process. Particularly, in every data exchange, the sender node firstly processes its message via the SHA and obtains the message digest. It then encrypts the message digest via its private key using a signature algorithm, e.g., Elliptic Curve Digital Signature Algorithm [44], [51], and obtains the digital signature. Finally, it transmits the data package consisting of the message and

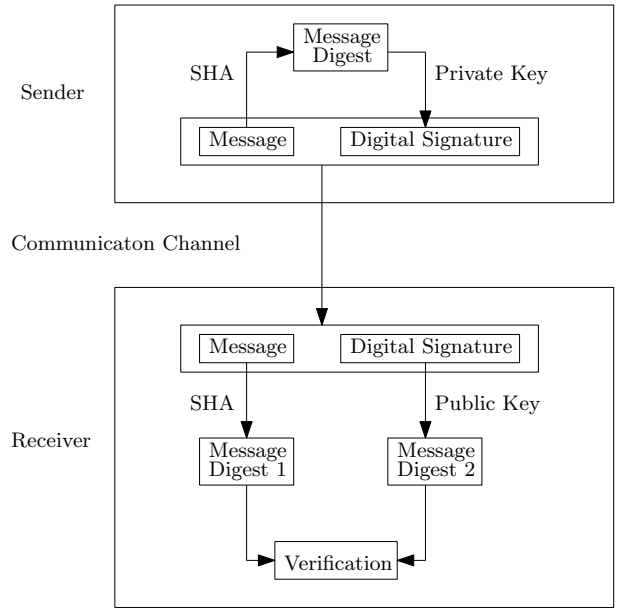

Fig. 2: Inter-node data exchanges based on the asymmetric encryption mechanism.

the corresponding digital signature. The receiver node then decrypts the received digital signature via the public key of the sender node and obtains a message digest. Moreover, it processes the received message via the SHA and obtains another message digest. Only if these two message digests exactly match, the integrity of the received message is verified.

In this procedure, the SHA provides security since it is computationally intractable to obtain the same message digest from two different messages [58], [59]. The SHA is a one-way function that outputs a fixed-length message digest for an arbitrary-size input message. Let $H(\cdot)$ denote the SHA and the output of the $H(\cdot)$ be n bits. Then, given a message m, the time complexity of finding $m' \neq m$ such that $H(m') = H(m)$ is $O(2^n)$ via brute-force search. This property implies that over a data exchange, if a malicious adversary aims to replace the actual message m with a fake message m' without being noticed ($H(m') = H(m)$ so that the receiver verifies the integrity of the message), and moreover if the adversary has a computational power of querying 2^θ possible fake messages, then the probability of a successful fake message is $2^{\theta-n}$. Here we assume that the adversary knows the SHA so that it can check whether $H(m') = H(m)$ while trying different fake messages m'. The probability of success is negligible in practical settings where $n = 256$, $n = 512$, etc. For example, if an adversary can query 2^{80} fake messages and $n = 256$, the probability of success is 2^{-176} for a single data package.

Furthermore, since the received digital signature can only be decrypted via the public key of the sender node, the receiver can verify the identity of the sender. Assuming (reasonably) that the private keys are kept secret and the digital signatures are k bits, the time complexity of generating a successful fake signature is $O(2^k)$ via brute-force search [44]. Then, if an adversary does not know the public key of the sender (so that it cannot check whether a fake signature is decrypted via the public key), the probability of generating a successful fake

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

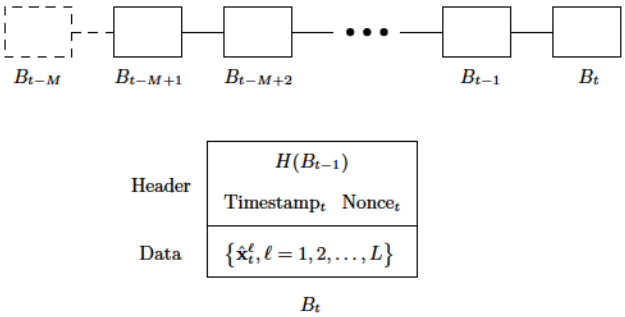

Fig. 3: The block structure and the distributed ledger at time t. B_t denotes the block generated at time t and $H(\cdot)$ denotes the cryptographic SHA. Each block contains a header and the estimates of the all local state vectors of the corresponding time. The ledger is composed of the M recent blocks where the oldest block, B_{t-M}, is pruned at each time t.

signature for a chosen fake message is 2^{-k}. On the other hand, obtaining public keys might be easier than private keys because the public keys are distributed over the network and all public keys can be accessed by hacking only one node. Then, if an adversary knows the public key of the sender, it can try different fake signatures for a chosen fake message and check whether the fake message is verified. In this case, if the adversary has the computational power of querying 2^θ possible fake signatures, then the probability of success is $2^{\theta-k}$. Again choosing k sufficiently high, such as $k = 256$, makes the success probability practically negligible. Notice that, however, if the private key of the sender is stolen, then the fake messages cannot be noticed at the receiver.

Over a data exchange, in case either the integrity of the received data package or the identity of the sender cannot be validated, then the received message is ignored and a retransmission can take place. Thereby, thanks to the asymmetric encryption procedure, the inter-node data exchanges are secured against attacks that manipulate either the message or the identity of the sender, such as man-in-the-middle and IP spoofing attacks.

2) Distributed Ledger and Consensus Mechanism: The ledger is a chronologically ordered sequence of blocks, stored at every node and synchronized over the network. In BC-based systems, the block content is application-specific. In our case (see Fig. 3), at each measurement sampling interval, a new block is produced, that includes (i) the state estimates of the current time and (ii) a header consisting of the discrete timestamp, hash value of the previous block that cryptographically links the current block to the previous block, and a random number called the nonce, which is the solution to a puzzle problem. Particularly, the nonce is determined, via brute-force search, such that the hash value of the current block satisfies a certain condition [44], [54]. We next explain the process of producing a new block and the corresponding update in the distributed ledger.

At each time t, after computing the local state estimates, each node ℓ broadcasts a data package, containing its local state estimates, $\hat{\mathbf{x}}_t^\ell$, as the message, to the entire network. Then, as explained above, every node checks the validity of each received data package. A broadcasted data package is verified over the network only if the majority of the nodes validate it. After all the broadcasted data packages are verified, all the local state estimates of the current time are recorded into a new block at each node.

In this process, to alleviate inconsistencies in the distributed ledger, we need to make sure that the ledger is synchronously updated over the network. Moreover, we seek a mutual consensus over the network for the update of the ledger. Using the common Proof-of-Work consensus mechanism [44], some nodes act as "miners", where the miners compete with each other to solve the puzzle and obtain the nonce value. Public BC-based systems such as Bitcoin [46] provides an incentive to miners, where the miner solving the puzzle first, gets a financial reward. In our case, however, this procedure is completely autonomous in that at each time, among all the local centers, a few nodes are randomly assigned as the miners. Then, the miner solving the puzzle first broadcasts the nonce value to the entire network. Each node then checks the puzzle solution. If the majority of nodes verify the solution, then the new block is produced and simultaneously connected to the existing ledger at every node. Here, random assignment of the miners enables a higher level of security to the BC mechanism compared to permanently assigning the miner nodes, that would then be the open targets for adversaries.

As explained before, for a reliable dynamic state estimation, we require that the most recent state estimates are secure. Moreover, in case of an anomaly over the network, e.g., a failure or an attack, we can recover the system state from the recent reliable state estimates (the details are presented in Sec. V and Sec. VI). For these reasons, in the distributed ledger, we propose to store a finite number of blocks that contain the recent state estimates in order to protect them against all kinds of manipulations. Let the number of blocks in the ledger be $M \geq 1$. Then, at each time, while a new block is connected to the existing ledger, the oldest block is pruned (see Fig. 3). This, in fact, solves the problem of monotonically increasing storage costs in the conventional BCs as well. We will explain how to choose M in the subsequent sections. Further, in our case, at each time interval, only one new block is generated and we have a main chain of blocks without any forks, unlike the well-known public BCs such as Bitcoin [46] and Ethereum [47], in which the ledger keeps record of the transactions between nodes and it is possible to observe multiple transactions at the same time. Finally, we assume that the update of the ledger is completed within one measurement sampling interval.

The distributed ledger is resistant to modifications due to the following reasons: (i) since each block is cryptographically linked to the previous block, to modify a single block without being noticed, all the subsequent blocks must be modified accordingly, and (ii) since updating the ledger requires mutual consensus over the network, in order to modify the ledger, malicious entities need to control the majority of nodes in the network. This also implies that the security improvements introduced with the BC-based system design are valid only if the majority of the nodes in the network are reliable. We expect that this condition is easily satisfied in large-scale smart grids with many nodes distributed over a geographically wide region, for which hacking the majority of the nodes

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

is practically quite difficult. From an attacker's perspective, letting $S \triangleq \lceil L/2 \rceil$ be the minimum number of nodes to be hacked in order to control the majority and hence to be able to arbitrarily modify the ledger, the best strategy would be attacking on the least costly set (in terms of the required efforts and resources for hacking) among the $\binom{L}{S}$ possible sets of nodes. Furthermore, an attacker may replace a block, say B_t, with a fake block B_t' such that $H(B_t') = H(B_t)$ with a practically negligible probability (similar to the analysis in Sec. III-B.1), however, the fake block can still be noticed by comparing the modified ledger with the other copies of the ledger in the network. Hence, for a successful fake block, the attacker still needs to control the majority of the nodes.

Remark 1: Private BC mechanisms can be much more efficient compared to public BCs [52], [54]. Moreover, recent research shows promising improvements in the speed of BC system operation, see e.g., [60], indicating that the BC efficiency can be safely improved in the near future, that enables using BC in time-sensitive real-time applications such as dynamic state estimation. Hence, in our mechanism, the update of the ledger can be completed within one measurement sampling interval using a fast private BC mechanism.

IV. Distributed State Estimation Assuming Regular System Operation

In this section, assuming that the grid operation always fits to the nominal system model (see (1) and (2)), we aim to perform optimal state estimation in a fully-distributed manner, where each node ℓ only estimates its local state vector $\mathbf{x}_t^\ell$.

For an optimal state estimation, the node ℓ needs to use all the measurements that bear information about $\mathbf{x}_t^\ell$, either fully or partially. The node ℓ has access to only the local measurements, $\mathbf{y}_t^\ell$, that are clearly informative about $\mathbf{x}_t^\ell$ (see (5)). On the other hand, due to the shared state variables, some measurements collected at the other nodes may also be informative about $\mathbf{x}_t^\ell$. Let

$$\mathcal{C}^\ell \triangleq \{j \,|\, \mathbf{x}_t^j \cap \mathbf{x}_t^\ell \neq \emptyset, j \in \{1,\dots,L\} \backslash \{\ell\}\}$$

be the set of nodes that share at least one state variable with the ℓth node. Let $j \in \mathcal{C}^\ell$ and $\mathbf{y}_t^{\ell,j}$ be the sub-vector of $\mathbf{y}_t^j$ that is informative about $\mathbf{x}_t^\ell$, i.e.,

$$\mathbf{y}_t^{\ell,j} \triangleq \{y_{k,t} \,|\, k \in \mathcal{R}^j, \mathcal{X}_{y_k} \cap \mathbf{x}_t^\ell \neq \emptyset\}.$$

Moreover, let $\mathbf{x}_t^{\bar{\ell},j} \triangleq \mathbf{x}_t^j \backslash \mathbf{x}_t^\ell$. Then, we can decompose $\mathbf{y}_t^{\ell,j}$ as

$$\mathbf{y}_t^{\ell,j} = \mathbf{H}^{\ell,j}\mathbf{x}_t^\ell + \mathbf{H}^{\bar{\ell},j}\mathbf{x}_t^{\bar{\ell},j} + \mathbf{w}_t^{\ell,j}, \tag{6}$$

where the matrices $\mathbf{H}^{\ell,j}$ and $\mathbf{H}^{\bar{\ell},j}$ are easily determined such that the equality in (6) is satisfied for all t. Moreover, $\mathbf{w}_t^{\ell,j}$ is the sub-vector of $\mathbf{w}_t$ corresponding to $\mathbf{y}_t^{\ell,j}$.

In (6), the term $\mathbf{H}^{\bar{\ell},j}\mathbf{x}_t^{\bar{\ell},j}$ is clearly non-informative and irrelevant to the state estimator of the ℓth node. On the other hand, the jth node estimates $\mathbf{x}_t^j$ (and hence $\mathbf{x}_t^{\bar{\ell},j}$) at each time t. Then, denoting the estimate of $\mathbf{x}_t^{\bar{\ell},j}$ by $\hat{\mathbf{x}}_t^{\bar{\ell},j}$, based on (6), we can write

$$\begin{aligned}\tilde{\mathbf{y}}_t^{\ell,j} &\triangleq \mathbf{y}_t^{\ell,j} - \mathbf{H}^{\bar{\ell},j}\hat{\mathbf{x}}_t^{\bar{\ell},j} \\ &= \mathbf{H}^{\ell,j}\mathbf{x}_t^\ell + \mathbf{H}^{\bar{\ell},j}(\mathbf{x}_t^{\bar{\ell},j} - \hat{\mathbf{x}}_t^{\bar{\ell},j}) + \mathbf{w}_t^{\ell,j} \\ &= \mathbf{H}^{\ell,j}\mathbf{x}_t^\ell + \tilde{\mathbf{w}}_t^{\ell,j},\end{aligned} \tag{7}$$

where

$$\tilde{\mathbf{w}}_t^{\ell,j} \triangleq \mathbf{H}^{\bar{\ell},j}(\mathbf{x}_t^{\bar{\ell},j} - \hat{\mathbf{x}}_t^{\bar{\ell},j}) + \mathbf{w}_t^{\ell,j}. \tag{8}$$

We propose that the jth node subtracts $\mathbf{H}^{\bar{\ell},j}\hat{\mathbf{x}}_t^{\bar{\ell},j}$ from $\mathbf{y}_t^{\ell,j}$ to compute $\tilde{\mathbf{y}}_t^{\ell,j}$ and then transmits it to the ℓth node at each time t in order to facilitate the local state estimation at the ℓth node. Henceforth, we call $\tilde{\mathbf{y}}_t^{\ell,j}$ as the processed measurements (at the jth node for the ℓth node).

Notice that for each node ℓ, (4) defines the local state transition model. Moreover, the local measurements (see (5)) and the processed measurements to be received from the other nodes together form the overall measurement vector of the ℓth node (for the local state estimation task). Let the overall measurement vector of the ℓth node be denoted with $\tilde{\mathbf{y}}_t^\ell$ and as an example, let $\mathcal{C}^\ell = \{i,j\}$. Then, $\tilde{\mathbf{y}}_t^\ell$ is simply obtained as follows:

$$\tilde{\mathbf{y}}_t^\ell \triangleq \begin{bmatrix}\mathbf{y}_t^\ell \\ \tilde{\mathbf{y}}_t^{\ell,i} \\ \tilde{\mathbf{y}}_t^{\ell,j}\end{bmatrix}.$$

For the ℓth node, we then have the following linear state-space equations:

$$\begin{aligned}\mathbf{x}_t^\ell &= \mathbf{A}^\ell\mathbf{x}_{t-1}^\ell + \mathbf{v}_t^\ell, \\ \tilde{\mathbf{y}}_t^\ell &= \tilde{\mathbf{H}}^\ell\mathbf{x}_t^\ell + \tilde{\mathbf{w}}_t^\ell,\end{aligned} \tag{9}$$

where $\tilde{\mathbf{H}}^\ell$ is determined based on $\mathbf{H}^\ell$, $\mathbf{H}^{\ell,i}$, and $\mathbf{H}^{\ell,j}$, and $\tilde{\mathbf{w}}_t^\ell$ is the noise vector corresponding to $\tilde{\mathbf{y}}_t^\ell$ (see (5) and (7)):

$$\tilde{\mathbf{H}}^\ell \triangleq \begin{bmatrix}\mathbf{H}^\ell \\ \mathbf{H}^{\ell,i} \\ \mathbf{H}^{\ell,j}\end{bmatrix}, \ \tilde{\mathbf{w}}_t^\ell \triangleq \begin{bmatrix}\mathbf{w}_t^\ell \\ \tilde{\mathbf{w}}_t^{\ell,i} \\ \tilde{\mathbf{w}}_t^{\ell,j}\end{bmatrix}. \tag{10}$$

The Kalman filter is an iterative real-time estimator consisting of prediction and measurement update steps at each iteration. For the linear system given in (9), the following equations describe the Kalman filter iteration at time t, where $\hat{\mathbf{x}}_{t|t'}^\ell$ denotes the state estimates of the ℓth node at time t ($t' = t-1$ for prediction and $t' = t$ for measurement update) under the assumption that the correlation between state variables belonging to different local centers is zero, that leads to the independence of $\mathbf{x}_t^\ell$ from $\tilde{\mathbf{w}}_t^{\ell,i}$ and $\tilde{\mathbf{w}}_t^{\ell,j}$ (and hence from $\tilde{\mathbf{w}}_t^\ell$):

Prediction:

$$\begin{aligned}\hat{\mathbf{x}}_{t|t-1}^\ell &= \mathbf{A}^\ell\hat{\mathbf{x}}_{t-1|t-1}^\ell, \\ \mathbf{P}_{t|t-1}^\ell &= \mathbf{A}^\ell\mathbf{P}_{t-1|t-1}^\ell\mathbf{A}^{\ell^{\mathrm{T}}} + \sigma_v^2\mathbf{I}_{N^\ell},\end{aligned} \tag{11}$$

Measurement Update:

$$\begin{aligned}\mathbf{G}_t^\ell &= \mathbf{P}_{t|t-1}^\ell\tilde{\mathbf{H}}^{\ell^{\mathrm{T}}}(\tilde{\mathbf{H}}^\ell\mathbf{P}_{t|t-1}^\ell\tilde{\mathbf{H}}^{\ell^{\mathrm{T}}} + \mathbf{R}_t^\ell)^{-1}, \\ \hat{\mathbf{x}}_{t|t}^\ell &= \hat{\mathbf{x}}_{t|t-1}^\ell + \mathbf{G}_t^\ell(\tilde{\mathbf{y}}_t^\ell - \tilde{\mathbf{H}}^\ell\hat{\mathbf{x}}_{t|t-1}^\ell), \\ \mathbf{P}_{t|t}^\ell &= \mathbf{P}_{t|t-1}^\ell - \mathbf{G}_t^\ell\tilde{\mathbf{H}}^\ell\mathbf{P}_{t|t-1}^\ell,\end{aligned} \tag{12}$$

where $\mathbf{P}_{t|t-1}^\ell$ and $\mathbf{P}_{t|t}^\ell$ denote the estimates of the state covariance matrix of the ℓth node at time t based on the measurements up to $t-1$ and t, respectively. Moreover, $\mathbf{G}_t^\ell$ is the Kalman gain matrix of the ℓth node at time t and $\mathbf{R}_t^\ell$ denotes the covariance matrix of $\tilde{\mathbf{w}}_t^\ell$.

We now go back to the process of obtaining $\tilde{\mathbf{y}}_t^{\ell,j}$ at the jth node. We see through (7) that this process contains estimation

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

errors due to the term $\mathbf{x}_t^{\bar{\ell},j} - \hat{\mathbf{x}}_t^{\bar{\ell},j}$. Our aim is to statistically characterize the estimation errors in order to compute the statistics of $\tilde{\mathbf{w}}_t^{\ell,j}$ (see (8)), which is, in fact, required for an optimal state estimation at the ℓth node. Note that the jth node estimates its local state vector $\mathbf{x}_t^j$ (and hence $\mathbf{x}_t^{\bar{\ell},j}$) via its local Kalman filter. We propose that

$$\hat{\mathbf{x}}_t^{\bar{\ell},j} = \hat{\mathbf{x}}_{t|t-1}^{\bar{\ell},j},$$

where $\hat{\mathbf{x}}_{t|t-1}^{\bar{\ell},j}$ is the sub-vector of $\hat{\mathbf{x}}_{t|t-1}^j$ corresponding to $\mathbf{x}_t^{\bar{\ell},j}$. Then, the following lemma states the distribution of $\tilde{\mathbf{w}}_t^{\ell,j}$.

Lemma 1:

$$\tilde{\mathbf{w}}_t^{\ell,j} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Delta}_t^{\ell,j}), \tag{13}$$

where

$$\mathbf{\Delta}_t^{\ell,j} \triangleq \mathbf{H}^{\bar{\ell},j}\mathbf{P}_{t|t-1}^{\bar{\ell},j}\mathbf{H}^{\bar{\ell},j^{\mathrm{T}}} + \sigma_w^2\,\mathbf{I}_{K^{\ell,j}}, \tag{14}$$

where $\mathbf{P}_{t|t-1}^{\bar{\ell},j}$ is the estimate of the state covariance matrix of $\hat{\mathbf{x}}_t^{\bar{\ell},j}$, that can be obtained from $\mathbf{P}_{t|t-1}^j$ and $K^{\ell,j}$ denotes the size of the vector $\tilde{\mathbf{w}}_t^{\ell,j}$.

Proof. See Appendix A. □

Through (4) and (5), we know that $\mathbf{v}_t^\ell$ and $\mathbf{w}_t^\ell$ are independent zero-mean multivariate Gaussian noise vectors. Furthermore, from Lemma 1, we know that the noise term $\tilde{\mathbf{w}}_t^{\ell,j}$ for the processed measurements (see (7)) is also zero-mean multivariate Gaussian. That implies $\tilde{\mathbf{w}}_t^\ell$ is a zero-mean multivariate Gaussian vector (see (10)). Then, since all the noise terms are Gaussian for the linear system in (9), the local Kalman filter given in (11)–(12) is the optimal state estimator for the ℓth node in minimizing the mean squared state estimation error [61].

Based on Lemma 1, we can write (see (10))

$$\tilde{\mathbf{w}}_t^\ell \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t^\ell),$$

where

$$\begin{aligned}\mathbf{R}_t^\ell &= \begin{bmatrix} \mathrm{E}\big[\mathbf{w}_t^\ell \mathbf{w}_t^{\ell^{\mathrm{T}}}\big] & \mathrm{E}\big[\mathbf{w}_t^\ell \tilde{\mathbf{w}}_t^{\ell,i^{\mathrm{T}}}\big] & \mathrm{E}\big[\mathbf{w}_t^\ell \tilde{\mathbf{w}}_t^{\ell,j^{\mathrm{T}}}\big] \\ \mathrm{E}\big[\tilde{\mathbf{w}}_t^{\ell,i} \mathbf{w}_t^{\ell^{\mathrm{T}}}\big] & \mathrm{E}\big[\tilde{\mathbf{w}}_t^{\ell,i} \tilde{\mathbf{w}}_t^{\ell,i^{\mathrm{T}}}\big] & \mathrm{E}\big[\tilde{\mathbf{w}}_t^{\ell,i} \tilde{\mathbf{w}}_t^{\ell,j^{\mathrm{T}}}\big] \\ \mathrm{E}\big[\tilde{\mathbf{w}}_t^{\ell,j} \mathbf{w}_t^{\ell^{\mathrm{T}}}\big] & \mathrm{E}\big[\tilde{\mathbf{w}}_t^{\ell,j} \tilde{\mathbf{w}}_t^{\ell,i^{\mathrm{T}}}\big] & \mathrm{E}\big[\tilde{\mathbf{w}}_t^{\ell,j} \tilde{\mathbf{w}}_t^{\ell,j^{\mathrm{T}}}\big] \end{bmatrix} \\ &= \begin{bmatrix} \sigma_w^2\,\mathbf{I}_{K^\ell} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Delta}_t^{\ell,i} & \mathrm{E}\big[\tilde{\mathbf{w}}_t^{\ell,i} \tilde{\mathbf{w}}_t^{\ell,j^{\mathrm{T}}}\big] \\ \mathbf{0} & \mathrm{E}\big[\tilde{\mathbf{w}}_t^{\ell,j} \tilde{\mathbf{w}}_t^{\ell,i^{\mathrm{T}}}\big] & \mathbf{\Delta}_t^{\ell,j} \end{bmatrix}. \end{aligned} \tag{15}$$

Here, we observe that the computation of $\mathbf{R}_t^\ell$ requires the correlation between $\tilde{\mathbf{w}}_t^{\ell,i}$ and $\tilde{\mathbf{w}}_t^{\ell,j}$. Using (8), we can write

$$\mathrm{E}\Big[\tilde{\mathbf{w}}_t^{\ell,i} \tilde{\mathbf{w}}_t^{\ell,j^{\mathrm{T}}}\Big] = \mathbf{H}^{\bar{\ell},i}\,\mathrm{E}\Big[\big(\mathbf{x}_t^{\bar{\ell},i} - \hat{\mathbf{x}}_t^{\bar{\ell},i}\big)\big(\mathbf{x}_t^{\bar{\ell},j} - \hat{\mathbf{x}}_t^{\bar{\ell},j}\big)^{\mathrm{T}}\Big]\mathbf{H}^{\bar{\ell},j^{\mathrm{T}}},$$

which requires the correlation between the state estimation errors for $\mathbf{x}_t^{\bar{\ell},i}$ and $\mathbf{x}_t^{\bar{\ell},j}$. However, since it is possible that $\mathbf{x}_t^{\bar{\ell},\bar{j},i} \triangleq \mathbf{x}_t^{\bar{\ell},i} \backslash \mathbf{x}_t^{\bar{\ell},j} \neq \emptyset$ or $\mathbf{x}_t^{\bar{\ell},\bar{i},j} \triangleq \mathbf{x}_t^{\bar{\ell},j} \backslash \mathbf{x}_t^{\bar{\ell},i} \neq \emptyset$ and the correlations for such state variables that belong to different local centers are not computed over the network, we approximate the correlation terms involving the state variables in $\mathbf{x}_t^{\bar{\ell},\bar{j},i}$ and $\mathbf{x}_t^{\bar{\ell},\bar{i},j}$ as zero. Notice that the same assumption is made in the local Kalman filter equations above and in fact, it is the only assumption consistently made in this study, that looses the optimality of the proposed solution. Through simulations, we observe that it only slightly increases the state estimation error compared to the optimal centralized Kalman filter. Hence, in the rest of the design and analysis (in Sec. V and Sec. VI), we assume that the proposed distributed dynamic state estimator achieves near-optimal performance. Finally, we note that the main contributions of the proposed distributed dynamic state estimator are in the processing of the measurements that are acquired at the other local centers $j \in \mathcal{C}^\ell$ and only partially relevant to the local state vector $\mathbf{x}_t^\ell$.

Remark 2: Each node ℓ needs the knowledge of $\mathbf{R}_t^\ell$ for the measurement update step of its local Kalman filter, where computing $\mathbf{R}_t^\ell$ requires $\mathbf{\Delta}_t^{\ell,\bar{j}}, \forall j \in \mathcal{C}^\ell$ (see (15)). We observe through (14) that $\mathbf{\Delta}_t^{\ell,j}$ depends on $\mathbf{H}^{\bar{\ell},j}$ and $\mathbf{P}_{t|t-1}^{\bar{\ell},j}$. Here, $\mathbf{H}^{\bar{\ell},j}$ is determined based on the network topology, which is available at every node. On the other hand, $\mathbf{P}_{t|t-1}^{\bar{\ell},j}$ is extracted from the state covariance matrix of the jth node, $\mathbf{P}_{t|t-1}^j$, which is primarily computed at the jth node. Nevertheless, we see through (11) and (12) that the (iterative) computation of $\mathbf{P}_{t|t-1}^j$ does not depend on online sensor measurements and hence it can be computed offline at each node in the network. Moreover, in the proposed distributed trust management scheme (see Sec. VI), each node needs to compute the state covariance matrices of all other nodes. Hence, the proposed state estimation mechanism does not introduce further computational complexity beyond the trust management scheme.

V. Secure State Estimation against Measurement Anomalies

The state estimator proposed in the previous section is based on the assumption that the network operation always fits to the nominal system model. However, in practice, various kinds of anomaly might appear all over the network, e.g., measurement anomalies, due to cyber-attacks or network faults. In order to achieve secure state estimation against measurement anomalies, we propose to quickly and reliably detect them and then eliminate their effects as much as possible.

Considering that the attackers can be advanced, strategic, or adaptive to the system and detector dynamics, it is hard to model all attack types [7], [8]. Moreover, considering the complex cyber-physical nature of the smart grid, it is also difficult to model all types of network faults. Hence, in this study, we do not focus on particular anomaly types and rather we assume the anomaly type is totally unknown. On the other hand, once we detect an anomaly, this prevents us to recover the useful part of the anomalous measurements (if any). Our anomaly mitigation strategy is then to reject/neglect the anomalous measurements in the state estimation process and predict the system state, until the system is recovered back to the regular operating conditions, based on (i) the previous reliable state estimates, securely recorded in the distributed ledger and (ii) the nominal system model. As the sensors are distributed over the network, each node analyzes only its local measurements. We next explain the proposed measurement anomaly detection scheme at the ℓth node and then the corresponding state recovery over the network.

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

A. Real-time Detection of Local Measurement Anomalies

During the regular system operation, it is possible to observe infrequent outliers and the Kalman filter is known to be effective to compensate (suppress) small errors due to such infrequent outliers [62]. Hence, we are particularly interested in long-term anomalies where there exist a temporal relation between anomalous measurements. Our aim is to detect such anomalies timely and reliably using the measurements that become available sequentially over time. In this problem, although we can statistically characterize the nominal measurements sufficiently accurately based on the nominal system model and the online state estimates under regular system operation, measurements can take various unknown statistical forms in case of an anomaly. Hence, we follow a solution strategy in that we derive and monitor (over time) a univariate summary statistic that is informative about a possible deviation of the online measurements from their nominal model, as detailed next.

At each time t, the ℓth node locally observes $\mathbf{y}_t^\ell$ (see (5)). Moreover, based on the local Kalman filter, we have

$$\mathbf{x}_t^\ell - \hat{\mathbf{x}}_{t|t-1}^\ell \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{t|t-1}^\ell). \tag{16}$$

Using (5) and (16), we can write

$$\begin{aligned}\mathbf{y}_t^\ell - \mathbf{H}^\ell \hat{\mathbf{x}}_{t|t-1}^\ell &= \mathbf{H}^\ell(\mathbf{x}_t^\ell - \hat{\mathbf{x}}_{t|t-1}^\ell) + \mathbf{w}_t^\ell \\ &\sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_t^\ell),\end{aligned} \tag{17}$$

where

$$\mathbf{\Sigma}_t^\ell \triangleq \mathbf{H}^\ell \mathbf{P}_{t|t-1}^\ell \mathbf{H}^{\ell^{\mathrm{T}}} + \sigma_w^2 \mathbf{I}_{K^\ell}.$$

Then, based on (17),

$$\chi_t^\ell \triangleq \left(\mathbf{y}_t^\ell - \mathbf{H}^\ell \hat{\mathbf{x}}_{t|t-1}^\ell\right)^{\mathrm{T}} \mathbf{\Sigma}_t^{\ell^{-1}} \left(\mathbf{y}_t^\ell - \mathbf{H}^\ell \hat{\mathbf{x}}_{t|t-1}^\ell\right) \tag{18}$$

is a chi-squared random variable with K^ℓ degrees of freedom. Notice that χ_t^ℓ has a time-invariant (stationary) distribution under regular system operation.

Let $F_{K^\ell}(\cdot)$ be the cumulative distribution function (cdf) of a chi-squared random variable with K^ℓ degrees of freedom. If the p-value (right tail probability) corresponding to χ_t^ℓ satisfies

$$p_t^\ell \triangleq 1 - F_{K^\ell}(\chi_t^\ell) < \alpha, \tag{19}$$

then the corresponding local measurement vector $\mathbf{y}_t^\ell$ is considered as an outlier for the significance level of α. In case of an anomaly, we expect that the chi-squared statistic χ_t^ℓ takes higher values compared to its nominal values and hence, we expect to observe more frequent outliers. Then, we can model an anomaly as persistent outliers, as in [63] and [64].

Based on (19), for an outlier $\mathbf{y}_t^\ell$, we have

$$s_t^\ell \triangleq \log\left(\frac{\alpha}{p_t^\ell}\right) > 0, \tag{20}$$

and similarly, for a non-outlier $\mathbf{y}_t^\ell$, we have $s_t^\ell \leq 0$. Hence, we can consider s_t^ℓ as a (positive/negative) statistical evidence for anomaly at time t. Then, similar to the accumulation of the log-likelihood ratios in the well-known cumulative sum (CUSUM) test, we can accumulate s_t^ℓ's over time and declare a measurement anomaly only if there is a strong/reliable evidence, that results in the following CUSUM-like test [63]:

$$\begin{aligned}\Gamma^\ell &= \inf\{t : g_t^\ell \geq h\}, \\ g_t^\ell &= \max\{0, g_{t-1}^\ell + s_t^\ell\},\end{aligned} \tag{21}$$

where $g_0^\ell = 0$ and Γ^ℓ denotes the stopping time at which an anomaly is detected at the ℓth node.

Let τ^ℓ be the unknown change-point at which a local measurement anomaly happens at the ℓth node and continues thereafter. The CUSUM test always keeps the maximum likelihood estimate of the change-point and update it as the measurements become available over time [65, Sec. 2.2]. Let $\hat{\tau}^\ell$ be the change-point estimate of the proposed test. Initializing $\hat{\tau}^\ell = 0$ at $t = 0$, whenever the decision statistic g_t^ℓ reaches zero, we make the following update: $\hat{\tau}^\ell \leftarrow t$. In other words, $\hat{\tau}^\ell$ is the latest time-instant at which the decision statistic reaches zero. The final change-point estimate is determined when an anomaly is declared at the stopping time Γ^ℓ. Hence, we have

$$\hat{\tau}^\ell \triangleq \max\{t : g_t^\ell = 0, t < \Gamma^\ell\}.$$

The change-point estimate will be useful for state recovery (see Sec. V-B).

For the CUSUM-like test in (21), to achieve a lower false alarm rate (equivalently a larger average false alarm period), the significance level α is chosen smaller and/or the test threshold h is chosen higher, that, on the other hand, leads to larger detection delays (see (20) and (21)). Let $\mathrm{E}_\infty[\Gamma^\ell]$ be the average false alarm period, i.e., the average stopping time in the absence of anomalies ($\tau^\ell = \infty$). The following corollary (to Theorem 2 of [63]) describes how to choose α and h to obtain a desired lower bound L on the average false alarm period.

Corollary 1: For a chosen $\alpha < 1/e$ and

$$h \geq \frac{\log(L)}{1 - W(\alpha \log(\alpha))/\log(\alpha)}, \tag{22}$$

we have

$$\mathrm{E}_\infty[\Gamma^\ell] \geq L,$$

where $W(\cdot)$ denotes the Lambert-W function[1].

Proof. See Appendix B. □

B. State Recovery

Once the proposed CUSUM-like detection scheme in (21) declares an anomaly, our purpose is to recover the current (and the future) state estimates. Since the local measurements observed after the (unknown) change-point, i.e., $\mathbf{y}_t^\ell, t > \tau^\ell$, are not reliable, we can recover the state estimates from the latest reliable state estimates computed at the change-point estimate $\hat{\tau}^\ell$.

The smart grid is a highly interconnected network, as in the proposed mechanism, the local state estimation is performed using the local measurements as well as the processed measurements received from some other nodes. Hence, if an

[1]There exists a built-in MATLAB function **lambertw**.

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

anomaly happens at a node, the whole network is affected by the anomaly to some extent. Then, whenever a measurement anomaly is detected at a node, say the ℓth one at time t, the ℓth node immediately broadcasts $\hat{\tau}^\ell$ to the entire network. Then every node $j \in \{1, \ldots, L\}$ makes the following state recovery:

$$\hat{\mathbf{x}}^j_{t|t} = \mathbf{A}^{j^{(t-\hat{\tau}^\ell)}} \hat{\mathbf{x}}^j_{\hat{\tau}^\ell|\hat{\tau}^\ell}, \tag{23}$$

that essentially corresponds to the case where we replace all the measurements $\tilde{\mathbf{y}}^j_t$ during the anomaly interval $(\hat{\tau}^\ell, t]$ with the corresponding pseudo measurements $\hat{\mathbf{y}}^j_t \triangleq \tilde{\mathbf{H}}^j \hat{\mathbf{x}}^j_{t|t-1}$, that makes the measurement innovation signal zero (see (11) and (12)).

The regular network operation requires the participation of every node in the state estimation process. Hence, whenever a measurement anomaly is detected at the ℓth node, we propose to raise an alarm flag, calling for further investigation at the ℓth subregion and the neighboring subregions $j \in \mathcal{C}^\ell$ considering the possibility that the processed measurements received from the neighboring nodes may also lead to an anomaly in the local state estimation process. The investigation should be performed considering also the possibility of false alarms. After the investigation process and a possible recovery of the system, the predesigned regular network operation is restarted. Our main purpose here is to decrease the state estimation errors due to anomalies and hence to provide more reliable state estimates during the anomaly mitigation/system recovery period.

Remark 3: For the state recovery, the distributed ledger consisting of the M recent blocks needs to include the state estimates computed at the change-point estimate $\hat{\tau}^\ell$, where $\hat{\tau}^\ell$ is not known ahead of time. However, since we expect quick detection, we do not expect to observe a $\hat{\tau}^\ell$ that is quite far away to the stopping time Γ^ℓ. Hence, in practice M can be chosen reasonably small. In the case where the ledger does not contain the state estimates of the time $\hat{\tau}^\ell$, we can recover the state estimates from the oldest state estimates available in the ledger considering that they are more reliable compared to the other alternatives.

VI. Distributed Trust Management

In BC-based distributed networks, malicious adversaries may obtain illegitimate access to the system, e.g., via stealing the digital identity of some nodes, malware propagation, etc., [11], [66], and additionally some nodes may get faulty during the system operation. Moreover, as the network is fully-distributed, there is no centralized trusted node to check whether all nodes are safe and trustable, i.e., whether the nodes are functioning according to the predesigned network rules. Therefore, against the possibility of misbehaving nodes, we need a distributed trust management mechanism over the network, in which all nodes collectively verify the trustability of each node. Recall that every node knows (i) the nominal system model and the network configuration, and (ii) a finite history of recent state estimates of all the nodes stored in the shared distributed ledger (see Sec. III-B.2). Using only (i) and (ii), each node votes on the trustability of all other nodes. Then, at each time, trustability of each node is decided via majority-voting. We explain below how the ℓth node is evaluated by the other regular (non-misbehaving) nodes.

Suppose that at an unknown time η^ℓ, an unexpected event happens at the ℓth node: the node gets faulty or an attacker hacks and takes control of the node. Then, we can no longer expect that the behavior of the ℓth node fits to its pre-defined regular operation. Furthermore, similar to the measurement anomalies, it is quite difficult to model the (anomalous) behavior of the ℓth node after time η^ℓ. We would like to detect misbehaving nodes as quickly as possible to timely mitigate the corresponding effects on the state estimation process. Hence, for the evaluation of the ℓth node, we propose that each node decides whether the state estimates provided by the ℓth node exhibit an anomalous pattern over time. In this direction, we next derive the nominal evolution (over time) model of the local state estimates of the ℓth node. Then, similar to Sec. V-A, we derive a univariate summary statistic that is informative about a possible deviation of the local state estimates of the ℓth node from their nominal evolution model and monitor this statistic over time.

Based on the local Kalman filter iteration of the ℓth node at time t (see (11) and (12)), we can write

$$\begin{aligned}\hat{\mathbf{x}}^\ell_{t|t} &= \hat{\mathbf{x}}^\ell_{t|t-1} + \mathbf{G}^\ell_t\big(\tilde{\mathbf{y}}^\ell_t - \tilde{\mathbf{H}}^\ell \hat{\mathbf{x}}^\ell_{t|t-1}\big) \\ &= \mathbf{A}^\ell \hat{\mathbf{x}}^\ell_{t-1|t-1} + \mathbf{G}^\ell_t\big(\tilde{\mathbf{H}}^\ell(\mathbf{x}^\ell_t - \hat{\mathbf{x}}^\ell_{t|t-1}) + \tilde{\mathbf{w}}^\ell_t\big) \qquad (24)\\ &\sim \mathcal{N}\big(\mathbf{A}^\ell \hat{\mathbf{x}}^\ell_{t-1|t-1}, \mathbf{\Psi}^\ell_t\big), \qquad (25)\end{aligned}$$

where

$$\mathbf{\Psi}^\ell_t \triangleq \mathbf{G}^\ell_t\big(\tilde{\mathbf{H}}^\ell \mathbf{P}^\ell_{t|t-1} \tilde{\mathbf{H}}^{\ell^{\mathrm{T}}} + \mathbf{R}^\ell_t\big)\mathbf{G}^{\ell^{\mathrm{T}}}_t. \tag{26}$$

Here, (24) is obtained using (9) and (11). Moreover, $\mathbf{\Psi}^\ell_t$ is obtained using $\mathbf{x}^\ell_t - \hat{\mathbf{x}}^\ell_{t|t-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}^\ell_{t|t-1})$ (see (16)), $\tilde{\mathbf{w}}^\ell_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}^\ell_t)$, and approximating the correlation terms involving the state variables belonging to different local centers as zero, as in Sec. IV.

Notice that (25) statistically characterizes the local state estimates at time t, given the local state estimates at time $t-1$, under regular system operation. Based on (25), we can write

$$\hat{\mathbf{x}}^\ell_{t|t} - \mathbf{A}^\ell \hat{\mathbf{x}}^\ell_{t-1|t-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Psi}^\ell_t),$$

that implies

$$\pi^\ell_t \triangleq \big(\hat{\mathbf{x}}^\ell_{t|t} - \mathbf{A}^\ell \hat{\mathbf{x}}^\ell_{t-1|t-1}\big)^{\mathrm{T}} \mathbf{\Psi}^{\ell^{-1}}_t \big(\hat{\mathbf{x}}^\ell_{t|t} - \mathbf{A}^\ell \hat{\mathbf{x}}^\ell_{t-1|t-1}\big) \tag{27}$$

is a chi-squared random variable with N^ℓ degrees of freedom under regular system operation.

If node ℓ is misbehaving, we expect that the local state estimates provided by it deviate from the nominal evolution model in (25), that makes π^ℓ_t larger than its nominal values. Then, similar to the real-time detection of measurement anomalies in Sec. V-A, the following CUSUM-like test can be employed at a regular node j to decide on the trustability of the ℓth node:

$$\begin{aligned}&\Gamma^\ell_j = \inf\{t : g^\ell_{j,t} \geq h^\ell\},\ g^\ell_{j,t} = \max\{0, g^\ell_{j,t-1} + s^\ell_{j,t}\},\\ &s^\ell_{j,t} = \log\left(\alpha^\ell / p^\ell_{j,t}\right),\ p^\ell_{j,t} = 1 - F_{N^\ell}(\pi^\ell_t), \qquad (28)\end{aligned}$$

where $g^\ell_{j,t}$ is the decision statistic at time t, $g^\ell_{j,0} = 0$, $h^\ell > 0$ is the test threshold, $s^\ell_{j,t}$ is the statistical evidence

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

(for misbehavior) at time t, α^ℓ is the significance level, $p^\ell_{j,t}$ is the p-value corresponding to π^ℓ_t, and $F_{N^\ell}(\cdot)$ denotes the cdf of a chi-squared random variable with N^ℓ degrees of freedom. A regular node j then evaluates the ℓth node as trusted until time Γ^ℓ_j, and misbehaving after Γ^ℓ_j. Furthermore, as before, the unknown change-point η^ℓ is estimated by the jth node as the latest time-instant at which the decision statistic reaches zero before the stopping time Γ^ℓ_j:

$$\hat{\eta}^\ell_j \triangleq \max\{t : g^\ell_{j,t} = 0, t < \Gamma^\ell_j\}.$$

The overall decision on the trustability/misbehavior of the ℓth node is made by the majority of the nodes. Let the vote of the jth node on the trustability of the ℓth node at time t be denoted with a binary variable $d^\ell_{j,t}$, where $d^\ell_{j,t} = 0$ or $d^\ell_{j,t} = 1$ if the jth node evaluates the ℓth node as trusted or misbehaving, respectively. Notice that if the jth node is regular, then it votes based on the designed test in (28) that gives rise to

$$d^\ell_{j,t} = \mathbb{I}(\Gamma^\ell_j \leq t).$$

The time at which the ℓth node is declared misbehaving over the network is then determined as follows:

$$\Gamma^\ell_{\text{net}} \triangleq \inf\left\{t : \sum_{j\,\in\,\{1,\ldots,L\}\backslash\{\ell\}} d^\ell_{j,t} > \frac{L-1}{2}\right\}. \tag{29}$$

Notice that this decision mechanism works as intended unless the majority of the nodes are misbehaving. In other words, as long as the majority of the nodes regularly employ the proposed misbehavior detection scheme in (28) and vote accordingly, then the trustability of the ℓth node is reliably evaluated over the network, for every node ℓ.

Under the nominal system operation, all nodes are regular and hence the misbehavior detector (for the ℓth node) in (28) is identical at all the nodes j in the network. Then, the false alarm rate of (29) is equal to the false alarm rate of (28). Hence, the proposed trust management scheme achieves the same false alarm guarantees through Corollary 1 (after replacing the parameters α and h in the corollary with α^ℓ and h^ℓ, respectively).

If (29) declares the ℓth node as misbehaving, an alarm flag is raised, calling for an investigation at the ℓth node and the neighboring nodes $j \in \mathcal{C}^\ell$. Moreover, due to the inter-node data exchanges, a node misbehavior affects all the nodes in the network to some extent. Then, as before, until the system is recovered back to the regular operating conditions, the local states of each node $j \in \{1, \ldots, L\}$ can be predicted based on the nominal system model and the latest reliable estimates computed at the change-point estimate $\hat{\eta}^\ell_j$ as follows:

$$\hat{\mathbf{x}}^j_{t|t} = \mathbf{A}^{j\,(t-\hat{\eta}^\ell_j)}\hat{\mathbf{x}}^j_{\hat{\eta}^\ell_j|\hat{\eta}^\ell_j},$$

where $\hat{\mathbf{x}}^j_{\hat{\eta}^\ell_j|\hat{\eta}^\ell_j}$ can be obtained from the distributed ledger.

The proposed trust management scheme requires that each node $j \in \{1, \ldots, L\}\backslash\{\ell\}$ computes π^ℓ_t and $\mathbf{\Psi}^\ell_t$ for every other node $\ell \in \{1, \ldots, L\}\backslash\{j\}$ (see (26)–(29)). Since the nominal system model is known by every node and the local state estimates provided by the other nodes are available (via the distributed ledger) to every node, the node j already knows $\mathbf{A}^\ell$, $\tilde{\mathbf{H}}^\ell$, $\hat{\mathbf{x}}^\ell_{t|t}$, and $\hat{\mathbf{x}}^\ell_{t-1|t-1}$ at each time t for every node ℓ. On the other hand, $\mathbf{R}^\ell_t$, $\mathbf{G}^\ell_t$, and $\mathbf{P}^\ell_{t|t-1}$ are not directly available to the jth node. Fortunately, the Kalman gain matrix $\mathbf{G}^\ell_t$ and the estimate of the state covariance matrix $\mathbf{P}^\ell_{t|t-1}$ can be computed offline without requiring online sensor measurements. Moreover, $\mathbf{R}^\ell_t$ is computed based on the estimates of the state covariance matrices (see also Remark 2). Hence, at each node j, we propose to compute $\mathbf{G}^\ell_t$ and $\mathbf{P}^\ell_{t|t-1}$ (iteratively through (11) and (12)) for every other node ℓ in the network.

Remark 4: Both the local measurement anomaly detection and the distributed trust management schemes depend on a CUSUM-like anomaly detection algorithm, whose general form is first presented in [63] along with an asymptotic performance analysis. On the other hand, the summary statistics used in the both schemes are designed in a novel task-specific manner. Further, although voting-based consensus is not new, achieving consensus based on the real-time misbehavior detection algorithms employed at every regular node is novel. Finally, the proposed state recovery scheme is novel, as after an anomaly/misbehavior is detected in real-time based on the sequence of the derived summary statistics, the change-point is estimated and then the state estimates are recovered accordingly from the latest reliable state estimates without assuming any anomaly/misbehavior model.

VII. Summary of the Proposed Mechanism

We summarize the proposed procedure at the ℓth node in Fig. 4, where the procedure is identical at every node. Since the proposed mechanism requires an investigation after the detection of a measurement anomaly at any node or a misbehavior of any node, the overall stopping time of the network is given by

$$\Gamma_{\text{net}} \triangleq \inf\{\Gamma^\ell, \Gamma^\ell_{\text{net}}, \ell = 1, \ldots, L\}. \tag{30}$$

If several detection mechanisms simultaneously give alarms, we can recover the state estimates based on the oldest among the corresponding change-point estimates. Moreover, if the state estimates corresponding to the (oldest) change-point estimate are not included in the distributed ledger consisting of the M recent blocks, then we can choose $t_R = t - M + 1$ as the state recovery point, that corresponds to the oldest state estimates available in the ledger, where the state recovery point is denoted with t_R in Fig. 4. In case an anomaly is declared over the network, after an investigation and possibly the recovery of the system, the proposed mechanism is restarted.

VIII. Simulation Results

In this section, we evaluate the performance of the proposed mechanism via simple case studies over an IEEE-14 bus power system. We consider a special case where system state includes the voltage phase angles only and the measurements are the real power flow and power injection measurements (assuming the voltage magnitudes are known). The system consists of 4 subregions, $N = 14$ buses, and $K = 23$ sensors (see Fig. 5). The bus 6 is chosen as the reference bus, the state transition matrix is chosen to be an identity matrix, and the measurement matrix $\mathbf{H}$ is determined based on the network topology. The

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

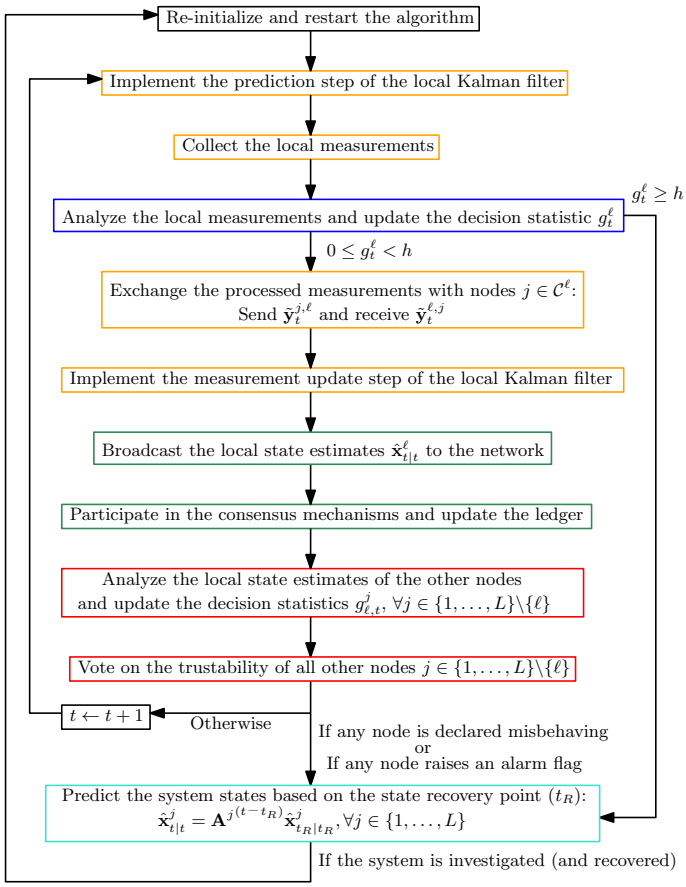

Fig. 4: Flowchart of the proposed procedure at the ℓth node. The green, orange, blue, and red rectangles mainly refer to the components of the proposed solution design discussed in Sec. III, Sec. IV, Sec. V, and Sec. VI, respectively. Moreover, the turquoise rectangle refers to the state recovery mechanism discussed in Sec. V and Sec. VI.

noise variances are chosen as $\sigma_v^2 = \sigma_w^2 = 10^{-4}$, and the initial state variables (voltage phase angles) are determined via the DC optimal power flow algorithm for case-14 in MATPOWER [67]. For the proposed detection schemes, to achieve $\mathrm{E}_\infty[\Gamma^\ell], \mathrm{E}_\infty[\Gamma_{\text{net}}^\ell] \geq 10^6, \forall \ell \in \{1,2,3,4\}$, we choose $\alpha = \alpha^1 = \cdots = \alpha^4 = 0.2$ and $h = h^1 = \cdots = h^4 = 21.3527$ (see (22)). Then, we obtain via a Monte Carlo simulation that the average false alarm period of the network (see (30)) is $\mathrm{E}_\infty[\Gamma_{\text{net}}] \backsimeq 1.26 \times 10^6$. Moreover, we choose the number of blocks in the distributed ledger as $M = 200$. In the following, we present simulation results firstly for a measurement anomaly case and then for a node misbehavior case.

A. Case 1: Measurement Anomaly

As an example, we consider FDI attacks against the sensor measurements in the subregion ℓ:

$$\mathbf{y}_t^\ell = \mathbf{H}^\ell \mathbf{x}_t^\ell + \mathbf{a}_t^\ell + \mathbf{w}_t^\ell, \ t \geq \tau^\ell, \tag{31}$$

where $\mathbf{a}_t^\ell = [a_{1,t}^\ell, \ldots, a_{K^\ell,t}^\ell]^{\mathrm{T}}$ denotes the injected false data at time t. We assume that subregions 1 and 2 are under FDI attack after time $\tau \triangleq \tau^1 = \tau^2$ with $a_{i,t}^\ell \sim \mathcal{U}[0,\rho], \forall i \in \{1,\ldots,K^\ell\}$, $\forall \ell \in \{1,2\}$, $\forall t \geq \tau$, where $\mathcal{U}[0,\rho]$ denotes a uniform random variable in the range of $[0,\rho]$.

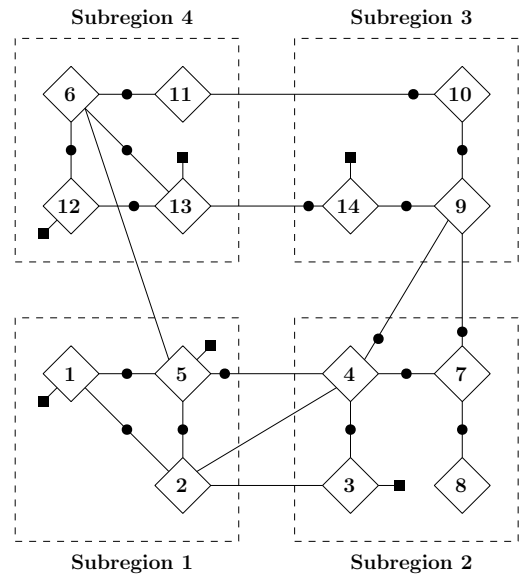

Fig. 5: Wide-area monitoring model of the IEEE-14 bus power system. Four subregions are illustrated with dashed rectangles. The circles on the branches represent the power flow measurements, and the squares represent the power injection measurements.

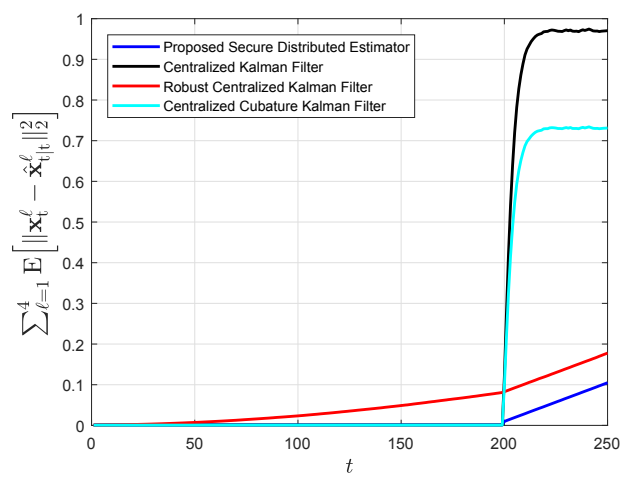

Fig. 6: Mean squared state estimation error vs. time in case of an FDI attack (with $\rho = 0.3$) performed after time $\tau = 200$ targeting the sensor measurements in the subregions 1 and 2.

Firstly, assuming $\tau = 200$ and $\rho = 0.3$, we present in Fig. 6 the sum of the mean squared state estimation errors over all local centers for both the pre-attack period, i.e., $t < 200$, and for the attacking period of $t \in [200, 250]$. We present the performance of the proposed distributed secure state estimation mechanism, the centralized Kalman filter, the centralized cubature Kalman filter [68] (since it is reported to be robust against attacks [34] although mainly designed for nonlinear systems), and the robust centralized Kalman filter that rejects gross outliers and replaces the corresponding measurements with the pseudo measurements, similar to [32]. In particular, the robust Kalman filter computes a chi-squared statistic at each time t using all the measurements $\mathbf{y}_t$, similar to (18), and compute the corresponding p-value based on the chi-squared distribution with K degrees of freedom. The significance level

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

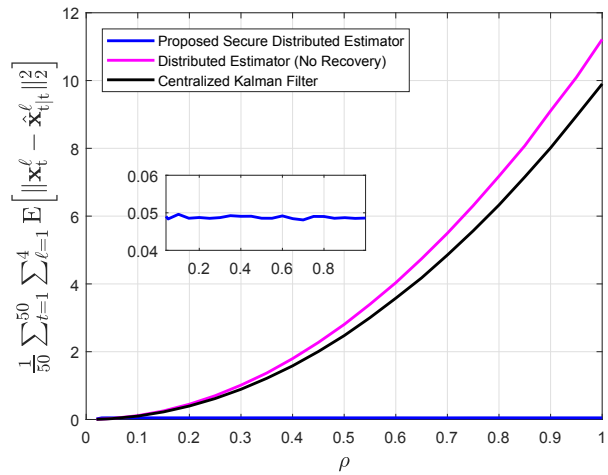

Fig. 7: Average state estimation error over an attacking period vs. attack magnitudes in case of FDI attacks to the sensor measurements in the subregions 1 and 2.

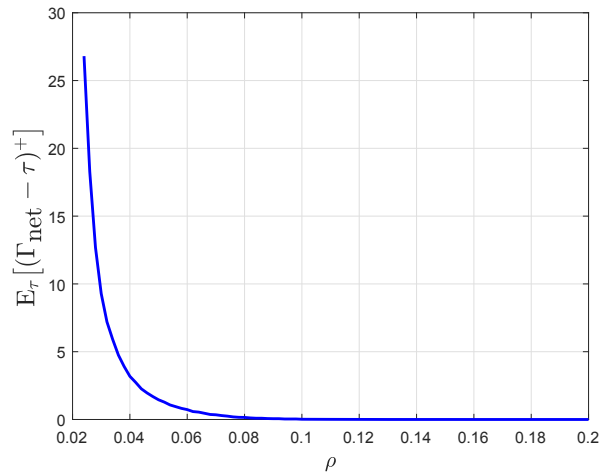

Fig. 8: Average detection delay vs. attack magnitudes in case of FDI attacks targeting the subregions 1 and 2.

for outliers is chosen as 0.01. Then, if the p-value is less than 0.01, the corresponding measurements $\mathbf{y}_t$ are replaced with the pseudo measurements $\hat{\mathbf{y}}_t \triangleq \mathbf{H}\hat{\mathbf{x}}_{t|t-1}$ in the robust centralized Kalman filter where $\hat{\mathbf{x}}_{t|t-1}$ denotes the predicted states at time t. We observe via Fig. 6 that the proposed estimator performs near-optimally in the pre-attack period and reduces the state estimation errors in the attacking period. Since the proposed mechanism does not use any measurements after the anomaly detection, state estimation/prediction error increases over time in the attacking period, as expected. Hence, after anomaly detection, the anomaly needs to be quickly mitigated to prevent large errors. We also observe that although the cubature Kalman filter shows some performance improvement over the Kalman filter in the attacking regime, it still significantly deviates from the actual system state after the attack is being launched. Moreover, although the robust Kalman filter reduces errors over the attacking period, it deviates from the system state under regular conditions due to the rejection of nominal outliers.

Next, we present in Fig. 7 the average state estimation performance over an attacking period as the FDI attack magnitude ρ varies. Particularly, we assume $\tau = 1$ and compute the sum of the mean squared state estimation errors over local centers, averaged over an attacking period of $t \in [1, 50]$. Recall that after the detection of measurement anomalies and until the system is back to the regular conditions, the proposed mechanism predicts the system state based on the nominal system model and the previous reliable state estimates. Hence, after the detection, the state estimation/prediction performance of the proposed mechanism does not depend on the attack magnitudes. On the other hand, for the centralized Kalman filter and the distributed state estimator designed assuming ideal operating conditions (see Sec. IV), the state estimation error increases with the attack magnitude as there are no anomaly detection/state recovery mechanisms integrated into them.

Finally, we illustrate in Fig. 8 how the average detection delay varies as the attack magnitude/intensity changes, assuming $\tau = 1$, that results in the worst-case detection delays since $g_t = 0$ at $t = 0$. As expected, for stronger attacks, we observe more significant outliers (smaller p-values) that leads to quicker detection (see (20) and (21)).

B. Case 2: Node Misbehavior

We assume that at time η, the node 3 is hacked and controlled by a malicious entity that injects false data to the local measurements and also does not employ the local measurement anomaly detection mechanism. Similar to the Case 1 above, we consider the FDI attack in (31) with $a_{i,t}^3 \sim \mathcal{U}[0, \rho]$, $\forall i \in \{1, \ldots, 5\}, \forall t \geq \eta$. The difference is that the node 3 does not notify the other local centers about the local measurement anomaly at all. In this case, we expect that the misbehavior is detected over the network via the distributed trust management scheme. Assuming $\eta = 1$, we firstly present in Fig. 9 the average detection delay versus the attack magnitude curve. Then, we present in Fig. 10 the average state estimation performance over an attacking period of $t \in [1, 50]$. We observe that via the trust management scheme and the state recovery mechanism employed after misbehavior detection, the state estimation error can be significantly reduced compared to the estimators with no recovery mechanisms.

IX. Concluding Remarks

We have studied the secure distributed dynamic state estimation in the wide-area smart grids. We have proposed a novel BC-based resilient mechanism in which each local center supervises a subregion and collaborates with the other local centers to perform reliable state estimation in a fully-distributed manner under adversarial settings. We have proposed to enhance the security of each part of the network against attacks and data manipulations. Particularly, we reduce the risk of manipulations in the network database and the network communication channels via cryptography and voting-consensus mechanisms. Moreover, we timely and reliably

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

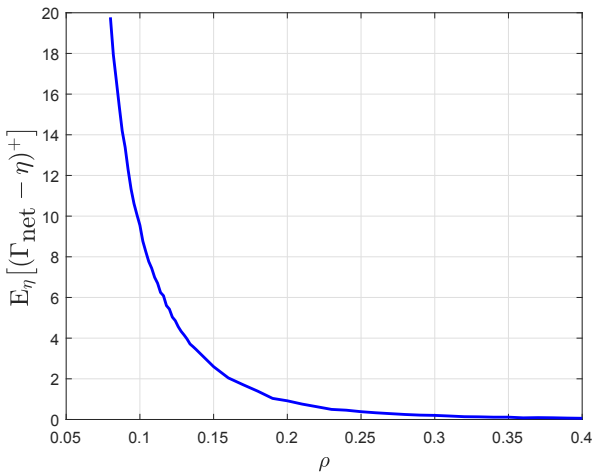

Fig. 9: Average detection delay vs. attack magnitudes. False data are injected at the local measurements of the node 3 and no anomaly detection mechanism is employed at the node 3.

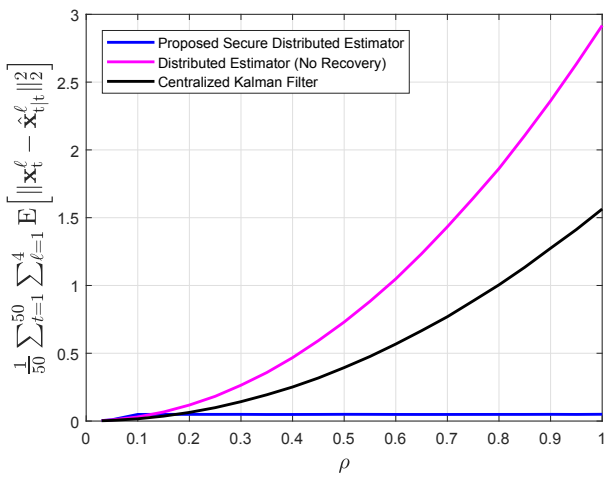

Fig. 10: Average state estimation error over an attacking period vs. attack magnitudes in case of a node misbehavior.

detect and eliminate the effects of measurement anomalies and misbehavior of the local control centers. Numerical studies illustrate the promising performance of the proposed mechanism.

We consider this work as a proposal for secure distributed state estimation in the wide-area smart grids, where all parts of the solution design, namely the BC mechanism, the distributed state estimation mechanism, the measurement anomaly detection scheme, the trust management scheme, and the state recovery scheme are open to improvements. Moreover, although we have specifically focused on the power networks, the proposed mechanism can be used for any linear dynamic networked control system after simple modifications. Particularly, the following future research directions can be considered:

- The proposed BC-based mechanism introduces many advantages in terms of data security, availability, accessibility, and process transparency, etc., however, some new challenges arise such as storage redundancy over the network, higher network congestion, synchronization issues due to possible communication delays, higher computational complexity over the network, etc. Hence, the applicability of the proposed mechanism should be analyzed with proper metrics that quantify the benefits and the costs associated with the proposed mechanism and the traditional SCADA-based centralized implementation.
- After early anomaly/misbehavior detection, identification of anomalies (types and causes) and the development of the corresponding mitigation/recovery strategies are needed to achieve a completely autonomous network operation.
- Secure resilient wide-area distributed dynamic state estimation needs to be studied for the nonlinear AC power system model with the extensive use of Phasor Measurement Units, as the future grid operation is expected to be more complex, dynamic, and uncertain with the broad integration of renewable and distributed energy resources [69], [70].

Appendix

A. Proof of Lemma 1

In the Kalman filter, given the measurements up to time $t-1$, we have

$$\mathbf{x}_t^j \sim \mathcal{N}(\hat{\mathbf{x}}_{t|t-1}^j, \mathbf{P}_{t|t-1}^j),$$

that implies

$$\mathbf{x}_t^j - \hat{\mathbf{x}}_{t|t-1}^j \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{t|t-1}^j). \tag{32}$$

Hence, similarly to (32), we can write the following:

$$\mathbf{x}_t^{\bar{\ell},j} - \hat{\mathbf{x}}_t^{\bar{\ell},j} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{t|t-1}^{\bar{\ell},j}). \tag{33}$$

Notice that (33) statistically characterizes the estimation error term, $\mathbf{x}_t^{\bar{\ell},j} - \hat{\mathbf{x}}_t^{\bar{\ell},j}$. Based on (8) and (33), we then have

$$\tilde{\mathbf{w}}_t^{\ell,j} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Delta}_t^{\ell,j}),$$

where

$$\mathbf{\Delta}_t^{\ell,j} \triangleq \mathbf{H}^{\bar{\ell},j}\mathbf{P}_{t|t-1}^{\bar{\ell},j}\mathbf{H}^{\bar{\ell},j^{\mathrm{T}}} + \sigma_w^2\,\mathbf{I}_{K^{\ell,j}}.$$

□

B. Proof of Corollary 1

Since χ_t^ℓ is a continuous random variable, its cdf and hence the tail probability, p_t^ℓ, are uniformly distributed $\mathcal{U}[0,1]$ [71]. Then, for the rest of the proof, we refer the reader to [63, Theorem 2]. The only difference in the proofs is that in [63], the tail probability is uniformly distributed in an asymptotic regime, while it is exactly uniform here. □

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

References

[1] G. Liang, J. Zhao, F. Luo, S. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2016.

[2] Y. Lopes, N. C. Fernandes, T. B. de Castro, V. dos Santos Farias, J. D. Noce, J. P. Marques, and D. C. Muchaluat-Saade, "Vulnerabilities and threats in smart grid communication networks," in *Security Solutions and Applied Cryptography in Smart Grid Communications*. IGI Global, 2017, pp. 1–28.

[3] D. Grochocki, J. H. Huh, R. Berthier, R. Bobba, W. H. Sanders, A. A. Cárdenas, and J. G. Jetcheva, "AMI threats, intrusion detection requirements and deployment recommendations," in *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2012, pp. 395–400.

[4] M. Mylrea and S. N. G. Gourisetti, "Blockchain: A path to grid modernization and cyber resiliency," in *2017 North American Power Symposium (NAPS)*, Sept. 2017, pp. 1–5.

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

[6] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proceedings of the 16th ACM Conference on Computer and Communications Security*, ser. CCS '09. New York, NY, USA: ACM, 2009, pp. 21–32.

[7] M. N. Kurt, Y. Yilmaz, and X. Wang, "Real-time detection of hybrid and stealthy cyber-attacks in smart grid," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 2, pp. 498–513, Feb 2019.

[8] M. N. Kurt, O. Ogundijo, C. Li, and X. Wang, "Online cyber-attack detection in smart grid: A reinforcement learning approach," *IEEE Transactions on Smart Grid*, pp. 1–1, 2018.

[9] Y. Yilmaz and S. Uludag, "Mitigating IoT-based cyberattacks on the smart grid," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Dec 2017, pp. 517–522.

[10] S. Soltan, P. Mittal, and H. V. Poor, "BlackIoT: IoT botnet of high wattage devices can disrupt the power grid," in *27th USENIX Security Symposium, 2018*.

[11] E. Bertino and N. Islam, "Botnets and Internet of Things security," *Computer*, vol. 50, no. 2, pp. 76–79, 2017.

[12] A. Ashok, M. Govindarasu, and J. Wang, "Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1389–1407, July 2017.

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

[14] A. S. Debs and R. E. Larson, "A dynamic estimator for tracking the state of a power system," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 7, pp. 1670–1678, Sept. 1970.

[15] N. R. Shivakumar and A. Jain, "A review of power system dynamic state estimation techniques," in *2008 Joint International Conference on Power System Technology and IEEE Power India Conference*, Oct 2008, pp. 1–6.

[16] Q. Yang, L. Chang, and W. Yu, "On false data injection attacks against kalman filtering in power system dynamic state estimation," *Security and Communication Networks*, vol. 9, no. 9, pp. 833–849, 2016.

[17] A. L. Da Silva, M. Do Coutto Filho, and J. De Queiroz, "State forecasting in electric power systems," in *IEE Proceedings C (Generation, Transmission and Distribution)*, vol. 130, no. 5. IET, 1983, pp. 237–244.

[18] Z. Huang, K. Schneider, and J. Nieplocha, "Feasibility studies of applying Kalman filter techniques to power system dynamic state estimation," in *Power Engineering Conference, 2007. IPEC 2007. International*. IEEE, 2007, pp. 376–382.

[19] M. A. Gandhi and L. Mili, "Robust kalman filter based on a generalized maximum-likelihood-type estimator," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2509–2520, May 2010.

[20] L. Chang and K. Li, "Unified form for the robust gaussian information filtering based on m-estimate," *IEEE Signal Processing Letters*, vol. 24, no. 4, pp. 412–416, April 2017.

[21] Y. Zhang, Y. C. Soh, and W. Chen, "Robust information filter for decentralized estimation," *Automatica*, vol. 41, no. 12, pp. 2141–2146, 2005.

[22] J. Zhao, M. Netto, and L. Mili, "A robust iterated extended kalman filter for power system dynamic state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3205–3216, 2017.

[23] Y. Huang, Y. Zhang, N. Li, Z. Wu, and J. A. Chambers, "A novel robust student's t-based kalman filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 3, pp. 1545–1554, June 2017.

[24] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "Approximate inference in state-space models with heavy-tailed noise," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5024–5037, Oct 2012.

[25] H. Wang, H. Li, W. Zhang, and H. Wang, "Laplace l1 robust kalman filter based on majorization minimization," in *2017 20th International Conference on Information Fusion (Fusion)*, July 2017, pp. 1–5.

[26] Y. H. Chang, Q. Hu, and C. J. Tomlin, "Secure estimation based kalman filter for cyber–physical systems against sensor attacks," *Automatica*, vol. 95, pp. 399–412, 2018.

[27] Y. Mo and E. Garone, "Secure dynamic state estimation via local estimators," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec 2016, pp. 5073–5078.

[28] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1454–1467, June 2014.

[29] M. C. Graham, J. P. How, and D. E. Gustafson, "Robust state estimation with sparse outliers," *Journal of Guidance, Control, and Dynamics*, vol. 38, no. 7, pp. 1229–1240, 2015.

[30] Y. Shoukry, P. Nuzzo, A. Puggelli, A. L. Sangiovanni-Vincentelli, S. A. Seshia, and P. Tabuada, "Secure state estimation for cyber-physical systems under sensor attacks: A satisfiability modulo theory approach," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 4917–4932, Oct 2017.

[31] H. Wang, H. Li, J. Fang, and H. Wang, "Robust gaussian kalman filter with outlier detection," *IEEE Signal Processing Letters*, vol. 25, no. 8, pp. 1236–1240, Aug 2018.

[32] A. M. L. da Silva, M. B. D. C. Filho, and J. M. C. Cantera, "An efficient dynamic state estimation algorithm including bad data processing," *IEEE Transactions on Power Systems*, vol. 2, no. 4, pp. 1050–1058, Nov 1987.

[33] H.-Q. Mu and K.-V. Yuen, "Novel outlier-resistant extended kalman filter for robust online structural identification," *Journal of Engineering Mechanics*, vol. 141, no. 1, p. 04014100, 2014.

[34] J. Qi, A. F. Taha, and J. Wang, "Comparing kalman filters and observers for power system dynamic state estimation with model uncertainty and malicious cyber attacks," *IEEE Access*, vol. 6, pp. 77 155–77 168, 2018.

[35] G. Valverde and V. Terzija, "Unscented kalman filter for power system dynamic state estimation," *IET Generation, Transmission Distribution*, vol. 5, no. 1, pp. 29–37, Jan 2011.

[36] M. S. Mahmoud and H. M. Khalid, "Distributed kalman filtering: a bibliographic review," *IET Control Theory Applications*, vol. 7, no. 4, pp. 483–501, March 2013.

[37] A. G. O. Mutambara, *Decentralized Estimation and Control for Multisensor Systems*, 1st ed. Boca Raton, FL, USA: CRC Press, Inc., 1998.

[38] Y. Yilmaz, M. N. Kurt, and X. Wang, "Distributed dynamic state estimation and lqg control in resource-constrained networks," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 3, pp. 599–612, Sep. 2018.

[39] R. Olfati-Saber, "Distributed kalman filtering for sensor networks," in *2007 46th IEEE Conference on Decision and Control*, Dec 2007, pp. 5492–5498.

[40] M. Rostami and S. Lotfifard, "Distributed dynamic state estimation of power systems," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3395–3404, Aug 2018.

[41] D. Ding, Q. Han, Z. Wang, and X. Ge, "A survey on model-based distributed control and filtering for industrial cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2483–2499, May 2019.

[42] D. Ding, Z. Wang, Q. Han, and G. Wei, "Security control for discrete-time stochastic nonlinear systems subject to deception attacks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 5, pp. 779–789, May 2018.

[43] D. Puthal, N. Malik, S. P. Mohanty, E. Kougianos, and C. Yang, "The blockchain as a decentralized security framework [future directions]," *IEEE Consumer Electronics Magazine*, vol. 7, no. 2, pp. 18–21, March 2018.

[44] G.-T. Nguyen and K. Kim, "A survey about consensus algorithms used in blockchain." *Journal of Information Processing Systems*, vol. 14, no. 1, 2018.

[45] T. T. A. Dinh, R. Liu, M. Zhang, G. Chen, B. C. Ooi, and J. Wang, "Untangling blockchain: A data processing view of blockchain systems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 7, pp. 1366–1385, July 2018.

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2019.2928207, IEEE Transactions on Information Forensics and Security

[46] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008, [Online]. Available: http://bitcoin.org/bitcoin.pdf, Accessed: 2018-09-18.

[47] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum Project Yellow Paper*, 2014, [Online]. Available: http://ethereum.github.io/yellowpaper/paper.pdf.

[48] D. Tapscott and A. Tapscott, *Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business, and the World*. Penguin Publishing Group, 2016.

[49] K. Kotobi and S. G. Bilen, "Secure blockchains for dynamic spectrum access: A decentralized database in moving cognitive radio networks enhances security and user access," *IEEE Vehicular Technology Magazine*, vol. 13, no. 1, pp. 32–39, March 2018.

[50] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain-based decentralized trust management in vehicular networks," *IEEE Internet of Things Journal*, 2018, Early Access.

[51] N. Z. Aitzhan and D. Svetinovic, "Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 5, pp. 840–852, Sep. 2018.

[52] F. Lombardi, L. Aniello, S. D. Angelis, A. Margheri, and V. Sassone, "A blockchain-based infrastructure for reliable and cost-effective IoT-aided smart grids," in *Living in the Internet of Things: Cybersecurity of the IoT - 2018*, March 2018, pp. 1–6.

[53] "Exergy: Building a robust value mechanism to facilitate transactive energy," [Online]. Available: http://exergy.energy/wp-content/uploads/2017/12/Exergy-Whitepaper-v8.pdf, accessed: 2018-09-18.

[54] G. Liang, S. R. Weller, F. Luo, J. Zhao, and Z. Y. Dong, "Distributed blockchain-based data protection framework for modern power systems against cyber attacks," *IEEE Transactions on Smart Grid*, 2018, Early Access.

[55] D. Zhaoyang, L. Fengji, and G. Liang, "Blockchain: a secure, decentralized, trusted cyber infrastructure solution for future energy systems," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 5, pp. 958–967, Sept. 2018.

[56] K. Nishiya, J. Hasegawa, and T. Koike, "Dynamic state estimation including anomaly detection and identification for power systems," *IEE Proceedings C - Generation, Transmission and Distribution*, vol. 129, no. 5, pp. 192–198, Sep. 1982.

[57] A. Abur and A. Gomez-Exposito, *Power System State Estimation: Theory and Implementation*, 01 2004, vol. 24.

[58] M. Pilkington, "11 blockchain technology: principles and applications," *Research handbook on digital transformations*, p. 225, 2016.

[59] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han, "When intrusion detection meets blockchain technology: A review," *IEEE Access*, vol. 6, pp. 10 179–10 188, 2018.

[60] "University of Sydney's super-fast blockchain gets even faster," [Online]. Available: http://sydney.edu.au/news-opinion/news/2017/10/25/university-of-sydneys-super-fast-blockchain-gets-even-faster.html, accessed: 2019-06-07.

[61] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[62] M. Khosravi, M. Banejad, and H. T. Shandiz, "Forecast aided measurements data synchronisation in robust power system state estimation," *IET Generation, Transmission Distribution*, vol. 10, no. 10, pp. 2379–2388, 2016.

[63] M. N. Kurt, Y. Yilmaz, and X. Wang, "Real-Time Nonparametric Anomaly Detection in High-Dimensional Settings," *ArXiv e-prints arXiv:1809.05250*, Sep. 2018.

[64] Y. Yilmaz, "Online nonparametric anomaly detection based on geometric entropy minimization," in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 3010–3014.

[65] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[66] Y. Guo, C. Ten, S. Hu, and W. W. Weaver, "Preventive maintenance for advanced metering infrastructure against malware propagation," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1314–1328, May 2016.

[67] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb 2011.

[68] I. Arasaratnam and S. Haykin, "Cubature kalman filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, June 2009.

[69] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. M. Begovic, and A. Phadke, "Wide-area monitoring, protection, and control of future electric power networks," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 80–93, Jan 2011.

[70] J. Qi, A. Hahn, X. Lu, J. Wang, and C.-C. Liu, "Cybersecurity for distributed energy resources and smart inverters," *IET Cyber-Physical Systems: Theory & Applications*, vol. 1, no. 1, pp. 28–39, 2016.

[71] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

Mehmet Necip Kurt (S'19) received the B.S. and the M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2014 and 2016, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with Columbia University, New York, NY, USA. His current research interests include sequential analysis, statistical signal processing, and machine learning with applications to cybersecurity and cyber-physical systems.

Yasin Yılmaz (S'11-M'14) received the Ph.D. degree in electrical engineering from Columbia University, New York, NY, USA in 2014. He is currently an Assistant Professor of electrical engineering with the University of South Florida, Tampa, FL, USA. His research interests include statistical signal processing, machine learning, and their applications to cybersecurity, cyber-physical systems, IoT networks, communication systems, energy systems, transportation systems, social systems, and environmental systems. He received the Collaborative Research Award from Columbia University in 2015, and Research Initiation Award from the Southeastern Center for Electrical Engineering Education in 2017.

Xiaodong Wang (S'98-M'98-SM'04-F'08) received the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA in 1998. He is currently a Professor of electrical engineering at Columbia University, New York, NY, USA. Dr. Wang's research interests fall in the general areas of computing, signal processing and communications, and has published extensively in these areas. Among his publications is a book entitled "Wireless Communication Systems: Advanced Techniques for Signal Reception", published by Prentice Hall in 2003. His current research interests include wireless communications, statistical signal processing, and genomic signal processing. Dr. Wang received the 1999 NSF CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He has served as an Associate Editor for the *IEEE Transactions on Communications*, the *IEEE Transactions on Wireless Communications*, the *IEEE Transactions on Signal Processing*, and the *IEEE Transactions on Information Theory*. He is a Fellow of the IEEE and listed as an ISI Highly-cited Author.

1556-6013 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.