IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 4 July 2022, accepted 20 July 2022, date of publication 25 July 2022, date of current version 29 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3193781

RESEARCH ARTICLE

False Data Detection in a Clustered Smart Grid Using Unscented Kalman Filter

MUHAMMAD RASHED[1], JOARDER KAMRUZZAMAN[1], IQBAL GONDAL[2], AND SYED ISLAM[1], (Fellow, IEEE)
[1]Internet Commerce Security Laboratory, School of Engineering, Federation University, Mount Helen, VIC 3350, Australia
[2]SCT, STEM College, RMIT University, Melbourne, VIC 3001, Australia

Corresponding author: Muhammad Rashed (muhammadrashed@students.federation.edu.au)

ABSTRACT The smart grid accessibility over the Internet of Things (IoT) is becoming attractive to electrical grid operators as it brings considerable operational and cost efficiencies. However, this in return creates significant cyber security challenges, such as fortification of state estimation data such as state variables against false data injection attacks (FDIAs). In this paper, a clustered partitioning state estimation (CPSE) technique is proposed to detect FDIA by using static state estimation, namely, weighted least square (WLS) method in conjunction with dynamic state estimation using minimum variance unscented Kalman filter (MV-UKF) which improves the accuracy of state estimation. The estimates acquired from the MV-UKF do not deviate like WLS as these are purely based on the previous iteration saved in the transition matrix. The deviation between the corresponding estimations of WLS and MV-UKF are utilised to partition the smart grid into smaller sub-systems to detect FDIA and then identify its location. To validate the proposed detection technique, FIDAs are injected into IEEE 14-bus, IEEE 30-bus, IEEE 118-bus, and IEEE 300-bus distribution feeder using MATPOWER simulation platform. Our results clearly demonstrate that the proposed technique can locate the attack area efficiently compared to other techniques such as chi square.

INDEX TERMS Smart grid, unscented Kalman filter, state estimation, FDIA.

I. INTRODUCTION

IoT refers to a network of smart devices where each device is assigned a unique IP address that helps with its identification and connectivity over the global network [1]. A smart grid is an intricate intelligent network of power lines and equipment connecting buses, nodes, generators, control center, and meters [2]. It provides electricity and power to the consumers using smart techniques, ranging from a single user to critical businesses such as defense facilities, hospitals, and airports [3]. Within an IoT based smart grid, cyber-attack to a certain part of the grid means the intruder will not limit itself to one bus or a phasor measurement unit (PMU) but will aim to take control of the complete infrastructure and drive the smart grid towards shutdown [4].

The fact that smart grid interconnects generation resources such as renewable, thermal, hydro, and solar, a cyber-attack on IoT based smart grids may risk all the facilities and thus could result in cascaded blackouts, overloading and service disruptions to its consumers and critical infrastructures [5]. Therefore, the protection of smart grid and IoT systems from cyber-attacks and FDIAs is critical and should be addressed with utmost urgency [6]. FDIAs are the most studied cyber-physical attacks in smart grid security. One of the key targets of FDIA is state estimation data [7].

State estimations collect measurements through the use of sensors and metering devices for voltage magnitudes, line flows, and power to monitor the operational status of the grid [8]. State estimation also helps to indicate the presence of false data within these measurements [9]. In traditional state estimation such as WLS, the measurement data is collected from various buses to estimate the state of the grid. A hypothesis test known as chi-square test or residual test is then applied to collected measurements where a measured value is tested against a threshold value to determine the presence of false data. However, since cyber-attacks have become more

The associate editor coordinating the review of this manuscript and approving it for publication was Bin Zhou.

intricate in recent times, WLS alone fail to detect FDIA, especially, when the attackers possess the knowledge of network configuration [10]. Furthermore, the chi-square tests are no longer effective when applied to a complete gird due to large number of redundant measurements [11].

One drawback of WLS is that it only considers one set of measurements. Although it provides satisfactory results [12], but detection accuracy needs improvement in countering cyber-attacks like FDIAs. Static state estimation cannot capture the time history of the state estimates; therefore, it cannot predict the future estimates of the system [13]. The other problem with static WLS estimator is that it must be reinitialised for every new set of measurements without using any state prediction from previous estimations and this increases the computational complexity [14]. For this reason, dynamic state estimation such as Kalman filter is a better alternative [15].

Additionally, the MV-UKF improves the estimation results compared to ordinary UKF when used in combination with WLS. MV-UKF is mainly intended for nonlinear systems and has the potential to tackle the problem of FDIA detection since WLS alone has failed to do so [16]. Therefore, we investigated the detection of FDIA by measuring the deviation of corresponding estimations using the MV-UKF and WLS based state estimations. We found that the state variables under attack deviated significantly between the MV-UKF and WLS based state estimations.

In this research, we propose a method to enhance the reliability of the existing WLS estimates to ensure secure operations by estimating the state variables such as voltage magnitudes of the system accurately even in the presence of a cyber-attack. The proposed algorithm utilises the measurements collected from the sensors, and metering devices to accurately estimate the state of the system. The combination of state prediction, based on linearization of the power flow equations, WLS and Kalman filter formulation improves the results of FDIA detection. Note that our use of MV-UKF makes the detection system computationally efficient.

Overall, in this work, we exploit the deviation of state variables estimation by WLS and MV-UKF to detect FDIA using CPSE. Our contributions are as follows.

- A false data detection technique based on MV-UKF is proposed which has not been used previously in conjunction with WLS state estimation;
- An attack magnitude that can deceive ordinary Kalman filter is tested on MV-UKF and has shown promising results;
- Our technique partitions the smart grid into smaller subsystems when necessary to identify attack location, and thereby reduces computational complexity as a smaller number of measurements are taken into account for WLS calculation;
- Scalability study of the proposed technique with small to large size IEEE bus systems shows promising results and better detection performance than competing existing work.

The remainder of the paper is organized as follows. The Related work is in Section II. The State Estimation and FDIA is explained in Section III. The Proposed Attack Detection Methodology is explained in Section IV. The Case studies & results are discussed in Section V. The paper concludes in Section VI.

II. RELATED WORKS

In [17], a robust massively parallel dynamic (RMPD) state estimation approach utilizing extended Kalman filter is proposed using graphics processing units (GPU) that can detect FDIA using a trusted set of measurements from optimized PMUs. Further, a Markov model was proposed considering the stochastic nature of the power system and the historic measurements of the system's dynamic behavior to improve the accuracy of the estimation results using the Euclidean distance metric. The detection accuracy of this work can be improved by increasing the number of GPU cores and processing power. However, further analysis of this work is required to identify multiple attacks as this will lead to computational complexity.

In [18], Ganjkhani *et al.* presented a nonlinear autoregressive exogenous (NARX), a specific configuration of an artificial neural network (ANN), based bad data detection processor to identify the FDIAs on static state estimation. The estimates were predicted using NARX network and compared with the computed state variables to identify the FDIA. The NARX network demonstrated high accuracy for the detection of the FDIAs on state estimation. However, the presented method was only tested on DC state estimation and needs to be extended for AC state estimation and tested by adding reactive power measurements.

Chen *et al.* in [19] proposed an online detection method of data injection attacks against dynamic state estimation in a smart grid by solving an optimal model using particle swarm optimization (PSO). The system's performance was tested by developing a data injection attack strategy with minimum attack residual increment. Based on the test results, an online chi-square detection method associated with two kinds of state estimates was proposed to make up for the system vulnerability. However, it is difficult to analyze the attack residual increment in multiple cyber-attacks which could lead to a much more complex residual model in a dynamic power system, which makes it difficult to detect the attack vectors directly.

In [20], a GPU enabled adaptive robust state estimator was proposed comprising of deep learning algorithm, long short-term memory, and a non-linear extended Kalman filter to deal with the massive connections and states generated by the state estimation data. It provides an online parametric state estimate based on software defined IoT controller. Two levels of online parametric state estimation were used to improve the reliability and security of the communication. However, the implementation of the proposed methodology requires 6G enabled smart grids to achieve a minimum latency for

countering the transients to minimize the chances of cascading failures within both static and dynamic state estimation.

In [21], Pei *et al.* suggest that FDIAs can be detected by a set of strategically selected measurements. The authors proposed a deviation-based detection method (DBDM) based on an additional Kalman filter estimator for dynamic state estimation with the historical states transition. The FDIAs were detected by using an exponential weighting function to enhance the robustness of the Kalman filter against attacks, however, the type of the Kalman filter was not mentioned. Since there exists a continuous variation of load and generation, methods that can capture those variations such as time variant state transition methods needs to be incorporated in this work to improve the detection accuracy.

In [22], Zhao and Mili proposed to handle non-gaussian noise and outliers with the use of generalized maximum likelihood unscented Kalman filter which allows the sigma points to reliably approximate the mean and covariance matrices of the predicted and corrected state vectors. Numerical results were collected using the IEEE 39-bus 10-machine system which demonstrated the effectiveness of the proposed method.

In [23], Junbo *et al.* proposed a generalized maximum likelihood iterated extended Kalman filter for tracking the dynamic states of a power system. Simulations were carried out on the IEEE-39-bus test system that demonstrated the statistical efficiency of the proposed method. The authors also highlighted the vulnerability to system parameters, topology errors and unreliable estimates under strong nonlinearities of the power system model.

III. STATE ESTIMATION AND FDIA

The state estimation uses meter measurements to formulate state variables that can be expressed by a nonlinear model [24]–[27] as

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

where z is the meter measurement, k is the iteration index, $h(x)$ is the function of state variable x that constructs a Jacobian matrix that depends on the impedance of the network topology, and e is the measurement noise that follows a Gaussian distribution of zero mean.

Attackers inject the FDIA vector by manipulating the measurements of metering devices at any bus. The system measurement then becomes:

$$z_a(k) = Hx + a + e \tag{2}$$

$$a = Hc \tag{3}$$

$$a = [a_1, a_2, \ldots, a_m]^T \tag{4}$$

where $a = [a_1, a_2, \ldots, a_m]^T$ denotes the attack vector at each bus, H is the Jacobian matrix and it depends on the topology structure of power grid, and $c=[c_1,c_2,\ldots,c_n]^T$ is an arbitrary vector.

The meter measurements can be rewritten as

$$z_a(k) = Hx + Hc + e \tag{5}$$

$$z_a(k) = H(x + c) + e \tag{6}$$

When the next measurement $z(k+1)$ is injected with false vector such as $a(k+1)$, the measurement $z(k+1)$ and estimate $\hat{z}(k+1)$ becomes false measurement $z_a(k+1)$ and false estimates $\hat{z}_a(k+1)$.

The difference for data injection attacks against dynamic state estimation can be written by the norm of measurement residuals as:

$$\left\| z_a(k+1) - \hat{z}_a(k+1) \right\| = 0 \tag{7}$$

To successfully launch FDIA against state estimation, the above conditions must be met.

This attack is undetectable in the chi-square detector as the manipulated state $(x+c)$ (Eq. (6)) is treated as the real value in the state estimator. The hypothesis test such as chi-square test identifies bad data only if the absolute value of the residual exceeds a certain threshold value [11]. However, in this case, the residue test fails to detect the carefully designed FDIAs because this will not affect the residue.

Therefore, the partitioning of the grid into smaller subsystems is necessary and effective for the successful detection of FDIA as it reduces the number of redundant measurements, hence making the chi-square tests more effective.

The below objective function $J\left(\hat{x}\right)$ is first computed to solve the WLS estimation problem:

$$J\left(\hat{x}\right) = \sum\nolimits_{i=1}^{m} \frac{(z_i - h_i(x))^2}{\sigma_i^2} \tag{8}$$

where m is the number of measurements, σ_i^2 is the standard deviation, and z_i is the i-th measurement from a meter.

The next step is the use of a hypothesis test within the state estimator known as chi-square test on normalized residual formed using collected state measurements. The chi-square test will use the following two conditions.

$$\begin{cases} H_0 : J\left(\hat{x}\right) \geq \chi^2_{(m-n),p} & \textit{bad data} \\ H_0 : J\left(\hat{x}\right) \leq \chi^2_{(m-n),p} & \textit{no bad data} \end{cases} \tag{9}$$

where $J\left(\hat{x}\right)$ is the normalised-sum square residual that follows $\chi^2_{(m-n)}$ distribution pattern based on the load profile, $\hat{x}$ is an estimate of x solved by WLS algorithm, $\chi^2_{(m-n),p}$ is the detection threshold corresponding to p, where p is the detection confidence which is taken as 95%, and n is the number of state variables.

The objective function $J(\hat{x}) \geq \chi^2_{(m-n),p}$ is tested against the value in a distribution table. If the resulting value is greater, then false data is detected; otherwise, if the value is below the threshold, then it is assumed to be free of false data.

IV. PROPOSED ATTACK DETECTION METHODOLOGY

As shown in Fig. 1, our algorithm comprises (i) a traditional WLS estimator, (ii) an MV-UKF dynamic estimator, and (iii) a graph establishment based on state variables and further partitioning of the grid using k-means clustering.

MV-UKF uses statistical linearization known as unscented transformation. In this method, the probability distribution

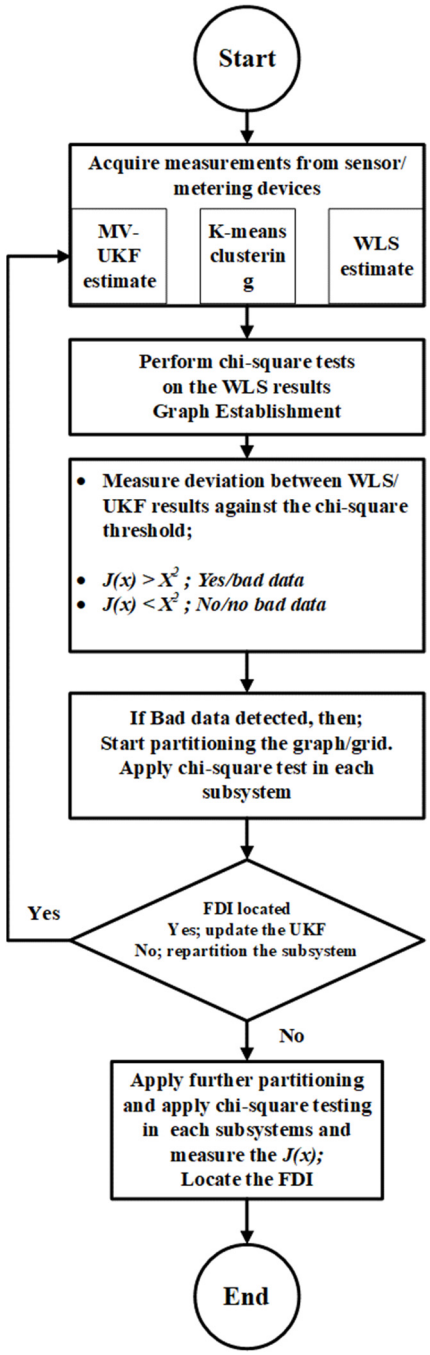


FIGURE 1. Proposed deviation based FDIA detection flow chart.

is approximated using a set of deterministic chosen points called sigma points [24], [28]–[33]. The MV-UKF achieves better performance when used in combination with distributed state estimation. As shown in Fig. 2, the main purpose of unscented transformation is to choose multiple sigma points with weights that are used to create a mean and covariance matrix of the estimates.

In MV-UKF, no calculation of Jacobian matrices is required for every new iteration which saves computational

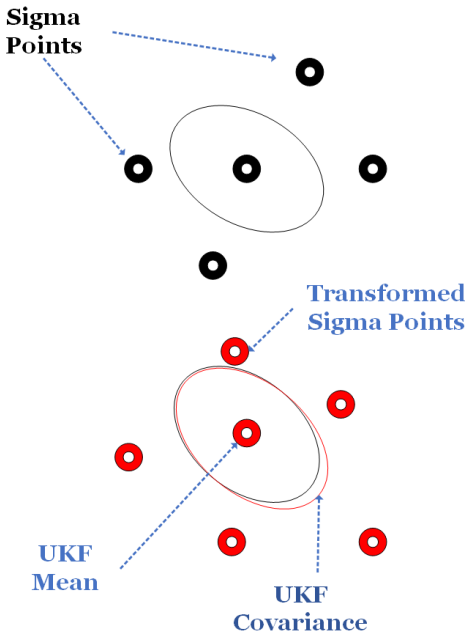


FIGURE 2. Mean and covariance propagation using MV-UKF.

complexity, reduces time, and makes it superior to other Kalman filters.

Unlike Kalman filter which uses a constant gain based on a single measurement, UKF in [28]–[33] chooses multiple sigma points around a mean to compute gain as shown in Fig. 2, hence making it superior to the ordinary Kalman filter. This technique can track errors more accurately and the false data detection becomes more precise. Once the smart grid is subjected to an attack, the false data within the measurements gradually becomes significant over time and an attacker can take advantage of this by keeping injecting false data to make the previous data look real to the ordinary Kalman filter. However, this is not the case with MV-UKF due to its multiple sigma measurements. Secondly, the unbiased minimum-variance state estimation is derived by minimizing the trace of the state error in the covariance matrix [28].

A nonlinear system can be described as:

$$x_k = f(x_{k-1}, u_k) + G_k d_k + w_k \tag{10}$$

$$z_k = h(x_k, u_k) + v_k \tag{11}$$

where $x_k \in \mathbb{R}$ is the state vector, $z_k \in \mathbb{R}$ is the measurement at time instant k, f and h are vector valued nonlinear functions that depend on network configuration and number of buses, G_k is a known matrix, d_k is the unknown input, u_k is the known input, w_k is the process noise, and v_k is the measurement noise.

For $2n$ sigma point weights in,
$\omega_i = 1/(2n), i = 1, \ldots, 2n,$
and the state vector $\chi^i_{k-1|k-1}$ can be expressed as:

$$\chi^i_{k-1|k-1} = \hat{x}_{k-1|k-1} \pm (\sqrt{n}\mathrm{P}_{k-1|k-1})_i \tag{12}$$

where $\hat{x}_{k-1|k-1}$ is the estimate at time instant k-1 and $P_{k-1|k-1}$ is its covariance matrix.

The predicted state $\hat{x}_{k|k-1}$ and its covariance matrix $P_{k-1|k-1}$ are computed as follows:

$$\hat{x}_{k|k-1} = \sum_{i=1}^{2n} \omega_i f(\chi^i_{k-1|k-1}, u_k) \tag{13}$$

$$P_{k-1|k-1} = \sum_{i=1}^{2n} \omega_i (\chi^i_{k|k-1} - \hat{x}_{k|k-1})(\chi^i_{k|k-1} - \hat{x}_{k|k-1})^T + Q_k \tag{14}$$

where Q is the covariance of the process noise.

$$\chi^i_{k|k-1} = f(\chi^i_{k-1|k-1}, u_k) \tag{15}$$

The new sigma points for the predicated estimate and its covariance matrix are generated as follows:

$$\chi^i_{k|k-1} = \hat{x}_{k|k-1} \pm (\sqrt{n P_{k|k-1}})_i \tag{16}$$

The predicted measurement vector $\hat{z}_{k|k-1}$ and its covariance matrix $P^{zz}_{k|k-1}$ are calculated as follows:

$$\hat{z}_{k|k-1} = \sum_{i=1}^{2n} \omega_i h(\chi^i_{k|k-1}, u_k) \tag{17}$$

$$P^{zz}_{k|k-1} = \sum_{i=1}^{2n} \omega_i (Z^i_{k|k-1} - \hat{z}_{k|k-1})(Z^i_{k|k-1} - \hat{z}_{k|k-1})^T + R_k \tag{18}$$

$$Z^i_{k|k-1} = h(\chi^i_{k|k-1}, u_k) \tag{19}$$

where R_k is the covariance matrix for measurement noise.

The nonlinear measurement function is the given by:

$$z_k = H_k(x_k - \hat{x}_{k|k-1}) + \hat{z}_{k|k-1} + \varepsilon_k \tag{20}$$

where ε_k is the error vector due to statistical linearization.

For MV-UKF, the state vector can be estimated using the following equations:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(z_k - \hat{z}_{k|k-1}) \tag{21}$$

$$P_{k|k} = P_{k|k-1} - H_k P_{k|k-1} \tag{22}$$

where K is Kalman gain.

The estimates generated by the MV-UKF in Eq. (21) are compared with the WLS estimates derived from Eq. (9). In the case of an FDIA at a sensor or a metering device, there is a significant deviation observed between the estimates generated by the MV-UKF and the WLS estimates. Once the deviation exceeds the pre-set threshold, our proposed methodology triggers the system graph to partition the smart grid into smaller sub-systems and then the chi-square test is applied within each sub-system. The test becomes more effective within the scope of these sub-systems because the number of redundant measurements is reduced hence reducing both χ^2 and $J(x)$.

In order to develop a pre-set threshold for the detection of false data within the estimates, we performed chi-square test on WLS estimates for several iterations. When the false data is not present, the pre-set threshold smoothed to a stable value which can only be exceeded when false data is present.

While knowing the initial network condition at iteration k-1, an MV-UKF is used to determine the network condition on the basis of the available meter measurements. A transition and covariance matrix are created using forecasted state data for future time instant k.

Algorithm 1: *Bad Data Detection Using CPSE & MV-UKF*

Result: *Initial estimation of estimate x and covariance p*
Initialize MV-UKF to acquire estimates;
while *Collect sensor measurements* **do**
  Develop graph based on collected measurements;
  Compute WLS using eq (8) $J(\hat{x}) = \sum_{i=1}^{m} \frac{(z_i - h_i(x))^2}{\sigma_i^2}$;
  if *deviation between measurements and MV-UKF estimates* $r_k = \frac{|x_k^{WLS} - x_k^{MV-UKF}|}{\sqrt{C_k}}$ *exceeds the pre-set threshold* $J(\hat{x}) \geq \chi^2_{(m-n),p}$; **then**
    Partition Graph G= { |V| → Bus number };
    Perform chi-square tests J(x) on a distributed cluster Subsystems using eq (9);
    Reduce the weights for the MV-UKF and give more weight to collected measurement;
    If weighted residual exceeds the threshold; then
    subdivide the graph and perform the chi-square test on subsystems and locate the FDIA;
    If *Bad data detected* ;
    Update the graph;
  else
    Send the estimates to MV-UKF for status update;
  end
end

Algorithm 1 describes how WLS, MV-UKF and graph partitioning are used in an iterative manner to detect FDIA in our proposed method. When a successful FDIA is launched, the state measurements disturb the estimates of the WLS based algorithm. However, on the contrary, the estimates acquired from the MV-UKF do not deviate as these are purely based on the previous iteration saved in the transition matrix. Therefore, a normalized residual r_k is calculated by measuring an absolute difference between a state variable acquired by the WLS based method x_k^{WLS} and the MV-UKF x_k^{MV-UKF}. The result is divided by the standard deviation $\sqrt{C_k}$; which is obtained from the covariance matrix. The residual r_k is thus calculated as:

$$r_k = \frac{\left| x_k^{WLS} - x_k^{MV-UKF} \right|}{\sqrt{C_k}} \tag{23}$$

The graph is established based on voltage magnitudes as well as bus numbers as follows:

$$G = \{|V| \rightarrow Bus_{number}\} \tag{24}$$

where $|V|$ is voltage magnitude, and Bus_{number} is the bus number from where the measurement is received.

Root mean square error (RMSE) compares a predicted value with an observed or known values. This represents an error between two data sets, hence smaller values illustrate improvement in the estimates.

TABLE 1. State estimates by wls and mv-ukf.

WLS	MV-UKF	FDIA	Graph
105	105	Not detected	No Partition
120	120	Not detected	No Partition
280	130	FDIA detected	Partition
135	135	Not detected	No Partition

TABLE 2. Chi-square test Chi-square test $J(x)$ when calculated over the complete system. Note: the value of $J(x)$ and threshold χ^2 are same as this is calculated over the entire system.

System	Bus	Real (kW)	False (kW)	Threshold χ^2	$J(x)$
Subsystem 1	1	100	105		
Subsystem 2	8	120	125		
Subsystem 3	22	135	140	102.5	
Subsystem 4	15	130	280		79.05

V. CASE STUDIES AND RESULTS

To demonstrate the performance of the proposed method, the simulations were performed using MATPOWER with a computer specification, 3.2 GHz Intel Core i5 processor and 4GB memory with a Window 10 system. The simulations were performed on smart grid IEEE 14-bus, IEEE 30-bus, IEEE 118-bus, & IEEE 300-bus. Below, we first present our detailed experiments with IEEE 30-bus and then show results with other IEEE systems.

As shown in Table 1; column 1 lists all the WLS estimates obtained through the meter measurements along with the dynamic estimates generated by MV-UKF in column 2. The decision to partition the grid is listed in column 4 which is only made once the deviation between the two estimates exceeds a certain threshold and FDIA detected.

In case of no FDIA, the weighted sum squared residual $J(x)$ stays below the threshold value for IEEE-30 bus system as per Eq. (9). The power flow on the transmission lines from bus 15 was modified from 130 kW to 280 kW to launch an FDIA shown in Fig. 3. The chi-square test is applied on all the measurements using MATPOWER. As shown in Table 2, the value of $J(x)$ doesn't exceed the threshold, indicating no FDIA on the system which is not true. The weighted sum-squared residual is calculated as $J(x) = 79.05$ which is lower than the threshold of the IEEE-30 bus system $\chi^2 = 102.5$, therefore no FIDA could be detected.

The incorrect measurement received by the estimator is 280 kW which is not the estimate generated by MV-UKF (130 kW), therefore system triggers the partitioning, as the deviation between the two estimates is high. The algorithm

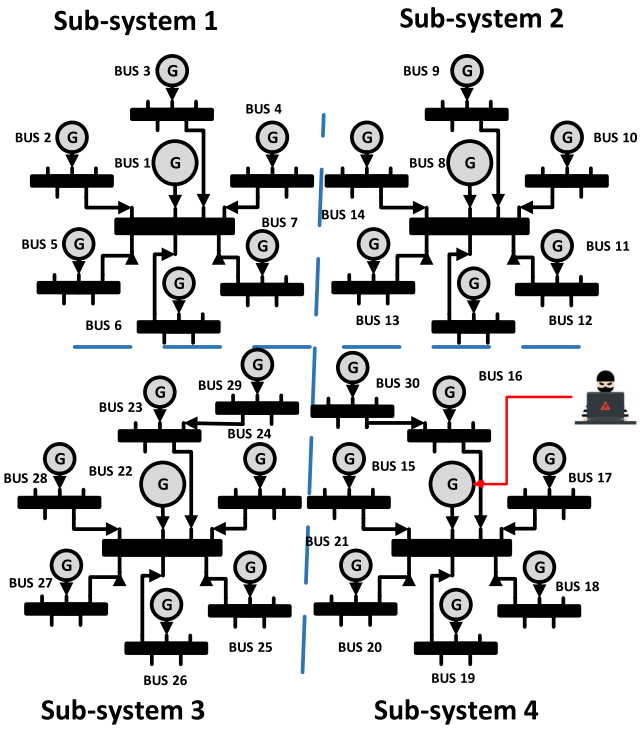


FIGURE 3. Partitioning over IEEE-30 bus system.

TABLE 3. Chi-square test $J(x)$ performed on the partitioned subsystem.

Subsystem	Threshold χ^2	$J(x)$
Subsystem 1	34.16	30
Subsystem 2	34.16	30
Subsystem 3	34.16	30
Subsystem 4	34.16	38.85

TABLE 4. Chi-square test $J(x)$ over partitioned subsystem 4.

Subsystem	Threshold χ^2	$J(x)$
Subsystem 41	15.85	14
Subsystem 42	15.85	14
Subsystem 43	15.85	28.73

triggers the need for partitioning the system into smaller subsystems and develops a graph. The graph is partitioned using a k-means clustering algorithm and initially the system is partitioned into subsystems 1, 2, 3 & 4 as shown in Fig. 3. It performs chi-square test on all the subsystems separately in order to make it more effective.

The number of redundant measurements is reduced with the reduction in the size of subsystems and the threshold drops to $\chi^2 = 34.16$. The weighted sum square residual $J(x)$ is shown in Table 3 calculated as $J(x) = 30$ for subsystem 1, 2 and 3, except for subsystem 4, where the FDIA is launched. In subsystem 4, $J(x) = 38.85$ exceeds the threshold $\chi^2 = 34.16$, hence, implying there is a presence of an FDIA in subsystem 4.

In order to locate the FDIA, the graph of subsystem 4 is further subdivided into subsystem 41, 42 & 43. The chi-square

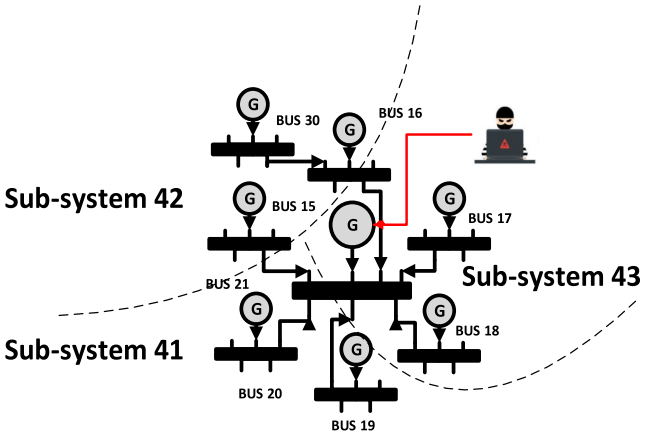


FIGURE 4. Further partitioning of subsystem 4.

test is applied on all these subsystems and the results are shown in Table 4. The threshold drops to $\chi^2 = 15.85$ for each subsystem and $J(x) = 14$ for subsystem 41 & 42. In the case of subsystem 43, where the FDIA is launched, weighted sum square residual is calculated as $J(x) = 28.73$, which exceeds the threshold $\chi^2 = 15.85$, indicating the attack is in subsystem 43.

For a large IEEE bus system, it is computationally complex to locate the attack because of the large number of redundant measurements.

To explore the efficiency further, estimates were obtained using the proposed method in Fig. 5 (a) IEEE 14-bus, (b) IEEE 30-bus, (c) IEEE 118-bus, and (d) IEEE 300-bus. The estimates obtained through the MV-UKF are plotted for RMPD [17] and DBDM [21] estimates where an additional Kalman filter was used. The real values in p.u. indicate that no FDIA has been launched.

FDIAs were injected at the following bus numbers (Bus no. 5, Bus no. 10, …, Bus no. 100) as shown in Fig. 6 (a) IEEE-118 bus and (b)IEEE-300 bus. The estimates obtained using the proposed algorithm are plotted at each bus with DBDM estimates. It can be easily seen that the MV-UKF estimates are improved as compared to the DBDM estimates.

As shown in Fig. 6 (a) IEEE-118 bus and (b) IEEE-300 bus, the voltage estimates obtained using the proposed method are much improved and closer to the real values than those produced by RMPD and DBDM methods. The simulation results prove the effectiveness of the proposed method over the other estimation results.

Our approach relies on generating estimates based on MV-UKF and the detection results which depend on a specified threshold. The threshold computed in Eq. (8) and (9) and its selection should be done carefully as lower and higher than necessary value increases the probability of false detections.

The careful selection of quality threshold can be improved by generating daily load curves for multiple days such as weekday, weekend and public holidays with the time, day and night specification. Adjustment to the threshold value can be done based on the load curves and use of daily load curves to improve the results of MV-UKF. Our results clearly show that combining graph partitioning with MV-UKF can be highly useful in this scenario.

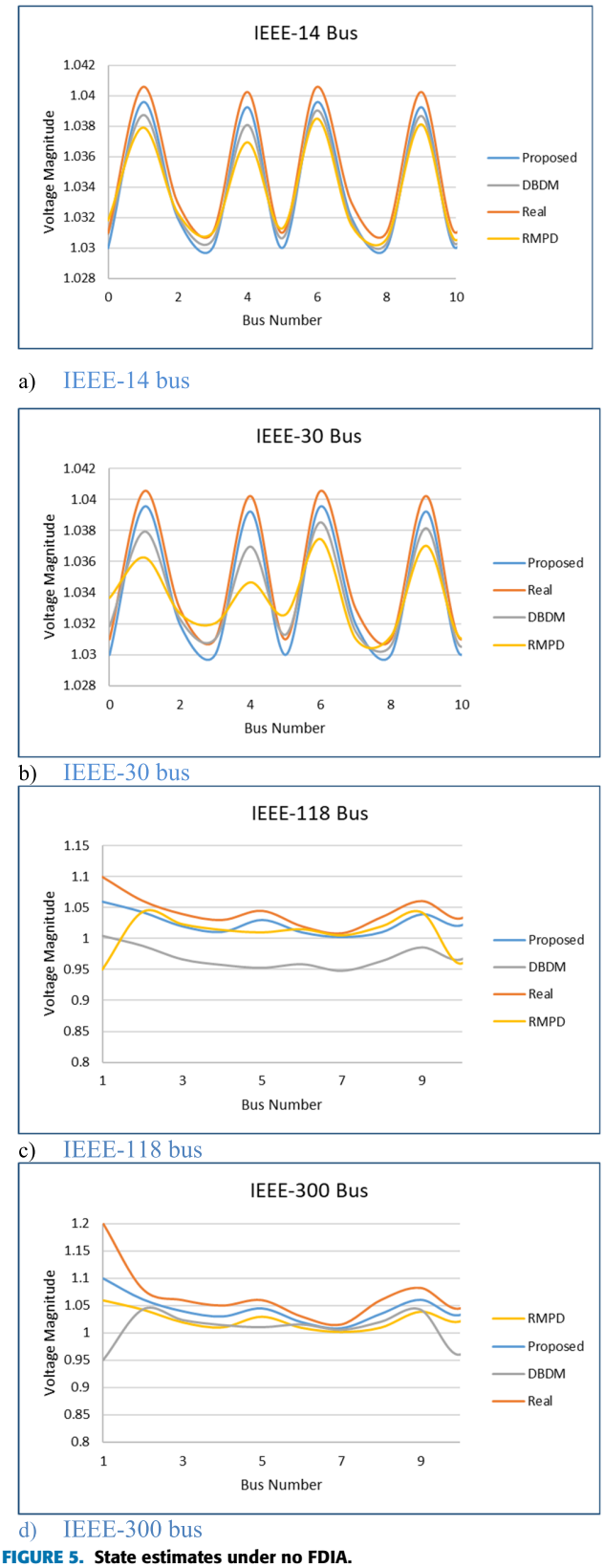
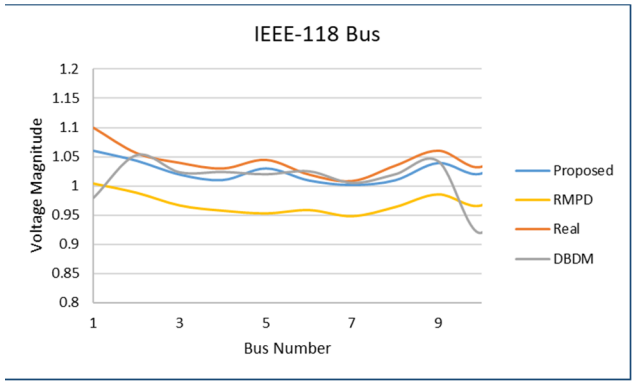
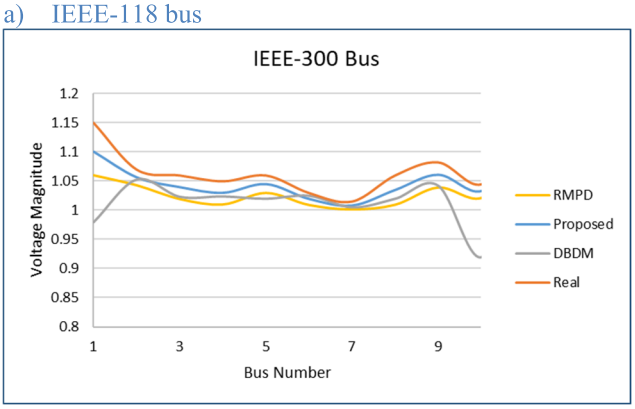


FIGURE 5. State estimates under no FDIA.



a) IEEE-118 bus



b) IEEE-300 bus

FIGURE 6. State estimates under FDIA.

TABLE 5. Comparison of RMSE values for different methods.

Bus Number	RMPD (RMSE)	DBDM (RMSE)	Proposed (RMSE)
1	$2.82\text{x}10^{-2}$	$1.51\ \text{x}10^{-2}$	$3.20\ \text{x}10^{-3}$
2	$2.81\ \text{x}10^{-2}$	$1.45\ \text{x}10^{-2}$	$3.13\ \text{x}10^{-3}$
3	$2.84\ \text{x}10^{-2}$	$1.47\ \text{x}10^{-2}$	$3.19\ \text{x}10^{-3}$
4	$2.79\ \text{x}10^{-2}$	$1.42\ \text{x}10^{-2}$	$2.75\ \text{x}10^{-3}$
5	$2.77\ \text{x}10^{-2}$	$1.38\ \text{x}10^{-2}$	$2.45\ \text{x}10^{-3}$
6	$2.74\ \text{x}10^{-2}$	$1.34\ \text{x}10^{-2}$	$1.95\ \text{x}10^{-3}$
7	$2.71\ \text{x}10^{-2}$	$1.30\ \text{x}10^{-2}$	$1.31\ \text{x}10^{-3}$
8	$2.70\ \text{x}10^{-2}$	$1.30\ \text{x}10^{-2}$	$1.37\ \text{x}10^{-3}$
9	$2.68\ \text{x}10^{-2}$	$1.25\ \text{x}10^{-2}$	$1.78\ \text{x}10^{-4}$
10	$2.67\ \text{x}10^{-2}$	$1.22\ \text{x}10^{-2}$	$1.32\ \text{x}10^{-4}$
11	$2.66\ \text{x}10^{-2}$	$1.19\ \text{x}10^{-2}$	$7.46\ \text{x}10^{-5}$
12	$2.67\ \text{x}10^{-2}$	$1.16\ \text{x}10^{-2}$	$8.34\ \text{x}10^{-5}$
13	$2.69\ \text{x}10^{-2}$	$1.10\ \text{x}10^{-2}$	$5.78\ \text{x}10^{-5}$
14	$2.73\ \text{x}10^{-2}$	$9.66\ \text{x}10^{-3}$	$6.74\ \text{x}10^{-5}$

Root Mean Square Error (RMSE) can be used to calculate the error in an estimation method by taking the differences between the estimated values and the real values. The RMSE values of the proposed method are calculated and compared with those of RMPD and DBDM as shown in Table 5. The proposed method values show the least error in the estimates.

VI. CONCLUSION

In this work, we have proposed a novel detection method for FDIA in a smart grid. Considering that the traditional chi-square detection fails in many cases in detecting the attacks, so we proposed a hybrid technique using WLS and MV-UKF state estimations in conjunction with a graph establishment and subsystem partitioning of the grid. The detection of FDIA improves significantly when estimates from MV-UKF are used to measure the deviation with the results obtained from WLS. A graph is developed when the deviation exceeds preset threshold and the state data from different buses are classified using clustering-based algorithm. Our method is also able to locate the subsystem where FDIA was launched. To showcase the effectiveness of the proposed algorithm, the simulations were conducted on IEEE-14 bus, IEEE-30 bus, IEEE-118 bus, and IEEE-300 bus.

Finally, the MV-UKF can obtain priori results of state estimation based on historical values. Hence, if some meter or sensor measurements are lost, these missing measurements can be replaced by the estimates to improve the reliability of the state estimation. Exploiting the proposed method's capability to use multiple sigma points, future work will focus on extending the method to handle multiple cyber-attacks.

REFERENCES

[1] K. Kimani, V. Oduol, and K. Langat, "Cyber security challenges for IoT-based smart grid networks," *Int. J. Crit. Infrastruct. Protection*, vol. 25, pp. 36–49, Jun. 2019, doi: 10.1016/j.ijcip.2019.01.001.

[2] F. Ahmad, A. Rasool, E. Ozsoy, R. Sekar, A. Sabanovic, and M. Elitaş, "Distribution system state estimation—A step towards smart grid," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 2659–2671, Jan. 2017, doi: 10.1016/j.rser.2017.06.071.

[3] S. Li, Y. Yilmaz, and X. Wang, "Quickest detection of false data injection attack in wide-area smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2725–2735, Nov. 2015, doi: 10.1109/TSG.2014.2374577.

[4] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000, doi: 10.1109/5.824004.

[5] A. Anwar and A. N. Mahmood, "Vulnerabilities of smart grid state estimation against false data injection attack," Tech. Rep., 2014.

[6] Z. Lv, L. Wang, Z. Guan, J. Wu, X. Du, H. Zhao, and M. Guizani, "An optimizing and differentially private clustering algorithm for mixed data in SDN-based smart grid," *IEEE Access*, vol. 7, pp. 45773–45782, 2019, doi: 10.1109/ACCESS.2019.2909048.

[7] N. Ahmadi, Y. Chakhchoukh, and H. Ishii, "Power systems decomposition for robustifying state estimation under cyber attacks," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 1922–1933, May 2021, doi: 10.1109/TPWRS.2020.3026951.

[8] S. Alromaihi, W. Elmedany, and C. Balakrishna, "Cyber security challenges of deploying IoT in smart cities for healthcare applications," in *Proc. 6th Int. Conf. Future Internet Things Cloud Workshops (FiCloudW)*, Aug. 2018, pp. 140–145.

[9] H. H. H. Aly, "A novel approach for harmonic tidal currents constitutions forecasting using hybrid intelligent models based on clustering methodologies," *Renew. Energy*, vol. 147, pp. 1554–1564, Mar. 2020, doi: 10.1016/j.renene.2019.09.107.

[10] H. H. H. Aly, "A proposed intelligent short-term load forecasting hybrid models of ANN, WNN and KF based on clustering techniques for smart grid," *Electr. Power Syst. Res.*, vol. 182, May 2020, Art. no. 106191, doi: 10.1016/j.epsr.2019.106191.

[11] G. Anagnostou and B. C. Pal, "Derivative-free Kalman filtering based approaches to dynamic state estimation for power systems with unknown inputs," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 116–130, Jan. 2018, doi: 10.1109/TPWRS.2017.2663107.

[12] C. Hernandez and P. Maya-Ortiz, "Comparison between WLS and Kalman filter method for power system static state estimation," in *Proc. Int. Symp. Smart Electr. Distrib. Syst. Technol. (EDST)*, Sep. 2015, pp. 47–52, doi: 10.1109/SEDST.2015.7315181.

[13] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Chin, and Z. Fan, "Sparse malicious false data injection attacks and defense mechanisms in smart grids," *IEEE Trans. Ind. Informat.*, vol. 11, no. 5, pp. 1–12, Oct. 2015, doi: 10.1109/TII.2015.2475695.

[14] Y. Al-Eryani and U. Baroudi, "An investigation on detecting bad data injection attack in smart grid," in *Proc. Int. Conf. Comput. Inf. Sci. (ICCIS)*, Apr. 2019, pp. 1–4.

[15] M. Rashed, I. Gondal, J. Kamruzzuman, and S. Islam, "State estimation in the presence of cyber attacks using distributed partition technique," in *Proc. Australas. Universities Power Eng. Conf. (AUPEC)*, Nov. 2020, pp. 1–6.

[16] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Online detection of stealthy false data injection attacks in power system state estimation," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1636–1646, May 2018, doi: 10.1109/TSG.2016.2596298.

[17] H. Karimipour and V. Dinavahi, "Robust massively parallel dynamic state estimation of power systems against cyber-attack," *IEEE Access*, vol. 6, pp. 2984–2995, 2017, doi: 10.1109/ACCESS.2017.2786584.

[18] M. Ganjkhani, S. N. Fallah, S. Badakhshan, S. Shamshirband, and K.-W. Chau, "A novel detection algorithm to identify false data injection attacks on power system state estimation," *Energies*, vol. 12, no. 11, p. 2209, Jun. 2019, doi: 10.3390/en12112209.

[19] R. Chen, X. Li, H. Zhong, and M. Fei, "A novel online detection method of data injection attack against dynamic state estimation in smart grid," *Neurocomputing*, vol. 344, pp. 73–81, Jun. 2019, doi: 10.1016/j.neucom.2018.09.094.

[20] M. Tariq, M. Ali, F. Naeem, and H. V. Poor, "Vulnerability assessment of 6G-enabled smart grid cyber–physical systems," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5468–5475, Apr. 2021, doi: 10.1109/JIOT.2020.3042090.

[21] C. Pei, Y. Xiao, W. Liang, and X. Han, "A deviation-based detection method against false data injection attacks in smart grid," *IEEE Access*, vol. 9, pp. 15499–15509, 2021, doi: 10.1109/ACCESS.2021.3051155.

[22] J. Zhao and L. Mili, "A robust generalized-maximum likelihood unscented Kalman filter for power system dynamic state estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 4, pp. 578–592, Aug. 2018, doi: 10.1109/JSTSP.2018.2827261.

[23] J. Zhao, M. Netto, and L. Mili, "A robust iterated extended Kalman filter for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3205–3216, Jul. 2016, doi: 10.1109/TPWRS.2016.2628344.

[24] N. Živković and A. T. Sarić, "Detection of false data injection attacks using unscented Kalman filter," *J. Modern Power Syst. Clean Energy*, vol. 6, no. 5, pp. 847–859, May 2018, doi: 10.1007/s40565-018-0413-5.

[25] M. Attia, S. M. Senouci, H. Sedjelmaci, E.-H. Aglzim, and D. Chrenko, "An efficient intrusion detection system against cyber-physical attacks in the smart grid," *Comput. Electr. Eng.*, vol. 68, pp. 499–512, May 2018.

[26] P. Bartolomey, E. Kotova, S. Semenenko, and E. Lebedev, "Phasor measurements impact on the *a priori* data filtration and power systems state estimation," in *Proc. 14th Int. Conf. Eng. Modern Electr. Syst. (EMES)*, Jun. 2017, pp. 51–54.

[27] R. A. S. Benedito, L. F. C. Alberto, N. G. Bretas, and J. B. A. London, "Power system state estimation: Undetectable bad data," *Int. Trans. Electr. Energy Syst.*, vol. 24, no. 1, pp. 91–107, Jan. 2014, doi: 10.1002/etep.1744.

[28] Z. Zheng, J. Zhao, L. Mili, and Z. Liu, "Robust unscented unbiased minimum-variance estimator for nonlinear system dynamic state estimation with unknown inputs," *IEEE Signal Process. Lett.*, vol. 27, pp. 376–380, 2020, doi: 10.1109/LSP.2020.2973116.

[29] R. Van der Merwe and E. A. Wan, "The square-root unscented Kalman filter for state and parameter-estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 6. May 2001, pp. 3461–3464, doi: 10.1109/ICASSP.2001.940586.

[30] G. Valverde and V. Terzija, "Unscented Kalman filter for power system dynamic state estimation," *IET Generat., Transmiss. Distrib.*, vol. 5, no. 1, pp. 29–37, Jan. 2011, doi: 10.1049/iet-gtd.2010.0210.

[31] G. Valverde and V. Terzija, "Unscented Kalman filter for power system dynamic state estimation," *IET Generat., Transmiss. Distrib.*, vol. 5, no. 1, pp. 29–37, Jan. 2011, doi: 10.1049/iet-gtd.2010.0210.

[32] J. Zhao and L. Mili, "A robust generalized-maximum likelihood unscented Kalman filter for power system dynamic state estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 4, pp. 578–592, Aug. 2018, doi: 10.1109/JSTSP.2018.2827261.

[33] J. Zhao and L. Mili, "Robust unscented Kalman filter for power system dynamic state estimation with unknown noise statistics," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1215–1224, Mar. 2019, doi: 10.1109/TSG.2017.2761452.

MUHAMMAD RASHED received the bachelor's degree from the University of Engineering and Technology Peshawar, Pakistan, in 2002, and the master's degree from the University of South Australia, in 2006. He is currently pursuing the Ph.D. degree with Federation University, Australia. Since then, he has been working in the field of telecommunications networks. His research interests include cyber security and smart grid.

JOARDER KAMRUZZAMAN received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from the Bangladesh University of Engineering and Technology, Dhaka, and the Ph.D. degree in information systems engineering from the Muroran Institute of Technology, Hokkaido, Japan, in 1993. He is currently a Professor with the School of Science, Engineering and Information Technology, Federation University, Australia. He has 29 years of tertiary level teaching experience in Australia and Bangladesh.

IQBAL GONDAL led the Internet Commerce Security Laboratory (ICSL), Federation University, as the Director for last seven years to conduct translational research in cybersecurity. He was responsible for establishing collaborative partnerships between Federation University and Australian Cyber Security Centre (ACSC) and Australian Federal Police (AFP). He was also the Associate Dean at the STEM School, Federation University, where he successfully developed partnerships with domestic and international institutions for teaching and research. He is currently the Associate Dean of cloud, systems and security at RMIT University. He has worked in industry and academia for 25 years in Singapore and Australia.

SYED ISLAM (Fellow, IEEE) received the B.Sc. degree in electrical engineering from the Bangladesh University of Engineering and Technology, Bangladesh, in 1979, and the M.Sc. and Ph.D. degrees in electrical power engineering from the King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, in 1983 and 1988, respectively. He is currently the Dean of the School of Engineering, Information Technology and Physical Sciences, Federation University, Australia.