Detecting Data Integrity Attacks on SCADA Systems Using Limited PMUs

Seemita Pal*, Biplab Sikdar† and Joe Chow*
*Rensselaer Polytechnic Institute and †National University of Singapore

***Abstract*—In the power grid, Supervisory Control and Data Acquisition (SCADA) systems are used for executing various applications for monitoring and controlling purposes, which in turn enable stable operation of the grid. The integrity and timely delivery of SCADA data is critical to the operation of the grid and this makes them an attractive target for cyber-attacks. However, SCADA systems have various vulnerabilities which may be exploited to launch attacks, leading to a number of security challenges. To address one of these security challenges, this paper proposes a technique for detecting data manipulation attacks on SCADA systems. The proposed methodology is based on utilizing synchrophasor measurements from Phasor Measurement Units (PMUs) that are increasingly being deployed in power grids. The proposed method exploits the correlation between SCADA and PMU data, and the classification of tampered and real data is done through a difference measure developed in this paper. Simulations performed on real grid data show that the detection technique is highly effective at attack detection.**

I. Introduction

SCADA systems have traditionally played a vital role in the control of various critical infrastructures like power grids, water distribution and irrigation networks, communication networks as well as oil and gas pipelines. In power grids, SCADA systems collect information from field instruments via Remote Terminal Units (RTUs) or Programmable Logic Controllers (PLCs) installed at various substations and transfer them to the central master station at the control center [1]. The RTUs commonly provide the active and reactive power flows, power injections, magnitudes of line currents and bus voltages measured at the substations at a refresh rate of about 2 to 5 seconds.

SCADA systems were developed decades ago as isolated networks but the need to remotely monitor and control systems has resulted in the interconnections with the enterprise communication infrastructures and thereby the Internet. Furthermore, since network security was a matter of little concern at the time when SCADA was developed, security was not designed into the system. The system design aimed to maximize functionalities, accessibility, and easier debugging, and some of these very features make it more vulnerable to cyber-attacks. In recent times, both capabilities to attack the SCADA systems as well as the number of cyber-attacks reported are on the rise [2]. In its annual Threat Report for 2015, Dell Security reported that worldwide cyber-attacks on SCADA systems have been increasing at an alarming rate with the number of incidents reported increasing from 163,228 in 2013 to 675,186 in 2014. Since many such incidents go unreported, the actual number may be much higher [3]. Vulnerabilities associated with SCADA systems include direct tampering of the RTUs, denial-of-service attacks, deletion of system files, modifying data logs, unauthorized changes to commands or alarm thresholds etc. Note that the use of leased lines, as preferred by utilities, does not ensure security. Tapping these lines as well as compromising frequency hopping spread spectrum radio and other wireless communication mechanisms, which are frequently used to control RTUs is not difficult [4].

This paper addresses the problem of detecting data modification attacks on SCADA systems in power grids. In such attacks the adversary modifies the contents of the data packet containing SCADA measurements, with the objective of biasing the state estimates of the grid. Such incorrect state estimates may lead the operator to take incorrect control actions or dispatch decisions, which in turn may cause disruption of operation, damage to equipments, injury to personnel, monetary losses, and even blackout. The proposed detection method is based on using measurements from synchrophasors, that are becoming increasingly popular for wide-area monitoring and control of the power grid. The synchrophasor system consists of PMUs which measure bus voltage magnitude and phase angle, branch current magnitudes and phase angles, frequency and rate of change of frequency at the buses where they are installed. These measurements are generated at a rate of 20, 30 , 50 or 60 samples per second and are accurately time-stamped using inbuilt or external GPS units. These fast and synchronized measurements enable dynamic state estimation over a wide geographical area, thereby aiding in increasing the reliability of the system. However, due to the high cost of PMUs as well as their communication facilities, limited number of PMUs have been installed grid-wide [5]. The deployment is taking place in phases and is expected to continue over the near future.

The main contribution of this paper is a technique for using already existing SCADA systems and a limited number of PMUs in order to verify the integrity of the SCADA datasets and detect cyber attacks, if any. The proposed methodology is based on exploiting the correlation between SCADA and PMU measurement data. We first develop a metric to quantify the difference between the SCADA and PMU measurements based on two factors: the divergence and the correlation coefficient. The metric is then used to detect malicious modifications in the SCADA data. The proposed detection technique works

This work was supported primarily by the ERC Program of the National Science Foundation and the Department of Energy under NSF Award Number EEC-1041877 and the CURENT Industry Partnership Program.

irrespective of the strategy used for PMU placement. Our methodology has been verified using simulations on real data from the New York power grid.

The rest of paper is organized as follows. Section II discusses the related work. In Section III, we define the threat model. Section IV explains the basis of the detection scheme. The attack detection scheme is then discussed in Section V. In Section VI, simulation results are provided for verifying the effectiveness of the detector. Finally Section VII concludes the paper.

II. Related Work

In the recent times, utilities and Independent System Operators (ISOs) are increasingly focusing on the security of SCADA and synchrophasor systems. Research efforts are underway for developing mechanisms for preventing as well as detecting cyber-attacks on these systems.

In 2009 it was first shown that, if the grid configuration information is available, it is possible for a malicious attacker to design coordinated attacks which can bias the system states without being detected by traditional bad data detection schemes [6]. Such attacks have been referred to as false data injection attacks. Indices were introduced for quantifying the least effort needed by attackers to design successful false-data injection attacks while avoiding bad data detection in [7]. There has been research on the formulation of attack vectors and optimal placement of sensors for injecting spurious data by the attacker. The authors in [8] have shown that it is a necessary but not a sufficient condition to protect at least a certain number of measurements in order to be able to ensure observability and enable detection of such attacks. In [9], a Bayesian framework which leverages the knowledge of prior distribution on the states for performing attack detection was proposed. The problem of estimating the smallest number of meters that are required to be compromised by the attacker has been modeled in [10], [11]. In [12], irreducible attacks have been defined and an algorithm which is based on graph theory for finding all irreducible attacks has been proposed. Two algorithms are proposed for determining the placement of encrypted devices in the system in order to maximize the detection of stealthy false-data attacks in [13]. In [14], a mechanism based on evaluation of the equivalent line impedances for detecting data manipulation attacks in synchrophasor data has been proposed. To the best of our knowledge, the existing false-data injection attack detection techniques assume that the defenders have some number of secure PMUs. But no PMU can be expected to provide absolutely accurate data and zero possibility of corruption by attackers at all times. Therefore, detection techniques are required that can detect manipulation of data by adversary without requiring such assumptions.

III. System and Threat Model

The system model assumed in this paper is that all buses are equipped with SCADA capabilities. Therfore, active or reactive power flows and power injection measurement data are available at these buses. However, PMUs are available only on a subset of the buses and these PMUs have sufficient channels to measure the voltage magnitude and angle of the bus, the current magnitudes and angles of all branches incident to that bus, as well as the frequency and the rate of change of frequency.

The threat model assumed is that the adversary has compromised one or more of the sensors, network routers or/and communications links. At each of the compromised nodes, the adversary has the ability to manipulate measurement data in order to bias the power system state estimates. PMUs are more sophisticated devices and due to the use of NASPInet architecture, the communication systems of synchrophasor systems are assumed to be more secure than that of SCADA [15]. It is assumed that the attacker has limited resources and can only successfully compromise the SCADA system. The data is assumed to be either unencrypted or the encryption has been broken. The attacker may even directly compromise the sensors and in that case it is not necessary to break encryption or steal cryptographic keys.

The data manipulated by the malicious attacker can bias the system state estimates and thereby influence the control center into taking suboptimal dispatch decisions or wrong control actions. This can lead to the adversary's monetary gains, operation disruption or equipment damage. To maximize the damage, the objective of the adversary is to manipulate data to the maximum extent possible since larger biasing is more likely to lead to erroneous actions of greater consequence. However, even relatively small changes can cause uneconomic dispatch choices. Our objective is to develop a detection technique that will effectively detect such cyber attacks.

The adversary may manipulate the SCADA measurements, i.e. the active or reactive power measurements. Two kinds of attacks are considered: ramp and step. In ramp attack, the attacker slowly and monotonously changes the data from its original value to make detection difficult. In step attack, the attacker abruptly changes the SCADA measurement data in order to influence the operators into taking immediate erroneous control actions which may be damaging for the system.

IV. The Basis of Detection Mechanism

In the system model assumed in this paper, PMUs are present on some buses while SCADA measurements are available on all buses. So we can classify the buses into two kinds: PMU buses and non-PMU buses. The non-PMU buses can be classified according to the degree of their connection with the closest PMU bus. The non-PMU buses connected directly to any PMU bus are called 1st degree non-PMU buses and their pseudo-measurements computed using the measurements from the assigned neighboring PMU bus are called 1st degree pseudo measurements. The buses connected to any 1st degree non-PMU bus and not connected to any PMU bus are called 2nd degree non-PMU buses and their computed pseudo measurements are called 2nd degree pseudo-measurements and so on.

For the PMU buses, both PMU and SCADA measurements are available. These SCADA and PMU measurements are essentially the sensor values of the same system and have a known relation. We quantify the difference measure between these two datasets to determine whether they conform with each other or if false data is present. The PMU data is used to compute the pseudo-measurements of the corresponding 1st degree non-PMU buses and communicated to them. These buses, in turn, calculate the difference measure between the 1st degree pseudo-measurements and the SCADA datasets to determine presence or absence of cyber-attacks and location, if any. The method is implemented at all the buses in a percolation-like manner till all the buses have been checked for possible data modification.

Consider a power system with N buses, labeled as $i = 1, 2, \cdots, N$. At the PMU buses, whenever a set of SCADA data comes in, the verification process begins. Every SCADA data typically consists of an associated time-tag, say t^S. Let the active and reactive power flowing from bus i to bus k, as measured by the SCADA system, be denoted by $P^S_{m,ik}$ and $Q^S_{m,ik}$, respectively. Here, m denotes the measurement serial number corresponding to the SCADA time-tag t^S. The PMU and SCADA datasets can be time-synchronized based on the correlation coefficient between the datasets since the coefficient will be maximum when the two measurements correspond to the same time.

The PMUs measure the voltage and all the currents incident on the bus where it is installed. The estimated bus voltage magnitudes and phase angles at bus i corresponding to time-tag t^S are denoted by $V_{i,m}$ and $\theta_{i,m}$, respectively. The magnitude and the phase angle of the current flowing from PMU-bus i to bus k are denoted by $I_{ik,m}$ and $\delta_{ik,m}$ respectively. One of the PMU buses is selected as the reference bus. The phase angle of the reference bus is subtracted from all the phase angle measurements to obtain the phase angles with respect to the reference bus. Line power flows are calculated using these voltage and current phasor measurements. Assuming the voltage magnitude is a phase-to-neutral value, which is the case for positive-sequence measurements, three-phase active and reactive power flows on lines can be calculated as follows:

$$P^P_{ik,m} = 3V_{i,m}I_{ik,m}\cos\phi_{ik,m} \tag{1}$$

$$Q^P_{ik,m} = 3V_{i,m}I_{ik,m}\sin\phi_{ik,m} \tag{2}$$

where $\phi_{ik,m} = \theta_{i,m} - \delta_{ik,m}$. Since all the derivations are provided for power flows on the line between buses i and k, from here we will omit the subscript 'ik' for the simplicity of notations.

V. Difference Measure and Attack Detection

In this section, we describe the difference measure that is used for determining whether the SCADA and PMU measurements conform with each other and detect presence of modified data, if any. The values of active and reactive power flow calculated using PMU data are compared with the power measurements obtained from the SCADA system using a difference measure described in this section. This difference measure takes in account both the relative Euclidean distance between the two datasets as well as the correlation between them. Thus, the difference measure consists of two factors: (i) Divergence factor and (ii) Miscorrelation factor.

A. Divergence Factor

The power flows computed using the two systems, i.e. SCADA and PMU system, are never exactly the same. This is because of different scaling factors, biases or calibration of the different sensors [16]. The difference between the two measurement systems can be modeled as the sum of a constant bias and noise. The difference in the active and reactive powers is given by,

$$P^P_m - P^S_m = K_p + v_{p,m} \tag{3}$$

$$Q^P_m - Q^S_m = K_q + v_{q,m} \tag{4}$$

where K is the constant bias that depends on the various scaling factors, biases or calibrations errors of the associated sensors and v is noise which is assumed to be normal with zero mean and standard deviation σ.

Therefore, the Euclidean distance between the the power flows computed using the two measurement systems can be monitored to determine whether the datasets are diverging from one another. Let the window for calculation be w samples points. The divergence factor is the ratio of the Euclidean distance between SCADA and PMU-generated active/reactive power for the current window (which consists of the latest w sample points) to the past window (which consists of w sample points preceding the current window). The expression for the divergence factor is:

$$e_{p,m} = \frac{\sqrt{\sum_{j=m-w+1}^{m}(P^S_j - P^P_j)^2}}{\sqrt{\sum_{j=m-2w+1}^{m-w}(P^S_j - P^P_j)^2}} \tag{5}$$

$$e_{q,m} = \frac{\sqrt{\sum_{j=m-w+1}^{m}(Q^S_j - Q^P_j)^2}}{\sqrt{\sum_{j=m-2w+1}^{m-w}(Q^S_j - Q^P_j)^2}}. \tag{6}$$

Using Equation (3) in Equation (5), we get,

$$e_{p,m} = \frac{\sqrt{\sum_{j=m-w+1}^{m}(K_p - v_{p,j})^2}}{\sqrt{\sum_{j=m-2w+1}^{m-w}(K_p - v_{p,j})^2}} = \sqrt{1 + \frac{\sigma^2_{p,m} - \sigma^2_{p,m-w}}{K_p^2 + \sigma^2_{p,m-w}}}. \tag{7}$$

If the distribution of the datasets in the current window and the past window are same, then $\sigma_{p,m} \approx \sigma_{p,m-w}$. Hence the second term in the expression will be almost zero, making the divergence factor approximately one. Also, it should be noted that, as the bias value i.e. K_p increases, the sensitivity of the divergence factor decreases. Hence, the expected bias value is initially estimated empirically using measurements from both the datasets and subtracted before computing the Euclidean distances. It is taken to be the mean of the difference between PMU and SCADA vectors for any particular electrical quantity. Therefore, the modified expression for the

divergence factor for mitigating the loss of sensitivity due to bias becomes:

$$e_{p,m} = \frac{\sqrt{\sum_{j=m-w+1}^{m}(P_j^S - P_j^P - K_p')^2}}{\sqrt{\sum_{j=m-2w+1}^{m-w}(P_j^S - P_j^P - K_p')^2}}. \quad (8)$$

Here, K_p' is the estimated expected bias between the SCADA and PMU datasets and $K_p' \sim K_p$. Even if the estimated bias is not exactly equal to the current bias value, we can say $K_p - K_p' = c_p$, where c_p is very small in value and negligible in most cases. The modified divergence factor in the normal case, when there is no attack is therefore given by:

$$e_{p,m} = \sqrt{1 + \frac{\sigma_{p,m}^2 - \sigma_{p,m-w}^2}{c_p^2 + \sigma_{p,m-w}^2}} = \sqrt{1 + s_n} \quad (9)$$

where, $s_n = \frac{\sigma_{p,m}^2 - \sigma_{p,m-w}^2}{c_p^2 + \sigma_{p,m-w}^2}$ is the sensitivity of the divergence factor in the normal case, when there is no attack.

The modified divergence factor for the reactive power datasets is similarly given by:

$$e_{q,m} = \frac{\sqrt{\sum_{j=m-w+1}^{m}(Q_j^S - Q_j^P - K_q')^2}}{\sqrt{\sum_{j=m-2w+1}^{m-w}(Q_j^S - Q_j^P - K_q')^2}}. \quad (10)$$

Also, the following expression can be derived for the reactive power datasets in case of no attack:

$$e_{q,m} = \sqrt{1 + \frac{\sigma_{q,m}^2 - \sigma_{q,m-w}^2}{c_q^2 + \sigma_{q,m-w}^2}}. \quad (11)$$

Let the elements of the attack vector be denoted by $a_{p,j}$. In the presence of an attack, the modified SCADA data is given by, $P_j^{S\prime} = P_j^S - a_{p,j}$. The divergence factor can be computed for the transition period, when the current window has attack (i.e. modified measurements) while the past window doesn't. For the current window, the difference between the two active powers is given by:

$$P_j^P - P_j^{S\prime} = K_p + v_{p,j} + a_{p,j} \quad (12)$$

where, $j = m-w+1, m-w+2, \cdots, m-1, m$. In case of step attack, if the attack vector elements are assumed to be constant at a, the divergence factor is given by:

$$e_{p,m} = \sqrt{1 + \frac{\sigma_{p,m}^2 - \sigma_{p,m-w}^2}{c_p^2 + \sigma_{p,m-w}^2} + \frac{a^2 + 2c_p a}{c_p^2 + \sigma_{p,m-w}}} \quad (13)$$

$$e_{p,m} = \sqrt{1 + s_n + s_a} \quad (14)$$

where, $s_a = \frac{a^2 + 2c_p a}{c_p^2 + \sigma_{p,m-w}}$ is the sensitivity to injected false data in case of attack. As the magnitude of the attack vector elements increases, s_a increases. Also, it is seen that the bias-correction leads to the divergence factor being more sensitive to attacks.

B. Miscorrelation Factor

If both the SCADA and PMU data sets are true (i.e. unmodifed), they will have a high correlation and the correlation coefficient is expected to be close to 1. However, if false-data is injected in any of the data sets, the correlation coefficient will decrease in value. The vector consisting of latest w data points for a given bias is denoted by:

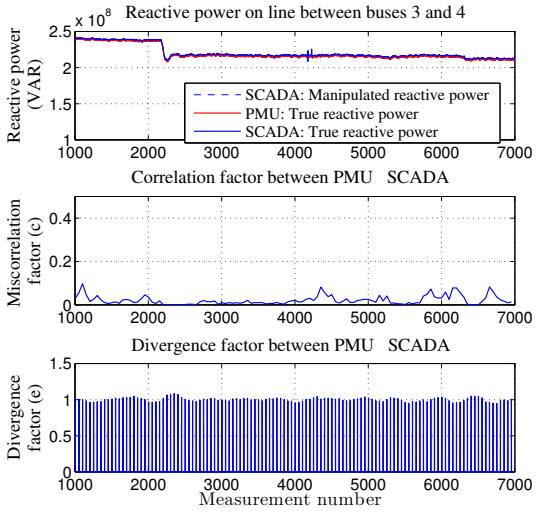

Fig. 1. Plots showing correlation factor and divergence factor in case of no attack.

$$\mathbf{P_m} = [P_{m-w+1}, \cdots, P_{m-1}, P_m]^T \quad (15)$$

$$\mathbf{Q_m} = [Q_{m-w+1}, \cdots, Q_{m-1}, Q_m]^T \quad (16)$$

The correlation coefficient between the two sets of data are calculated as follows:

$$r_{p,m} = \frac{Cov(\mathbf{P_m^P}, \mathbf{P_m^S})}{\sqrt{Var(\mathbf{P_m^P})}\sqrt{Var(\mathbf{P_m^S})}} \quad (17)$$

$$r_{q,m} = \frac{Cov(\mathbf{Q_m^P}, \mathbf{Q_m^S})}{\sqrt{Var(\mathbf{Q_m^P})}\sqrt{Var(\mathbf{Q_m^S})}}. \quad (18)$$

We know that in the normal case without attack,

$$P_j^P = P_j^S + K_p + v_{p,j} \quad j = 1, 2, \cdots, m. \quad (19)$$

Therefore, the covariance and the variances can be expressed as:

$$Cov(\mathbf{P_m^P}, \mathbf{P_m^S}) = Var(\mathbf{P_m^S}) + Cov(\mathbf{v_{p,m}}, \mathbf{P_m^S}) \quad (20)$$

$$Var(\mathbf{P_m^P}) = Var(\mathbf{P_m^S}) + Var(\mathbf{v_{p,m}}). \quad (21)$$

If the noise is assumed to be negligible compared to the true power values, $r_{p,m} \approx 1$. Thus, if the two datasets are true, they are highly correlated and the value of the correlation coefficient is almost equal to 1.

However, when false data in injected in the SCADA measurements, $P_j^P = P_j^{S\prime} + a_{p,j} + v_j + K_p \quad j$ in presence of attack, Therefore, the expressions for covariance and variance can be derived as follows,

$$Cov(\mathbf{P_m^P}, \mathbf{P_m^{S\prime}}) = Var(\mathbf{P_m^{S\prime}}) + Cov(\mathbf{v_{p,m}}, \mathbf{P_m^{S\prime}}) + Cov(\mathbf{a_{p,m}}, \mathbf{P_m^{S\prime}}) \quad (22)$$

$$Var(\mathbf{P_m^P}) = Var(\mathbf{P_m^{S\prime}}) + Var(\mathbf{v_{p,m}}) + Var(\mathbf{a_{p,m}}). \quad (23)$$

If the attack vector is non-negligible, the denominator of the correlation coefficient increases in value while the numerator may increase slightly or may decrease due to low or negative value of $Cov(\mathbf{a_{p,m}}, \mathbf{P_m^{S\prime}})$. Therefore, the correlation coefficient deviates from 1.

The miscorrelation factor is taken as the difference between 1 and the absolute value of the correlation coefficient.

$$c_{p,m} = 1 - |r_{p,m}| \tag{24}$$

$$c_{q,m} = 1 - |r_{q,m}|. \tag{25}$$

Therefore, as the correlation between the two datasets decreases, the miscorrelation factor increases in value. It is close to zero when the two datasets are highly correlated. Figure 1 shows the divergence factors and miscorrelation factors for the PMU and SCADA datasets of reactive power flowing in a line when there is no attack. The PMU data have been obtained from the NY power grid. Since corresponding SCADA data were not available, but representative state estimator data were available, we have used that as a substitute for SCADA data.

C. Difference measure

The difference measure between the two sets of data is expressed as the product of divergence factor and miscorrelation factor:

$$d_{p,m} = e_{p,m} \times c_{p,m} \tag{26}$$

$$d_{q,m} = e_{q,m} \times c_{q,m} \tag{27}$$

In case of true data, the divergence factor is expected to be close to one and the miscorrelation factor is expected to be much less than one. Hence, the difference measure is expected to be less than 1. However, in case of false-data attack, the divergence factor is expected to exceed 1 and the miscorrelation factor is expected to be closer to one. Therefore, the overall difference measure is expected to exceed 1.

If the difference measure computed for SCADA and PMU datasets at any point is computed to be more than one, then we can say that the datasets are not conforming with each other and the presence of false data is inferred. Figure 2 shows the plot of the divergence factors, miscorrelation factors and difference measures for a line when there is ramp attack on SCADA data between measurement number 4001 and 5000, with maximum manipulation of 5% at measurement number 4500. Figure 3, on the other hand, shows the plots when there is step attack of 5% on SCADA data between measurement numbers 4000 and 5000. From both these sample figures, it is evident that when there is no attack the divergence factor is close to 1 in value. However, when there is presence of attack, the divergence factor increases significantly. The miscorrelation factor remains low in normal cases but increases to close to 1, when there is attack. The resulting difference measure is observed to exceed 1 when there is attack.

VI. Simulation Results

In this section we present simulation results to verify the proposed detection mechanism. Real PMU data collected from various locations in New York have been used to verify the mechanism. Sets of more than 9000 samples of measurements of PMU data from two buses with two lines connected to one bus and three lines connected to the other, have been used for evaluating the effectiveness of the proposed detection method. Representative state estimator data has been used in place of SCADA data. To simulate an attack, the values of the SCADA measurements were altered. Two sets of simulations have been

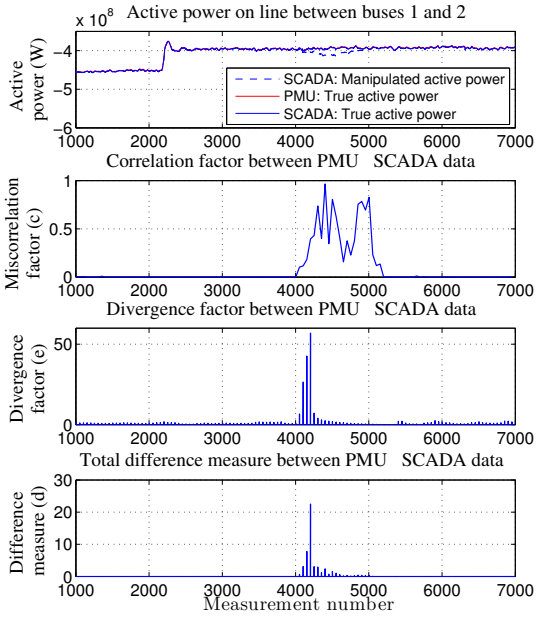

Fig. 2. Plots showing active power on line between buses 1 and 2 and the related factors and difference measure.

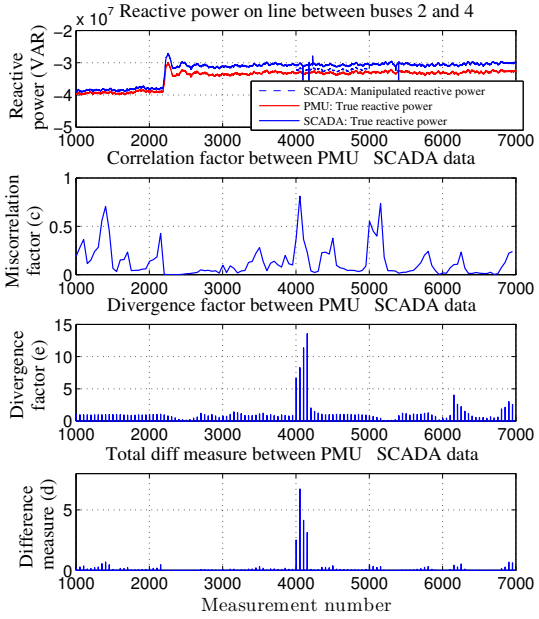

Fig. 3. Plots showing reactive power on line between buses 2 and 4 and the related factors and difference measure.

Algorithm 1 SCADA Data Attack Detection Using PMU Data

loop
  for arrival of m^{th} SCADA measurement at bus i **do**
    SCADA measurement on line ik: $P^S_{m,ik}$, $Q^S_{m,ik}$;
    SCADA measurement time-tag = t^S
    PMU measurements at t^S: V_i, θ_i, I_{ik}, δ_{ik};
    Calculate PMU-generated active power: $P^P_{j,ik}$
    Calculate PMU-generated reactive power: $Q^P_{j,ik}$
    Calculate divergence factors: $e_{p,ik,m}$, $e_{q,ik,m}$
    Calculate miscorrelation factors: $c_{p,ik,m}$, $c_{q,ik,m}$
    Determine difference measure for active and reactive power: $d_{p,ik,m}$, $d_{q,ik,m}$
    if $(d_{p,ik,m} > 1) \vee (d_{q,ik,m} > 1)$ **then**
      Generate data integrity attack alarm;
    end if
  end for
end loop when session is terminated

performed: firstly with modification in SCADA active power data and secondly with modification in SCADA reactive power data.

In order to evaluate the effectiveness of the proposed data manipulation attack detection mechanism, we consider two types of attacks that may be executed on SCADA data: ramp and step. Different levels of modifications have been simulated for both the attack types, the level of modification being the percentage of increase of the manipulated value from the true measurement value. For each set of SCADA data, five levels of modifications have been simulated, that is, 0% or no change, 2%, 5%, 10% and 20% changes.

We evaluate the performance of the proposed detection mechanism in terms of its accuracy (ACC), false positive (FP) rates, and false negative (FN) rates. The accuracy is the probability of correct detection. The false positive rate is the probability that a data manipulation attack alarm is raised when in reality there was no data manipulation. The false negative rate is the probability that a data manipulation attack goes undetected.

TABLE I
OVERALL DETECTION RESULTS USING PROPOSED DATA MODIFICATION ATTACK DETECTION ALGORITHM.

Modification percentage	Ramp change			Step change		
	A	FP	FN	A	FP	FN
0	80	0	20	80	0	20
2	85	0	15	80	0	20
5	100	0	0	90	0	10
10	100	0	0	95	0	5
20	100	0	0	100	0	0

The overall results of detecting data manipulation attacks are provided in Table I. The accuracy of the detection scheme is above 80% in all cases. The accuracy increases with increase in the modification level or if the modification is in the form of step-change.

It should be mentioned here, that the data used for the simulation are disturbance data. Even in the presence of disturbances, the proposed detection mechanism detects most of the attacks accurately.

VII. CONCLUSIONS

This paper proposes an effective and distributed method of detecting data attacks on SCADA systems in the power grid. It uses the available PMU measurements to perform validation of the SCADA datasets. It does not matter what placement scheme has been followed to place the PMUs as long as there are a few PMUs spread over the grid. The detection technique is found to be accurate, computationally inexpensive and can even detect attacks with manipulation as low as 2%.

REFERENCES

[1] K. Stouffer, J. Falco, and K. Kent, "Guide to supervisory control and data acquisition (SCADA) and industrial control systems security", *Special Publication NIST-SP-800-82-2006*, National Institute of Standards and Technology (NIST), 2006.

[2] http://www.eetimes.com/document.asp?doc_id=1327785, Last accessed: 04/07/2016.

[3] DELL, "Dell Security Annual Threat Report," https://software.dell.com/docs/2015-dell-security-annualthreat-report-white-paper-15657.pdf, Last accessed: 04/08/2016.

[4] Y. Wang and Y. Xiao, "Security and Privacy in Smart Grids", *CRC Press*, pp. 245-268, July 2013.

[5] R. F. Nuqui and A. G. Phadke, "Phasor measurement unit placement techniques for complete and incomplete observability," *IEEE Transactions on Power Delivery*, vol. 20, no. 4, pp. 2381-2388, Oct. 2005.

[6] Y. Liu, P. Ning and M. Reiter, "False data injection attacks against state estimation in electric power grids", *Proc. of ACM CCS,* Chicago, IL, November 2009.

[7] H. Sandberg, A. Teixeira, and K. Johansson, "On Security Indices for State Estimators in Power Networks", *Proc. of Workshop on Secure Control Systems,* Stockholm, Sweden, 2010.

[8] R. B. Bobba, K. M. Rogers, Q. Wang, and H. Khurana, "Detecting false data injection attacks on DC state estimation", *Proc. of Workshop on Secure Control Systems,* Stockholm, Sweden, 2010.

[9] O. Kosut J. Liyan, R. Thomas and T. Lang, "Malicious Data Attacks on Smart Grid State Estimation: Attack Strategies and Countermeasures," *Proc. of IEEE SmartGridComm,*, pp. 220-225, Gaithersberg, MD, Oct. 2010.

[10] T. Kim and H. V. Poor, "Strategic Protection Against Data Injection Attacks on Power Grids," *IEEE Transactions on Smart Grid,* vol. 2, no. 2, pp. 326-333, June 2011.

[11] C. Shuguang, H. Zhu, S. Kar, T. Kim, H. V. Poor and A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions,"*IEEE Signal Processing Magazine,* vol. 29, no. 5, pp. 106-115, Sept. 2012.

[12] A. Giani, R. Bent, M. Hinrichs, M. McQueen and K. Poolla, "Metrics for assessment of smart grid data integrity attacks,"*Proc. of IEEE Power and Energy Society General Meeting,*pp. 1-8, San Diego, CA, July 2012.

[13] G. Dan and H. Sandberg, "Stealth Attacks and Protection Schemes for State Estimators in Power Systems," *Proc. of IEEE SmartGridComm,* Gaithersburg, MD, 2010, pp. 214-219.

[14] S. Pal, B. Sikdar and J. Chow, "Detecting Malicious Manipulation of Synchrophasor Data," *Proc. of IEEE SmartGridComm,* Miami, FL, Nov. 2015.

[15] R. Bobba, E. Heine, H. Khurana and T. Yardley, "Exploring a tiered architecture for NASPInet," *Proc. of IEEE ISGT*, Gaithersburg, MD, 2010.

[16] S. Ghiocel, "Applications of synchronized phasor measurements for state estimation, voltage stability, and damping control", *Thesis (Ph.D.)*, Rensselaer Polytechnic Institute, May 2013.