This article has been accepted for publication in a future issue of IEEE Transactions on Smart Grid. Citation information: DOI 10.1109/TSG.2022.3150954.

Detection and Localization of PMU Time Synchronization Attacks via Graph Signal Processing

Ezzeldin Shereen, Raksha Ramakrishna, György Dán,
School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Sweden
E-mail: {eshereen, rakshar, gyuri}@kth.se

***Abstract*—Time Synchronization Attacks (TSAs) against Phasor Measurement Units (PMUs) constitute a major threat to modern smart grid applications. By compromising the time reference of a set of PMUs, an attacker can change the phase angle of their measured phasors, with potentially detrimental impact on grid operation and control. Going beyond traditional residual-based techniques in detecting TSAs, in this paper we propose the use of Graph Signal Processing (GSP) to model the power grid so as to facilitate the detection and localization of TSAs. We analytically show that modeling the state of the power system as a low-pass graph signal can significantly improve the resilience of the grid against TSAs. We propose TSA detection and localization methods based on GSP, leveraging state-of-the-art machine learning algorithms. We provide empirical evidence for the efficiency of the proposed methods based on extensive simulations on five IEEE benchmark systems. In fact, our methods can detect at least 77% more TSAs of significant impact and localize an additional 70% of the attacked PMUs compared to state-of-the-art techniques.**

***Keywords*—** Time Synchronization Attack, Phasor Measurement Unit, Graph Signal Processing, Power System State Estimation, Attack Detection and Identification, Machine Learning

I. Introduction

Wide Area Monitoring Protection and Control (WAMPAC) systems rely heavily on Phasor Measurement Units (PMUs) to enable a plethora of smart grid applications. With streaming rates of up to tens of measurements per second, PMUs offer much higher situational awareness than traditional Supervisory Control And Data Acquisition (SCADA) measurements. To provide accurate, timestamped, high-rate measurements of voltage and current phasors, precise time synchronization is essential for PMUs for which they can rely on space-based time synchronization (SBTS) or network-based (NBTS) mechanisms [1]. In the case of SBTS, a PMU synchronizes itself to the time reference received from a set of GPS satellites, which are equipped with accurate atomic clocks. In NBTS, the PMU clock acts as a *slave* device and adjusts its time to a reference received from a *master* clock device, which is equipped with an accurate clock. The most prominent protocol for NBTS is the Precision Time Protocol (PTP), and is able to achieve sub-microsecond time synchronization accuracy.

A major drawback of both synchronization mechanisms is the absence of integrated security controls, and therefore, both mechanisms have been shown vulnerable to Time Synchronization Attacks (TSAs) [2]. Civilian GPS signals used for SBTS are unauthenticated, and thus vulnerable to spoofing attacks [3]. PTP used for NBTS is vulnerable to software compromise [4] and to delay manipulations using, e.g., delay box insertion [5], despite the support for optional message authentication in the latest version, PTPv2.1.

In principle, applying Linear State Estimation (LSE) and associated Bad Data Detection (BDD) methods on the PMU measurements may detect TSAs [6], [7]. Nevertheless, recent work has shown that LSE is vulnerable to undetectable TSAs [8], [9], i.e., TSAs can be constructed such that they do not change the measurement residuals, and thus can bypass LSE-based BDD methods. The detection and localization of TSAs, i.e., identifying the attacked PMUs, are thus challenging open issues, at the same time they are of utmost importance to power system operators.

In this paper, we address the problem of detecting and localizing TSAs based on the framework of Graph Signal Processing (GSP) [10] for power grids [11], [12], and make three major contributions:

1) We extend the GSP framework for detecting TSAs by proposing novel detection metrics that capture the effect of a TSA.
2) We further exploit the GSP framework to propose methods for effectively localizing an attack.
3) We employ machine learning (ML) algorithms that leverage GSP features to further improve the detection and localization performance.

Unlike most of the previous work in the detection of TSAs, the proposed approach relies on data that is readily available to the system operator, works irrespective of the employed time synchronization mechanism, and does not require the use of larger system models [13]. Our numerical results show that the proposed methods outperform state-of-the-art TSA detection and localization methods.

The rest of this paper is organized as follows. Section II reviews prior work in the field of TSA detection and localization, as well as previous work using GSP in power systems. Section III presents the system and attack models. Section IV reviews concepts from GSP and motivates its use for detection and localization. The proposed GSP-based approaches for detecting and localizing TSAs are described in Section V, and evaluated and compared to state-of-the-art methods in Section VI. Section VII concludes the paper.

II. Related Work

The detection, localization, and correction of attacks against measurement integrity in power systems has received sig-

nificant attention in recent years [14], including attacks that bypass residual-based BDD methods, often called False Data Injection Attacks (FDIAs). FDIAs can target DC state estimation [15], [16], AC state estimation [17], [18], as well as PMU-based LSE [19]. Moreover, these attacks can bypass BDD methods even if only a few measurements are manipulated by the attacker [20], [21]. Research on the detection and protection against FDIAs focus on physical protection of strategically chosen sensors [16], [22], detection based on alternative state estimation models [23], [24], and data-driven detection [25]. One recent promising approach for detecting FDIAs is the framework of Graph Signal Processing [11], [26], which was also proposed for detecting anomalies in PMU data [27]. Importantly, [12] introduced the Grid-GSP framework and showcased the performance of GSP in different smart grid applications.

Due to the special nature of time synchronization attacks, the detection and localization of TSAs has often been studied separately from FDIAs. A number of works considered detecting TSAs using data from the PMU clock synchronization system [28], [29]. These approaches inherently localize the attacked PMUs, as the detection is per PMU, but are dependant on the underlying synchronization mechanism. For example, [28] detects GPS spoofing attacks by leveraging the fact that PMU locations are known, and hence the set of visible satellites at a given point in time can be estimated, as well as observing changes in the received GPS signal statistics. For time synchronization using PTP, [29] proposes the introduction of devices called "guard clocks" to the network in order to detect time reference manipulations.

An independent line of works analyze the effect of TSAs on PMU and SCADA measurements. Authors in [30] propose correcting a TSA, assuming that only one measurement is attacked. They do so by formulating an optimization problem with the objective of minimizing the SE residuals as a function of the attack phase angle shift. This approach is generalized by [6] to the case when multiple measurements are attacked. The proposed solution first identifies the attacked measurements by analyzing the SE residuals, and then corrects the phase angles of the identified measurements similar to [30]. Another residual-based approach is [31] which proposes the identification of the attacked measurements by observing the residuals of slightly perturbed measurements. Moreover, [32] proposes the detection of TSAs by estimating the rotor angles of the generators using dynamic state estimation utilizing fused PMU and SCADA measurements. Authors in [33] developed an alternating minimization approach to reconstruct and correct a TSA. Authors in [34] follow a two-step approach, which first identifies attacked PMUs using SE residuals, and then finds the sparsest TSA against the identified measurements. The paper also provides formal conditions on the network topology, PMU locations, and the number of spoofed PMUs, under which TSAs are identifiable using the proposed approach. The same approach was used in [35] on a low-dimensional representation of PMU data learned through Principal Components Analysis. Finally, [36] leverages PMU measurements at both ends of a transmission line to correct TSAs by assuming that the line admittance is unchanged during short periods of time.

Several works proposed combining information from the time synchronization system and PMU measurements to detect and localize TSAs. For instance, [37] proposed evaluating the trustworthiness of a PMU based on analyzing the Carrier-to-Noise Ratio of the received GPS signals at the physical layer and employing BDD on the measured data. The approach proposed in [38] detects TSAs by monitoring the correlation between frequency adjustments made to the PMU clock and the changes in phase angle measured by the PMU.

All of the aforementioned works mainly consider the detection and localization of naive TSAs that affect the SE residuals, but they do not consider the detection of TSAs that are constructed to be undetectable by traditional BDD methods [8], [9].

III. System Model

A. Power System and State Estimation Model

We consider a power system with N buses that is observable using a set $\mathcal{M}$ ($|\mathcal{M}| = M$) of voltage and current measurements taken by a set $\mathcal{T}$ ($|\mathcal{T}| = T \leq M$) of PMUs. Let $\mathcal{M}_\tau$ denote the set of measurements taken by PMU $\tau \in \mathcal{T}$. The relation between the measurements and the system state (voltage phasors at each bus) is captured by the linear measurement model

$$\boldsymbol{z} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}, \tag{1}$$

where $\boldsymbol{z} \in \mathbb{C}^M$ are the measurements from PMUs, $\boldsymbol{H} \in \mathbb{C}^{M\times N}$ is the complex measurement matrix, $\boldsymbol{x} \in \mathbb{C}^N$ is the system state and $\boldsymbol{e}$ is white Gaussian measurement noise. The PMU measurements could be nodal voltage phasors, branch current phasors, or nodal current injection phasors. Based on (1), Linear State Estimation (LSE) can be performed to compute a Least Squares (LS) estimate $\hat{\boldsymbol{x}}$ of the system state

$$\hat{\boldsymbol{x}} = (\boldsymbol{H}^\dagger \boldsymbol{H})^{-1}\boldsymbol{H}^\dagger \boldsymbol{z},$$

where $\boldsymbol{H}^\dagger$ is the conjugate transpose of $\boldsymbol{H}$. The estimation residual $\boldsymbol{r} = \boldsymbol{H}\hat{\boldsymbol{x}} - \boldsymbol{z}$ is typically used by BDD algorithms (e.g., the Largest Normalized Residual (LNR) and χ^2 test [39]) to detect faulty data. Faulty data means that the measurements do not match the measurement model (1), and thus leads to a relatively high $\boldsymbol{r}$. Defining the verification matrix $\boldsymbol{F} \in \mathbb{C}^{M\times M}$ as

$$\boldsymbol{F} = \boldsymbol{H}(\boldsymbol{H}^\dagger \boldsymbol{H})^{-1}\boldsymbol{H}^\dagger - \boldsymbol{I}_M$$

allows to express the residual as $\boldsymbol{r} = \boldsymbol{F}\boldsymbol{z}$, where $\boldsymbol{I}_M$ is the $M \times M$ identity matrix.

B. Attack Model

We consider an attacker that is able to manipulate the time reference of a set $\mathcal{P}$ of PMUs ($\mathcal{P} \subseteq \mathcal{T}, |\mathcal{P}| = P$) measuring the set $\mathcal{M}_\mathcal{P}$ of measurements. The dependence of measurements on the attacked PMUs can be captured by the attack-measurement matrix $\boldsymbol{\Psi} \in \{0,1\}^{M\times P}$ such that $\boldsymbol{\Psi}_{m,p} = 1$ if measurement $m \in \mathcal{M}$ is measured by PMU $\tau_p \in \mathcal{P}$ and $\boldsymbol{\Psi}_{m,p} = 0$ otherwise. For each attacked measurement $m \in \mathcal{M}_\mathcal{P}$, the TSA will shift the phase angle of the phasor z_m by

an angle α_m while leaving the phasor magnitude unchanged. Thus the attacked measurement becomes $z_m^a = z_m e^{j\alpha_m}$. Note that $\alpha_m = 0$ if $m \notin \mathcal{M}_\mathcal{P}$. The resulting measurement vector $\boldsymbol{z}^a$ is given by $\boldsymbol{z}^a = (z_1^a, \ldots, z_M^a)^\top = \boldsymbol{z} \odot \boldsymbol{u}$, where $\boldsymbol{u} = (e^{j\alpha_1}, \ldots, e^{j\alpha_M})^\top$ is the attack vector, $\top$ is the transpose operator, and $\odot$ is the Hadamard product. We further define $\boldsymbol{u}_\mathcal{P} \in \mathbb{C}^P$ as the subvector of $\boldsymbol{u}$ including only indices in $\mathcal{M}_\mathcal{P}$.

In practice, a TSA can be implemented by various means depending on the utilized time synchronization mechanism. For SBTS, the attacker can transmit a fake GPS signal that is stronger than the legitimate GPS signal and cause the GPS receiver in the PMU to track the wrong signal, thus potentially losing synchronization. This attack is known in the literature as the GPS spoofing attack [40]. For NBTS, the attacker can manipulate synchronization messages sent to PTP slaves (i.e., PMUs), to include incorrect timing information [4]. Alternatively, an attacker can manipulate a wired network using PTP by inserting a delay box [5], which is a device that can be connected to, e.g., a bidirectional optical fiber connection to create asymmetric delays, by having unequal fiber length in the two directions.

C. Undetectable TSAs

In general, an attacked measurement vector $\boldsymbol{z}^a$ is expected to have a different residual $\boldsymbol{F}\boldsymbol{z}^a \neq \boldsymbol{F}\boldsymbol{z}$, and the different residual may trigger an alarm by BDD algorithms. This is, however, not the case for so called undetectable TSAs.

Definition 1. *A TSA against a set $\mathcal{P}$ of PMUs is undetectable if it does not change the measurement residual, i.e., $\boldsymbol{F}\boldsymbol{z} = \boldsymbol{F}\boldsymbol{z}^a$.*

The necessary and sufficient condition for a TSA to be undetectable can be formulated as follows [8].

Lemma 1. *A TSA against a set $\mathcal{P}$ of PMUs is undetectable if and only if the vector $\boldsymbol{u}_\mathcal{P} \in \mathbb{C}^P$ satisfies*

$$\boldsymbol{W}(\boldsymbol{u}_\mathcal{P} - \mathbf{1}) = \mathbf{0}, \tag{2}$$

where $\boldsymbol{W} = \boldsymbol{\Psi}^T diag(\boldsymbol{z})^\dagger \boldsymbol{F}^\dagger \boldsymbol{F} diag(\boldsymbol{z}) \boldsymbol{\Psi}$ is the complex attack angle matrix, $\boldsymbol{W} \in \mathbb{C}^{P\times P}$, and is Hermitian, and $\mathbf{1}$ and $\mathbf{0}$ are the P-dimensional vectors of ones and zeros, respectively.

When $rank(\boldsymbol{W}) = 1$, undetectable TSAs can easily be computed [9], as only one row of $\boldsymbol{W}$ has to be considered in (2). If $rank(\boldsymbol{W}) > 1$ then approximately undetectable attacks can be computed using a rank-1 approximation of $\boldsymbol{W}$ [8].

In order to quantify the vulnerability of a set $\mathcal{P}$ of PMUs, the singular value decomposition of the corresponding $\boldsymbol{W}$ matrix is examined. Let $\sigma_i(\boldsymbol{W})$ be the i^{th} singular value of $\boldsymbol{W}$. Authors in [8] introduced the Index of Separation (IoS) metric as the ratio of the largest singular value and the sum of singular values of $\boldsymbol{W}$,

$$IoS_\mathcal{P} = \frac{\max_i |\sigma_i(\boldsymbol{W})|}{\sum_{i=1}^{P} |\sigma_i(\boldsymbol{W})|}. \tag{3}$$

The closer the IoS to 1, the closer $rank(\boldsymbol{W})$ is to 1, and the more vulnerable $\mathcal{P}$ is to undetectable TSAs. Note that the IoS value will depend on the measurement vector $\boldsymbol{z}$. To quantify the vulnerability of $\mathcal{P}$ irrespective of $\boldsymbol{z}$, [41] introduced an analogous metric, the Effective Rank Ratio (ERR), as

$$\text{ERR}_\mathcal{P} = \frac{\max_i |\sigma_i(\tilde{\boldsymbol{F}})|}{\sum_{i=1}^{P} |\sigma_i(\tilde{\boldsymbol{F}})|}, \tag{4}$$

where $\tilde{\boldsymbol{F}} \in \mathbb{C}^{M\times|\mathcal{M}_\mathcal{P}|}$ is the submatrix of $\boldsymbol{F}$ including only columns in $\mathcal{M}_\mathcal{P}$. Again, the closer $\text{ERR}_\mathcal{P}$ to 1, the more vulnerable $\mathcal{P}$ is to undetectable TSAs. Previous works utilizing $\text{IoS}_\mathcal{P}$ and $\text{ERR}_\mathcal{P}$ have used threshold values between 0.99 and 0.999 as an indication of vulnerability [13], [41].

D. Problem Formulation

Our objective is to detect and localize TSAs, including undetectable and approximately undetectable TSAs, i.e., TSAs where $IoS_\mathcal{P}$ and $\text{ERR}_\mathcal{P}$ are equal to or close to 1, respectively. Formally, a detector is a function $D : \mathbb{C}^M \rightarrow \{0, 1\}$ that outputs whether a measurement vector $\boldsymbol{z} \in \mathbb{C}^M$ was compromised by a TSA. Simiraily, localization can be represented by set valued function $L : \mathbb{C}^M \rightarrow \{0, 1\}^T$, i.e., it outputs for each PMU $\tau \in \mathcal{T}$ whether or not its time reference has been compromised, and hence the set $\mathcal{P}$ of attacked PMUs. In Section V, we propose detectors D and localization functions L based on a combination of Graph Signal Processing (GSP) and Machine Learning (ML). Before that, we present a brief review of the GSP Framework for the Power Grid.

IV. Rationale for Detection using Graph Signal Processing and Methodological Background

Several previous works dealing with PMU data have observed that the state vector $\boldsymbol{x}$ lies in a lower dimensional subspace [42]–[44]. This observation was leveraged using principal component analysis (PCA) for the recovery of missing data [42], event detection [43] and attack detection [44]. PCA does, however, not leverage knowledge of the structure of the low dimensional subspace. By considering PMU data as signals on a graph, i.e., the power grid, this structure can be made explicit and can be utilized for detecting attacks [11]. To make the argument more formal, we now review concepts from Graph Signal Processing (GSP) [10] that underlie the detectors we propose.

A. GSP Preliminaries

The GSP framework for power systems, introduced in [12], models the power grid as an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where nodes (vertices) $\mathcal{N}$ are buses, $|\mathcal{N}| = N$ and edges $\mathcal{E}$ are transmission lines. A graph signal is defined as a vector of values indexed by the vertices of the graph. In this context, the state vector $\boldsymbol{x} \in \mathbb{C}^N$ is a graph signal with respect to the power grid.

Analogous to the time shift operator considered in discrete-time signal processing (DSP), a graph shift operator (GSO), $\boldsymbol{Y} \in \mathbb{C}^{N\times N}$, is defined for subsequent GSP operations as a

linear operator, which when applied to a graph signal $\boldsymbol{x}$, produces a shifted graph signal $\boldsymbol{Y}\boldsymbol{x}$ whose entries at the vertices are a linear combination of the values at their corresponding neighboring vertices. Typically, the graph adjacency matrix or the graph Laplacian matrix is considered as the GSO [10]. For the power grid, we consider the complex-symmetric bus admittance matrix, $\boldsymbol{Y} \in \mathbb{C}^{N\times N}$ to be the GSO since the bus admittance matrix is equivalent to the weighted graph Laplacian associated with the power grid,

$$[\boldsymbol{Y}]_{ij} = \begin{cases} -y_{ij}, & i \neq j,\ (i,j) \in \mathcal{E} \\ \sum_{k|(i,k)\in\mathcal{E}} & y_{ik} \quad i = j. \end{cases} \tag{5}$$

where y_{ij} is the the admittance of the branch connecting buses i and j. With the help of the GSO, we can define the graph Fourier transform (GFT) and the graph frequencies by considering the eigenvalue decomposition of the GSO,

$$\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top. \tag{6}$$

In the GSP literature, the eigenvector matrix $\boldsymbol{U}$ is defined as the GFT basis. This concept was adopted so as to be analogous to DSP, where the Fourier basis corresponds to the eigenvectors of the Laplacian matrix of a directed circular graph where each node represents a particular time index of a periodic signal. Similarly, the eigenvalues of the GSO on the diagonal of $\boldsymbol{\Lambda}$ are called graph frequencies to be consistent with time-domain where the frequency is derived as a function of eigenvalues of the Laplacian matrix of a directed circular graph [45]. The ordering of graph frequencies is based on the *total variation (TV)* criterion $\|\boldsymbol{Y}\boldsymbol{x}\|_1$, which characterizes the smoothness in a graph signal. If neighboring node values are similar, then the graph signal is considered smooth. The *ascending* order of graph frequency magnitudes, $|\lambda_1| \leq |\lambda_2| \cdots \leq |\lambda_N|$, corresponds to increasing graph frequencies[1]. With that, the GFT of a graph signal $\boldsymbol{x}$ is defined as

$$\tilde{\boldsymbol{x}} = \boldsymbol{U}^\top \boldsymbol{x}, \quad \boldsymbol{x} = \boldsymbol{U}\tilde{\boldsymbol{x}}, \tag{7}$$

and entries in $\tilde{\boldsymbol{x}}$ are the graph Fourier (GF) coefficients.

Next, we review properties of graph filters that we use in attack detection. Like in DSP, a linear shift-invariant graph filter $\mathcal{H}(\boldsymbol{Y})$ has the property that the application of a GSO $\boldsymbol{Y}$ to the input of the filter is the same as applying the shift to the output of the filter, i.e., $\boldsymbol{x} = \mathcal{H}(\boldsymbol{Y})\boldsymbol{s} \Longleftrightarrow \boldsymbol{Y}\boldsymbol{x} = \mathcal{H}(\boldsymbol{Y})\boldsymbol{Y}\boldsymbol{s}$. Such graph filters can be written as a matrix polynomial in the GSO $\boldsymbol{Y}$, i.e.,

$$\mathcal{H}(\boldsymbol{Y}) = \sum_{i=0}^{N-1} h_i \boldsymbol{Y}^i = \boldsymbol{U}\left(\sum_{i=0}^{N-1} h_i \boldsymbol{\Lambda}^i\right)\boldsymbol{U}^\top, \tag{8}$$

where $\{h_i\}_i$ are the filter coefficients. The transfer function of $\mathcal{H}(\boldsymbol{Y})$ is $h(\lambda) = \sum_{i=0}^{N-1} h_i \lambda^i$. The graph filtering operation with input $\boldsymbol{s}$, i.e. $\boldsymbol{x} = \mathcal{H}(\boldsymbol{Y})\boldsymbol{s}$, can also be described in the graph frequency domain by observing from (8) that

$$\boldsymbol{U}^\top \boldsymbol{x} = \boldsymbol{U}^\top \mathcal{H}(\boldsymbol{Y})\boldsymbol{s} = \left(\sum_{i=0}^{N-1} h_i \boldsymbol{\Lambda}^i\right)\boldsymbol{U}^\top \boldsymbol{s} \tag{9}$$

$$\tilde{\boldsymbol{x}} = h(\boldsymbol{\lambda}) \odot \tilde{\boldsymbol{s}}, \tag{10}$$

where $h(\boldsymbol{\lambda})$ is the vectorized transfer function of $\mathcal{H}(\boldsymbol{Y})$ with $[h(\boldsymbol{\lambda})]_i = h(\lambda_i)$ and $\tilde{\boldsymbol{s}}$ is the GFT of the filter input. In (10), the graph frequency components of the filtered graph signal $\tilde{\boldsymbol{x}}$ are the product of the transfer function of the filter and the graph frequency components of the input signal. With the help of the graph filter transfer function $h(\boldsymbol{\lambda})$, analogous to its discrete-time counterpart, we can define low, high, and band-pass graph filters. In this work, we will use low-pass and high-pass graph filters. The ideal low-pass filter with a cutoff graph frequency λ_k is such that $h(\lambda_i) = 1, \forall i < k$ and zero otherwise. The ideal high-pass and band-pass filters are analogously defined.

B. GSP-based Generative Model for the State Vector $\boldsymbol{x}$

GSP can be used for defining a generative model for the voltage phasor state vector $\boldsymbol{x}$ [12]. The generative model stems from rewriting Ohm's law. By considering the vector of current phasors as the input $\boldsymbol{s}$, the state vector can be written as the output of a graph filter,

$$\boldsymbol{x} = \mathcal{H}(\boldsymbol{Y})\boldsymbol{s}, \quad \mathcal{H}(\boldsymbol{Y}) \triangleq \boldsymbol{Y}^{-1},\ h(\lambda) = \lambda^{-1}. \tag{11}$$

Consider now the transfer function $h(\lambda)$, and note that the response is inversely proportional to eigenvalues of $\boldsymbol{Y}$. In practice, most transmission grids are organized as communities with weak connectivity [46] and therefore the system admittance matrix (i.e., the GSO), $\boldsymbol{Y}$, tends to be sparse. Due to this, the condition number of $\boldsymbol{Y}$ is high and thereby the spectrum is such that the magnitude of eigenvalues exhibit a rapid decaying trend when arranged in descending order as shown in [12]. Hence, it is reasonable to consider that $\mathcal{H}(\boldsymbol{Y})$ is an approximate low-pass filter.

The above observation makes it possible to construct an approximate model for the system state. Let $\mathcal{H}_k(\boldsymbol{Y})$ be a *low-pass* filter with cutoff frequency λ_k, with transfer function $h_k(\lambda_i) = \lambda_i^{-1},\ i \leq k$ and zero otherwise. Also, let $\boldsymbol{\Lambda}_k$ be the diagonal matrix of eigenvalues with entries $\lambda_i, i = \{1, 2, \ldots, k\}$. Then, we can write

$$\boldsymbol{x} \approx \mathcal{H}_k(\boldsymbol{Y})\boldsymbol{s} = \boldsymbol{U}\begin{bmatrix} \boldsymbol{\Lambda}_k^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \overbrace{\boldsymbol{U}^\top \boldsymbol{s}}^{\tilde{\boldsymbol{s}}} \approx \boldsymbol{U}_k \boldsymbol{\Lambda}_k^{-1} \tilde{\boldsymbol{s}}_k, \tag{12}$$

where $\boldsymbol{U}_k \in \mathbb{C}^{N\times k}, k < N$ is the submatrix of the GFT basis $\boldsymbol{U}$ or eigenvectors corresponding to the first k graph frequencies and $\tilde{\boldsymbol{s}}_k$ are the first k graph frequency components of the input $\boldsymbol{s}$ (current phasors). From (12), we postulate that the state vector $\boldsymbol{x}$ lies approximately in a low-dimensional subspace spanned by $\boldsymbol{U}_k$. This approximation was empirically verified in [12] and is confirmed in Section VI.

Consequently, $\boldsymbol{x}$ is approximately a *low-pass* graph signal, and the GF coefficients $\tilde{\boldsymbol{x}}$ have negligible magnitude as the

[1]When the eigenvalues are complex-valued, the ordering may not be unique since two distinct eigenvalues may have the same magnitude. Such ties can be broken by comparing either the real or imaginary parts as applicable.

graph frequency increases, i.e.,

$$|\{\tilde{\boldsymbol{x}}\}_i| \ll |\{\tilde{\boldsymbol{x}}\}_k|, \quad k < i \leq N \tag{13}$$

Thus, we can approximate the state vector as

$$\boldsymbol{x} \approx \boldsymbol{U}_k \tilde{\boldsymbol{x}}_k. \text{ where } \tilde{\boldsymbol{x}}_k = \boldsymbol{\Lambda}_k^{-1} \tilde{\boldsymbol{s}}_k, \tag{14}$$

and we can approximate the measurement model (1) as

$$\boldsymbol{z} \approx \boldsymbol{H}_k \tilde{\boldsymbol{x}}_k + \boldsymbol{e}, \tag{15}$$

where $\boldsymbol{H}_k = \boldsymbol{H}\boldsymbol{U}_k$.

Having reviewed the low-pass GSP generative model for the state vector, in the following section, we describe the components of the proposed TSA detection and localization approach using the aforementioned GSP model.

V. TSA Detection and Localization using GSP

Leveraging the GSP based measurement model in (15), the state estimation problem can be solved for the vector $\tilde{\boldsymbol{x}}_k$, which is of lower dimension, instead of $\boldsymbol{x}$. This model restricts the space of feasible measurements $\boldsymbol{z}$ to a lower dimensional subspace compared to (1). The corresponding residuals will thus be

$$\boldsymbol{r}^k = \boldsymbol{F}_k \boldsymbol{z}, \ \boldsymbol{F}_k = \boldsymbol{H}_k (\boldsymbol{H}_k^\dagger \boldsymbol{H}_k)^{-1} \boldsymbol{H}_k^\dagger - \boldsymbol{I}_M, \tag{16}$$

and $\boldsymbol{F}_k \in \mathbb{C}^{M \times M}$ is the GSP-based verification matrix. We can use the GSP based verification matrix to provide an alternative to Definition 1 for the undetectability of TSAs, as follows.

Definition 2. *A TSA against a set $\mathcal{P}$ of PMUs is k-undetectable if and only if it does not change the GSP measurement residual considering k graph frequencies, i.e., $\boldsymbol{F}_k \boldsymbol{z} = \boldsymbol{F}_k \boldsymbol{z}^a$.*

Observe that for $k = N$ the definition above is equivalent to Definition 1[2]. For $k < N$ the two are not equivalent, however, as we show in the following. We start with a technical result concerning the rank of $\boldsymbol{F}_k$.

Lemma 2. *Consider that the system is observable for a set of M measurements, $M \geq N$, i.e., rank$(\boldsymbol{H}) = N$. Then rank$(\boldsymbol{F}_k) >$ rank$(\boldsymbol{F})$.*

Proof. Since $\text{rank}(\boldsymbol{H}) = N$, $\text{rank}(\boldsymbol{F}) = M - N$ as $\boldsymbol{F}$ is a projection matrix for the space orthogonal to the space spanned by $\boldsymbol{H}$. Also, $\text{rank}(\boldsymbol{H}_k) \leq \min(\text{rank}(\boldsymbol{H}), \text{rank}(\boldsymbol{U}_k)) \implies \text{rank}(\boldsymbol{H}_k) \leq k$. Therefore, $\text{rank}(\boldsymbol{F}_k) \geq M - k$. Since $k < N$, $\text{rank}(\boldsymbol{F}_k) > \text{rank}(\boldsymbol{F})$. □

Lemma 2 implies that the dimension of the null-space of $\boldsymbol{F}_k$ is smaller than that of $\boldsymbol{F}$ i.e.

$$\text{rank}(\boldsymbol{F}_k) > \text{rank}(\boldsymbol{F}) \implies \text{nullity}(\boldsymbol{F}_k) < \text{nullity}(\boldsymbol{F}), \tag{17}$$

which means that a solution for $\Delta\boldsymbol{z} = (\boldsymbol{z}^a - \boldsymbol{z})$ that satisfies $\boldsymbol{F}_k \Delta z = 0$ has to lie in a lower-dimensional subspace than a solution that satisfies $\boldsymbol{F}\Delta z = 0$. Consequently, by choosing a small enough value of k the dimension of the solution can be brought to zero (i.e. $\text{rank}(\boldsymbol{F}_k) = M \implies \text{nullity}(\boldsymbol{F}_k) = 0$). This means that even if the attacker knows the value of k chosen by the operator, they could be constrained by the dimensionality of the solution $\Delta\boldsymbol{z}$.

Our next result shows that Definitions 1 and 2 are indeed not equivalent, as k-undetectability implies undetectability, but the converse need not be true.

Proposition 1. *k-undetectability implies undetectability, i.e., $\boldsymbol{F}_k \Delta\boldsymbol{z} = 0 \implies \boldsymbol{F}\Delta\boldsymbol{z} = 0$, but $\boldsymbol{F}\Delta\boldsymbol{z} = 0 \not\Longrightarrow \boldsymbol{F}_k\Delta\boldsymbol{z} = 0$.*

Proof. Assume that $\boldsymbol{F}_k\Delta\boldsymbol{z} = 0 \implies \Delta\boldsymbol{z} = \boldsymbol{H}_k\boldsymbol{c}$ for some state $\boldsymbol{c}$. However,

$$\Delta\boldsymbol{z} = \boldsymbol{H}_k\boldsymbol{c} = \boldsymbol{H}\underbrace{\boldsymbol{U}_k\boldsymbol{c}}_{\triangleq \boldsymbol{c}_k} \implies \boldsymbol{F}\Delta\boldsymbol{z} = 0. \tag{18}$$

Thus, $\boldsymbol{F}_k\Delta\boldsymbol{z} = 0 \implies \boldsymbol{F}\Delta\boldsymbol{z} = 0$. Now, let $\boldsymbol{F}\Delta\boldsymbol{z} = 0$. This means that there is a modified state $\boldsymbol{x}^a$ such that $\Delta\boldsymbol{z} = \boldsymbol{H}\boldsymbol{x}^a$. Let us use the GFT basis from (7) for rewriting the modified state,

$$\boldsymbol{x}^a = \boldsymbol{U}\tilde{\boldsymbol{x}^a} = \boldsymbol{U}_k\tilde{\boldsymbol{x}^a}_k + \boldsymbol{U}_{N-k}\tilde{\boldsymbol{x}^a}_{N-k}, \tag{19}$$

where $\boldsymbol{U}_{N-k}$ are columns of $\boldsymbol{U}$ corresponding to $(k+1)^{th}$ graph frequency or eigenvalue and above. Thus we obtain

$$\Delta\boldsymbol{z} = \underbrace{\boldsymbol{H}\boldsymbol{U}_k}_{=\boldsymbol{H}_k}\tilde{\boldsymbol{x}^a}_k + \boldsymbol{H}\boldsymbol{U}_{N-k}\tilde{\boldsymbol{x}^a}_{N-k}. \tag{20}$$

Multiplying both sides by $\boldsymbol{F}_k$ we get

$$\boldsymbol{F}_k\Delta\boldsymbol{z} = 0 + \boldsymbol{F}_k\boldsymbol{H}\boldsymbol{U}_{N-k}\tilde{\boldsymbol{x}^a}_{N-k}. \tag{21}$$

Observe that the term $\boldsymbol{F}_k\boldsymbol{H}\boldsymbol{U}_{N-k}\tilde{\boldsymbol{x}^a}_{N-k}$ need not be zero, and thus $\boldsymbol{F}\Delta\boldsymbol{z} = 0 \not\Longrightarrow \boldsymbol{F}_k\Delta\boldsymbol{z} = 0$, which concludes the proof. □

From the proposition above, one can also conclude that undetectable TSAs against $\boldsymbol{F}_k$ form a subset of those against $\boldsymbol{F}_{k+1}$. If the attacker is aware of the utilization of the GSP framework for detection, it needs to ensure k^--undetectability for the attack to be undetectable when the GSP-based verification matrix $\boldsymbol{F}_k$ is used for $k^- \leq k$. However, Lemma 2 implies that the attacker is more constrained in finding attack vectors as the value of k decreases, due to dimensionality reduction of feasible solutions for $\Delta\boldsymbol{z}$. Thus, smaller k is desirable to constrain the attacker. However, from (14), we see that a smaller k translates to higher model approximation error. Therefore the operator has to choose k by striking a balance between the two aspects.

A. Detection of TSA using GSP Framework

We now propose four GSP-based methods to detect TSAs. All proposed methods output a detection score $\hat{\mathcal{D}}$ that corresponds to the certainty of a TSA against the measured data. The interrelation between the proposed detection methods is shown in Figure 1

1) GSP Residuals (GSP-R): This method uses the GSP-based verification matrix $\boldsymbol{F}_k$ to obtain the GSP residual $\boldsymbol{r}^k$ using (16). The detection score is then computed as the largest magnitude of elements in $\boldsymbol{r}^k$, i.e., $\hat{\mathcal{D}} = \max_{i \in \mathcal{M}} |\boldsymbol{r}_i^k|$.

[2]It is equivalent if $\boldsymbol{U}$ is an invertible matrix. This is true for most power grid topologies since the admittance matrix $\boldsymbol{Y}$ is generally diagonalizable.

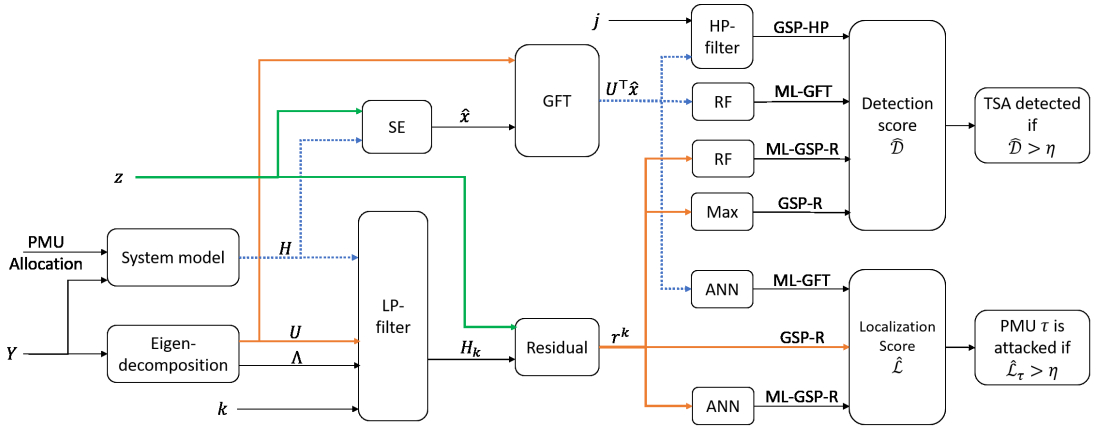


Fig. 1: A block diagram showing the interrelation between the proposed detection and localization methods.

2) GSP-based High-pass Filter (GSP-HP): Since the true state vector $\boldsymbol{x}$ is approximately low-pass, TSAs can be detected if the calculated energy of the estimated state after passing it through a high-pass graph filter is high. Consider the estimated state using the attacked measurements $\boldsymbol{z}^a$,

$$\hat{\boldsymbol{x}}^a = (\boldsymbol{H}^\dagger \boldsymbol{H})^{-1}\boldsymbol{H}^\dagger \boldsymbol{z}^a, \quad \hat{\boldsymbol{x}}^a \in \mathbb{C}^N \tag{22}$$

Define an ideal high-pass filter with cut-off frequency index $j < N$, $\mathcal{H}_{\text{HPF}}(\boldsymbol{Y})$ so that $h_{\text{HPF}}(\lambda_i) = 1, i > j$ and zero otherwise. Let $\hat{\boldsymbol{x}}^a$ be input to $\mathcal{H}_{\text{HPF}}(\boldsymbol{Y})$. Then,

$$\boldsymbol{d}^a \triangleq \mathcal{H}_{\text{HPF}}(\boldsymbol{Y})\hat{\boldsymbol{x}}^a, \quad \tilde{\boldsymbol{d}}^a = \boldsymbol{h}_{\text{HPF}} \odot (\boldsymbol{U}^\top \hat{\boldsymbol{x}}^a), \tag{23}$$

where $\tilde{\boldsymbol{d}}^a$ is the graph frequency response of $\boldsymbol{d}^a$. Since the high-pass filter is ideal, $\tilde{\boldsymbol{d}}^a$ contains the high-frequency ($i > j$) coefficients of $\boldsymbol{U}^\top \hat{\boldsymbol{x}}^a$. Consider now the test metric or detection score

$$\hat{\mathcal{D}} \triangleq \|\boldsymbol{h}_{\text{HPF}} \odot \boldsymbol{U}^\top \hat{\boldsymbol{x}}^a\|_2^2 \tag{24}$$

$$= \|\boldsymbol{h}_{\text{HPF}} \odot \boldsymbol{U}^\top (\boldsymbol{H}^\dagger \boldsymbol{H})^{-1}\boldsymbol{H}^\dagger \boldsymbol{z}\|_2^2. \tag{25}$$

Since $\boldsymbol{x}$ is an approximately low-pass graph signal, $\hat{\mathcal{D}}$ is lower when there is no attack as compared to when measurements are attacked.

3) Machine Learning using GSP Residuals (ML-GSP-R): Despite the evident advantage of using the GSP residuals $\boldsymbol{r}^k$ computed using (16) and graph frequency components $\boldsymbol{U}^\top \hat{\boldsymbol{x}}$ over traditional LSE residuals $\boldsymbol{r}$ in detecting TSAs, changes in the GSP metrics may often be too small to raise an alarm. As a solution to this problem, we propose to use a supervised machine learning model for attack detection. We performed initial experiments with Artificial Neural Network (ANN) classifiers and Random Forest (RF) classifiers [47]. We found that ANNs did not outperform RFs, and we thus opted for RF due to its shorter training time and its robustness to over-fitting. Given feature vectors $\boldsymbol{\mathcal{F}} \in \mathbb{R}^{2M}$, i.e., the real and imaginary parts of the GSP resdiduals $\boldsymbol{r}^k$, computed using (16), and corresponding ground-truth detection labels $\mathcal{D} \in \{0, 1\}$, the random forest is trained to produce a detection score $\hat{\mathcal{D}} \in [0, 1]$, s.t., $\hat{\mathcal{D}} \approx \mathcal{D}$.

4) Machine Learning using Graph Fourier Transform (ML-GFT): Similar to ML-GSP-R, the detection is done using a RF classifier, but with the real and imaginary parts of the GFT of the state estimate, i.e., $\boldsymbol{U}^\top \hat{\boldsymbol{x}}$, hence the input features $\boldsymbol{\mathcal{F}} \in \mathbb{R}^{2N}$.

B. Localization of TSA using GSP Framework

We propose three TSA localization methods based on GSP. Similar to the TSA detection methods, the proposed TSA localization methods output a localization score $\hat{\boldsymbol{\mathcal{L}}}_\tau$ for each measurement / PMU τ reflecting the certainty that τ is attacked. The interrelation between the proposed localization methods is shown in Figure 1

1) GSP-R: The localization score $\hat{\boldsymbol{\mathcal{L}}} \in \mathbb{R}^M$ is computed as the magnitude of the GSP residuals, i.e., $\hat{\boldsymbol{\mathcal{L}}}_\tau = |\boldsymbol{r}_i^k|$.

2) ML-GSP-R: The localization score is computed by an ANN, which takes as input features $\boldsymbol{\mathcal{F}} \in \mathbb{R}^{2M}$ the real and imaginary parts of the GSP residual $\boldsymbol{r}^k$ computed using (16), and is trained to perform multi-label classification with labels $\boldsymbol{\mathcal{L}} \in \{0, 1\}^T$ and scores $\hat{\boldsymbol{\mathcal{L}}} \in [0, 1]^T$. For a PMU $\tau \in \{1, \ldots, T\}$, $\boldsymbol{\mathcal{L}}$ indicates whether τ is attacked ($\boldsymbol{\mathcal{L}}_\tau = 1$) or not attacked ($\boldsymbol{\mathcal{L}}_\tau = 0$). The model is trained to produce $\hat{\boldsymbol{\mathcal{L}}} \approx \boldsymbol{\mathcal{L}}$.

3) ML-GFT: Similar to ML-GSP-R, the localization is done using an ANN, but using as input features the real and imaginary parts of the GFT of the state estimate, i.e., $\boldsymbol{U}^\top \hat{\boldsymbol{x}}$, hence the input features $\boldsymbol{\mathcal{F}} \in \mathbb{R}^{2N}$. The labels are the same as for ML-GSP-R.

VI. Numerical Results

In this section we evaluate the proposed detection and localization algorithms, and compare them to the state-of-the-art methods in the field. All simulations were carried out on a notebook with Intel Core i7-8550 CPU @ 1.8 GHz and 16 GB of RAM.

A. Evaluation Methodology

We considered two IEEE benchmark power systems for the evaluation of the proposed TSA detection and localization approaches, namely the IEEE 14-bus system ($N = 14$) and the IEEE 39-bus system ($N = 39$).

For the IEEE 14-bus system, we considered a PMU allocation with $M = 16$ measurements taken by $T = 10$ PMUs; 8 nodal voltage measurements at buses {1, 2, 4, 7, 8, 11, 12, 14}, and 8 current injection measurements at buses {2, 4, 6, 7, 8, 9, 11, 12}.

For the IEEE 39-bus system, we considered a PMU allocation with $M = 47$ measurements taken by $T = 33$ PMUs; 24 nodal voltage measurements at buses {1, 6, 8, 12, 13, 14, 15, 16, 17, 19, 20, 23, 24, 25, 28, 30, 31, 32, 33, 34, 35, 37, 38, 39} and 23 current injection measurements at buses {2, 3, 5, 6, 8, 9, 10, 13, 16, 17, 19, 21, 22, 23, 24, 25, 27, 28, 34, 35, 36, 37, 38}. For both systems, the considered PMU allocations ensure the observability of the system, i.e., $rank(\boldsymbol{H}) = N$. Moreover, no single measurement is allowed to be a critical measurement, i.e., the removal of any single measurement will not render the system unobservable. Formally, this last constraint means that the removal of any row of $\boldsymbol{H}$ will not decrease its rank. Moreover, we assume that voltage and current measurements taken by the same PMU share the same time reference. Thus, an attack targeting such a PMU will shift the phase angle of both measurements with the same attack angle α.

To model the two systems with GSP, the values of the cut-off frequency index k were chosen to achieve the best separation between attacked and non-attacked measurements in our datasets, as shown in Figure 2. The procedure for generating the dataset is shown later. In the figure, the average ℓ_2-norm of the GSP residual over the datasets is shown as a function of the parameter k. As seen from the figure, most values of k achieve good separation between attacked and non-attacked samples in the IEEE 14-bus system. We choose the value of $k = 8$ as an intermediate value. On the contrary, for the IEEE 39-bus system, only $32 \leq k \leq 38$ can achieve acceptable separation between the attacked and non-attacked measurements. Therefore, we choose the value $k = 35$.

To characterize the vulnerability of the considered allocations, we computed the vulnerability metric ERR for all possible combinations of triplets ($P = 3$) of PMUs based on $\boldsymbol{F}$ and based on $\boldsymbol{F}_k$. Figure 3 shows the distribution of ERR based on $\boldsymbol{F}$ and $\boldsymbol{F}_k$. It allows us to conclude that the ERR values computed with respect to $\boldsymbol{F}_k$ are significantly lower, suggesting the GSP-based residual $\boldsymbol{F}_k\boldsymbol{z}$ is more likely to change significantly due to a TSA. This is also aligned with the implications of Lemma 2 and Proposition 1 which characterize the difficulty an attacker has in constructing attacks when $\boldsymbol{F}_k$ is used as the verification matrix.

To generate PMU measurements for both the power systems, we used MATPOWER [48] base load power injections for the IEEE 14-bus and IEEE 39-bus cases and modified the power injections at a number of load buses in order to simulate a variety of system states. The number of modified load buses followed a discrete uniform distribution with end points 0 and N, and the injected powers, both real and reactive, were changed by scaling the base load power by a factor sampled from a uniform distribution with end points 0.5 and 1.5. Then, through a load flow analysis, the system state x was recomputed. The voltage and current injection measurements were generated according to the measurement model $\boldsymbol{H}$ and noise was added to the phasors according to the level of noise of 0.1-class PMUs. For each power system, we generated a non-attacked dataset containing 15000 measurement samples. Next, the following procedure was adopted to simulate TSAs. For each measurement sample in the non-attacked dataset, we randomly sampled a set $\mathcal{P}$ of $P = 3$ PMUs to attack from the T available PMUs, resulting in $3 \leq |\mathcal{M}_{\mathcal{P}}| \leq 6$ attacked measurements. Note that attacks on $P > 3$ PMUs have been successfully demonstrated in previous work [9] but we use $P = 3$ in this work for simplicity. Also, the proposed TSA detection and localization methods are independent of the number of attacked PMUs P.

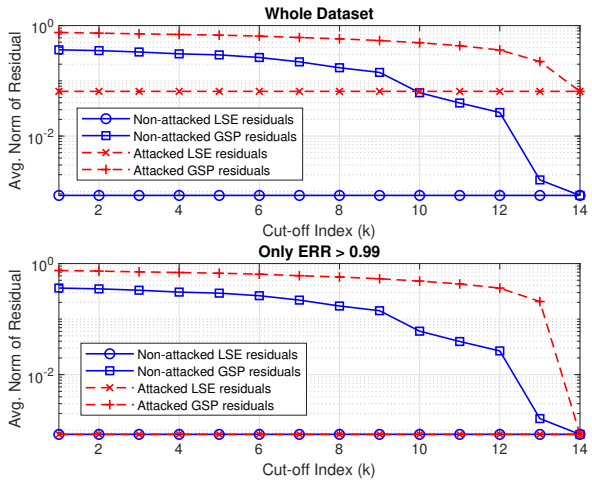
	ERR			
	IEEE 14-bus		**IEEE 39-bus**	
	$\leq \mathbf{0.99}$	$> \mathbf{0.99}$	$\leq \mathbf{0.99}$	$> \mathbf{0.99}$
$\mathbf{I \geq 100}$	606	655	1293	121
$\mathbf{10 \leq I < 100}$	1026	713	1837	209
$\mathbf{I < 10}$	2423	2077	3010	1030

TABLE I: Number of sampled TSAs categorized based on the vulnerability index ERR and the attack impact.

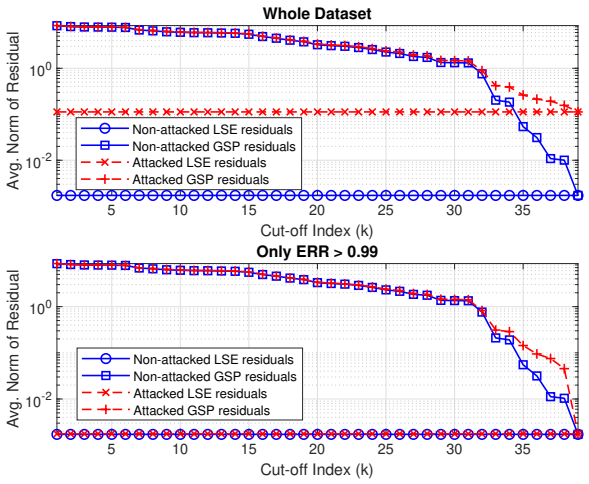
To generate a TSA for $\mathcal{P}$ and a measurement sample, we compute the corresponding $\boldsymbol{\Psi}$ and $\boldsymbol{W}$ matrices, and choose the three attack angles $(\alpha_1, \alpha_2, \alpha_3)$ of the TSA as follows. First, the feasible range for the angle α_3 is identified from the rank-1 approximation of $\boldsymbol{W}$ according to Proposition 1 in [9]. The attack angle α_3 is then chosen from a uniform distribution over the feasible range. Next, the angles (α_1, α_2) are computed for the chosen α_3 according to Proposition 2 in [9]. In case two tuples (α_1, α_2) are possible, one tuple is chosen at random. We used this procedure for generating an attacked dataset containing 15000 samples. The non-attacked and the attacked datasets were each split into training and test datasets, each consisting of 7500 samples.

To facilitate analysis, both the ERR metric and the attack impact were recorded for each sample, To compute the attack impact for each data sample, we first computed the estimated system state based on both the attacked and the non-attacked measurements. Using the estimated state, the apparent power flow on each transmission line can be computed using the power flow equations [39]. Next, the absolute difference between the apparent power flow with and without the TSA is computed for each transmission line. The maximum of those differences across all transmission lines is then defined as the impact of the TSA. Simulations performed using the total phase angle shift as attack impact led to similar results as the ones reported in the paper, and we omit them for brevity. Table I shows the distribution of attacks based on the ERR and the attack impact (I) in the test dataset.

For both non-attacked and attacked data samples, we then computed the measurement residuals $\boldsymbol{r} = \boldsymbol{F}\boldsymbol{z}$ and the GSP residuals $\boldsymbol{r}^k = \boldsymbol{F}_k\boldsymbol{z}$ (again, with $k = 8$ and $k = 35$ for the IEEE 14-bus and 39-bus systems, respectively). Furthermore, we computed the GFT of the system state, i.e. $\boldsymbol{U}^\top\hat{\boldsymbol{x}}$. Figure 4 shows the magnitude of $\boldsymbol{U}^\top\hat{\boldsymbol{x}}$ for the IEEE 14-bus and the IEEE 39-bus systems, for both non-attacked and attacked data



(a) IEEE 14-bus system



(b) IEEE 39-bus system

Fig. 2: LSE and GSP residual vs. the cut-off frequency index k for IEEE 14 and 39 bus test cases.

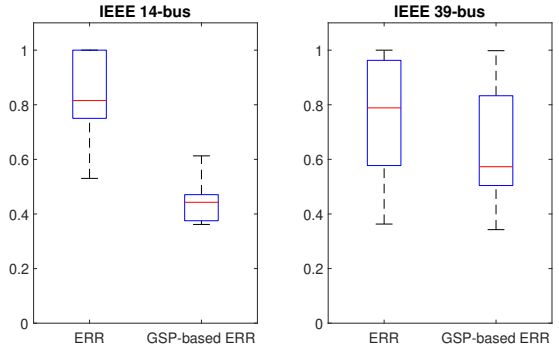


Fig. 3: ERR and the GSP-based ERR vulnerability metrics for triplets of PMUs in the IEEE 14-bus and the IEEE 39-bus systems. GSP-based ERR has significantly lower values, which indicates less vulnerable sets of PMUs.

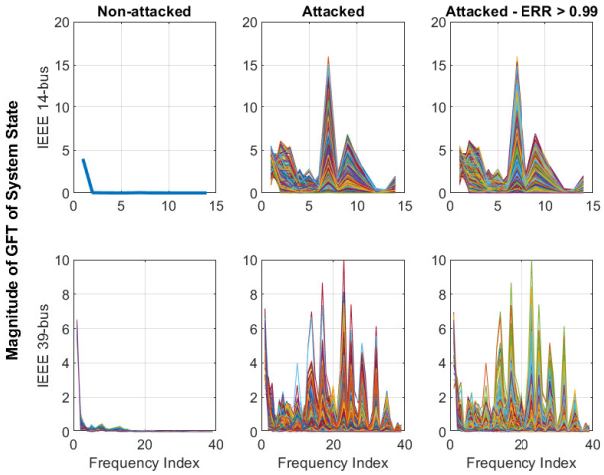


Fig. 4: Absolute values of the GFT of the estimated state, $\boldsymbol{U}^\top \hat{\boldsymbol{x}}$ for non-attacked and attacked measurements. Note the clear effect of TSAs as a change in the GFT of the estimate state, particularly the presence of high frequency components.

samples. It can be observed that the attacked samples show large magnitude of coefficients corresponding to the higher graph frequencies even when ERR is very high, suggesting that they can be easily distinguished from non-attacked samples using a high-pass graph filter.

To illustrate that the task of detecting TSAs is harder than that of detecting FDIAs, we implemented the FDIA detection method proposed in [12], which uses the ℓ_2-norm of the GSP residual vector r^k as the detection score. Figure 5 shows the distribution of the above mentioned metric for FDIAs and TSAs. To have a fair comparison, we simulated FDIAs and TSAs targeting the same number of PMUs (i.e., three PMUs). The FDIAs were simulated by manipulating the measurements as $\boldsymbol{z}^{FDIA} = \boldsymbol{z} + \boldsymbol{a}$ s.t. $\boldsymbol{a} = \boldsymbol{H}\boldsymbol{c}$ for some $\boldsymbol{c} \in \mathbb{C}^N$ for each sample in the dataset [15], [16], where $|\boldsymbol{a}|$ followed a uniform distribution (i.e., $|\boldsymbol{a}| \sim \mathcal{U}(1, 100)$). The figure shows that TSAs generate lower residuals than FDIAs, making them harder to distinguish from non-attacked measurements.

B. Performance Metrics: AUC and FNR

Given the ground truth labels $\boldsymbol{\mathcal{L}}_\tau$ for localization, the localization score $\hat{\boldsymbol{\mathcal{L}}}_\tau$, and a set threshold η, each decision can result in one of four outcomes:

1) True Positive: if $\boldsymbol{\mathcal{L}}_\tau = 1$ and $\hat{\boldsymbol{\mathcal{L}}}_\tau > \eta$
2) False Positive: if $\boldsymbol{\mathcal{L}}_\tau = 0$ and $\hat{\boldsymbol{\mathcal{L}}}_\tau > \eta$
3) True Negative: if $\boldsymbol{\mathcal{L}}_\tau = 0$ and $\hat{\boldsymbol{\mathcal{L}}}_\tau \leq \eta$
4) False Negative: if $\boldsymbol{\mathcal{L}}_\tau = 1$ and $\hat{\boldsymbol{\mathcal{L}}}_\tau \leq \eta$

Note that these definitions can easily be applied for the TSA detection problem, by substituting $\boldsymbol{\mathcal{L}}_\tau$ by $\mathcal{D}$ and $\hat{\boldsymbol{\mathcal{L}}}_\tau$ by $\hat{\mathcal{D}}$. Let TP, FP, TN, and FN denote the number of true positives, false positives, true negatives, and false negatives in the dataset, respectively. For both detection and localization, We can then compute the True Positive Rate (TPR) and the False Positive

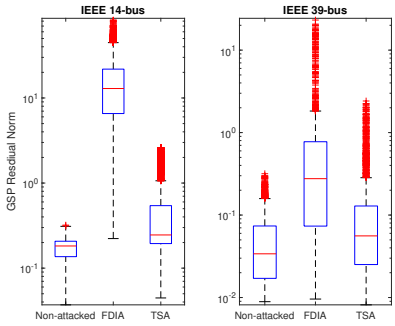


Fig. 5: GSP residual of FDIAs and TSAs for the considered PMU allocations in the IEEE 14-bus and the IEEE 39-bus systems. TSAs have significantly lower residual values making them harder to detect via simple metrics.

Rate (FPR),

$$\text{TPR} = \frac{\text{TP}}{\text{TP}+\text{FN}},$$
$$\text{FPR} = \frac{\text{FP}}{\text{TN}+\text{FP}}.$$

The TPR is the ratio of the number of attacked measurements that are detected by a localization method to the total number of attacked measurements, while the FPR is the ratio of false alarms raised by the localization method to the total number of non-attacked measurements.

Using the above quantities, we use two performance metrics for evaluating the proposed detection and localization solutions. The first metric is the Area Under the Curve (AUC), which is a widely used metric for evaluating detection models. The AUC is computed based on the Receiver Operating Characteristic (ROC) curve, which captures the trade-off between the TPR and the FPR, and is obtained by varying the detection threshold η across its domain.

The second metric is the FNR (i.e., $1-$TPR) for a particular FPR rate, and aims to reflect the practical usefulness of the algorithms. Due to the high streaming rates of PMUs, we considered the desired FPR to be very low (less than the reciprocal of the number of samples in the dataset). Unfortunately, finding the corresponding FNR for such low FNR is not a straightforward task. Therefore, we computed the best quadratic curve that fits the ROC curve for the interval $0 \leq \text{FPR} \leq 0.1$ and used the fitted curve to get the FNR at the required FPR.

C. TSA Detection

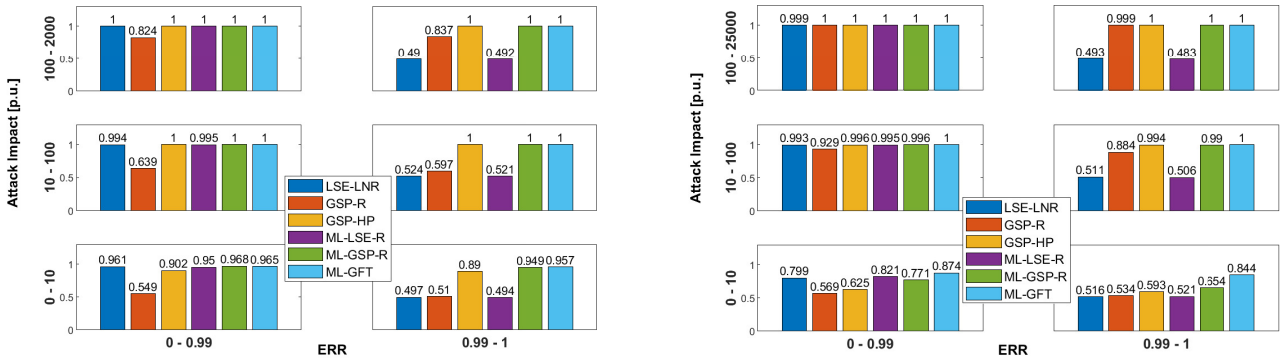
We first present results for the detection methods proposed in Section V-A. To compute $\hat{\mathcal{D}}$ in (24) for the GSP-HP method, we utilized cutoff frequencies of $j = 4$ and $j = 10$ for the high-pass graph filters for detection in the IEEE 14-bus and the IEEE 39-bus systems, respectively. In other words, we excluded the lowest 4 and 10 graph frequency components from $\boldsymbol{U}^\top \hat{\boldsymbol{x}}$. For the ML-based methods, the RF classifier consisted of 20 trees, with a maximum allowed tree depth of 10. We compared our methods to two state-of-the-art baseline detectors. The first baseline is the Largest Normalized Residual (LNR) test [39], where the maximum LSE residual over all measurements is used as the detection score $\hat{\mathcal{D}} = \max_{i\in\mathcal{M}} |\boldsymbol{r}_i^N|$, where $\boldsymbol{r}_i^N$ is the normalized residual obtained by dividing the residual of measurement i by its standard deviation. We refer to this as *LSE-LNR*. The second baseline is an RF classifier applied to the LSE residuals $\boldsymbol{r}$ as input features $\boldsymbol{\mathcal{F}}$. We refer to this as *ML-LSE-R*.

Figure 6 shows the AUC obtained by applying the six detection methods on the datasets for the IEEE 14-bus system (Figure 6a) and the IEEE 39-bus system (Figure 6b). For the sake of clarity, each subplot shows the results for a certain range of ERR and the attack impact. For each subplot, the ROC curve was computed using a set of samples, including the attacked samples in this category (ERR, Impact), as well as an equally-sized random subset of the non-attacked data. The figures clearly show that the detection performance of all methods is positively correlated with the attack impact. Moreover, the detection performance is negatively correlated with the ERR. However, the effect of the ERR on the detection performance is different for different methods. While LSE-LNR (and hence ML-LSE-R) work very well for relatively low ERR values, they fail for high ERRs. This was expected since a TSA with high ERR will not significantly change the LSE residual. On the contrary, the GSP-based methods (and their ML equivalents) perform almost equally well irrespective of the ERR. Interestingly, the GSP-HP can almost perfectly detect TSAs regardless of the ERR when the attack impact is fairly high (above 10 p.u.). However, it performs poorly for attacks with smaller impacts. For such subtle attacks, the ML methods based on GSP features (i.e., ML-GSP-R and ML-GFT) can achieve significantly better performance.

Figure 7 shows the FNR achieved by the detection methods for the IEEE 39-bus system (Results for the IEEE 14-bus system are excluded for brevity). The figure shows the fraction of attacks that will bypass detection when the system operator sets the FPR to be one false alarm per day (FNR = $1/(50{*}60{*}60{*}24) = 2.315{*}10^{-7}$ assuming 50 measurements per second). For ML methods trained on GSP features, the achieved FNR is very low except for TSAs with very little impact. In fact, more that 94% (FNR < 0.06) of TSAs with impact greater than 100 p.u. and more than 80% (FNR < 0.2) of TSAs with impact greater than 10 p.u. can be detected by the proposed methods, even when using such an extremely low FPR. This holds true even for TSAs targeting PMUs with high ERR. Compared to LSE-LNR and ML-LSE-R which can detect at most 3% (FNR > 0.97) of attacks with impact greater than 10 p.u. and high ERR, the proposed methods can detect at least 77% more attacks in this category. The figure shows however that the GSP-based detection methods without ML are not as efficient, e.g., GSP-HP can only detect around 50% of TSAs with impact higher that 10 p.u.. This low value of TPR is mainly due to the extremely low value of the desired FPR. Raising the desired FPR might improve the detection performance, but may lead to alarm fatigue.

D. TSA Localization

We now turn to evaluating the localization performance. To implement the proposed ML-based localization methods in



(a) (b)

Fig. 6: The relation between the Area under the ROC curve for the considered TSA detection methods, the vulnerability of the attacked PMUs, and the attack impact, for the (a) IEEE 14-bus system and (b) the IEEE 39-bus system. The proposed methods outperform state-of-the-art methods, performing equally good irrespective of ERR.

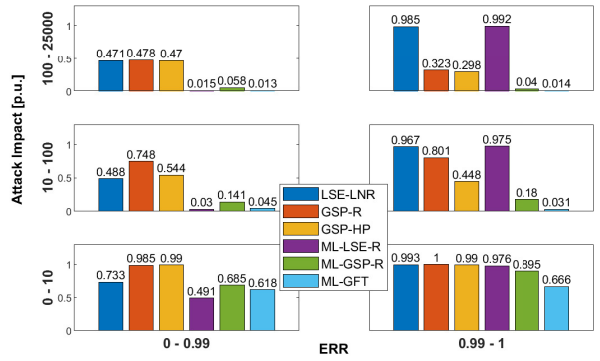
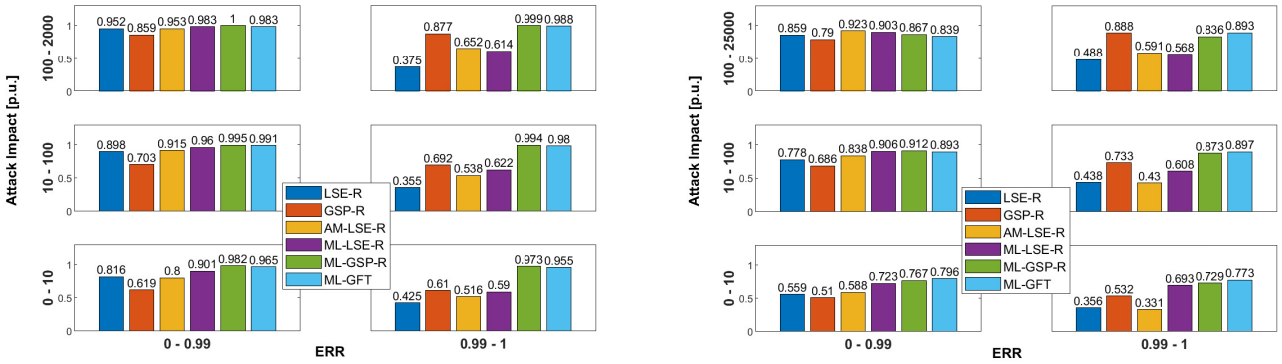


Fig. 7: The relation between the FNR (at one false positive per day) for the considered TSA detection methods, the vulnerability of the attacked PMUs, and the attack impact, for the IEEE 39-bus system. The proposed methods are able to detect more attacks, especially when ERR is high and when the TSA has non-negligible impact.

Section V-B, the ANN was implemented using the Pytorch Python library [49]. The dimension of the ANN input layer depended on the considered method ($2M$ for ML-GSP-R and $2N$ for ML-GFT). The ANN consisted of one hidden layer of 200 neurons using ReLU as activation function, and an output layer of T neurons using Sigmoid as activation function. We used three state-of-the-art methods as a baseline for the evaluation. The first baseline computes the localization score $\hat{\boldsymbol{\mathcal{L}}} \in \mathbb{R}^M$ as the magnitude of the normalized LSE residual, i.e., $\hat{\boldsymbol{\mathcal{L}}}_\tau = |\boldsymbol{r}_\tau^N|$. We refer to this as *LSE-R*. The second baseline computes the localization score $\hat{\boldsymbol{\mathcal{L}}} \in \mathbb{R}^M$ via an alternating minimization algorithm as proposed in [33]. The algorithm (Alg. 3 in [33]) recovers the state vector $\boldsymbol{x}$ and the attack vector $\boldsymbol{u}$ by alternating between their estimation while keeping one of them constant. We refer to this as Alternating Minimization based on the LSE Residuals *(AM-LSE-R)*. The localization score is given by the magnitude of $\boldsymbol{u} - \mathbf{1}$, since a value of 1 at a certain measurement index of $\boldsymbol{u}$ indicates no attack, i.e., $\hat{\boldsymbol{\mathcal{L}}} = |\boldsymbol{u} - \mathbf{1}|$. The third baseline uses an ANN multi-label classifier based on the LSE residuals $\boldsymbol{r}$ as input features $\boldsymbol{\mathcal{F}}$. We refer to this as *ML-LSE-R*.

Figure 8 shows the AUC obtained by applying the six localization methods on the datasets for the IEEE 14-bus system (Figure 8a) and the IEEE 39-bus system (Figure 8b). Similar to Figure 6, each subplot shows the results for a certain range of the ERR and the attack impact. Comparing to Figure 6, we observe that the AUC values for localization are in general slightly lower than for detection. This is expected, as localization is significantly more challenging due to the number of decisions involved (notice that the performance on the IEEE 39-bus is therefore slightly inferior to the IEEE 14-bus system). Comparing the different TSA localization methods, the relative performance is very similar to that in Figure 6. First, the two state-of-the-art methods (LSE-R and AM-LSE-R) show very similar performance since they both rely on LSE. Even though their performance is acceptable when ERR is relatively low, they fail in localizing the TSAs when the ERR is high. The proposed GSP-based methods perform consistently better, and in general we can observe that using ML significantly improves the accuracy of localization, even for the cases when the TSA impact is relatively low and the ERR is relatively high.

Finally, Figure 9 shows the FNR achieved by the localization methods for the IEEE 39-bus system. Unlike TSA detection, the shown values are the fraction of attacked measurements / PMUs that will be declared as non-attacked when the system operator sets the false alarm rate (i.e., FPR) to 5% (that is, 5% of non-attacked measurements are deemed to be attacked). Observe that we set the desired FPR to be higher than in Section VI-C since TSA localization is not needed unless a TSA is detected by a detection algorithm. Therefore, allowing a higher FPR should not lead to alarm fatigue. Focusing on the proposed methods (ML trained on GSP features), the achieved FNR is very low except for TSAs with very little impact. For example, at least 79% of attacked



(a) (b)

Fig. 8: The relation between the Area under the ROC curve for the considered TSA localization methods, the vulnerability of the attacked PMUs, and the attack impact, for the (a) IEEE 14-bus system and (b) the IEEE 39-bus system. The proposed methods yield higher AUC values compared to state-of-the-art methods, especially for high values of ERR.

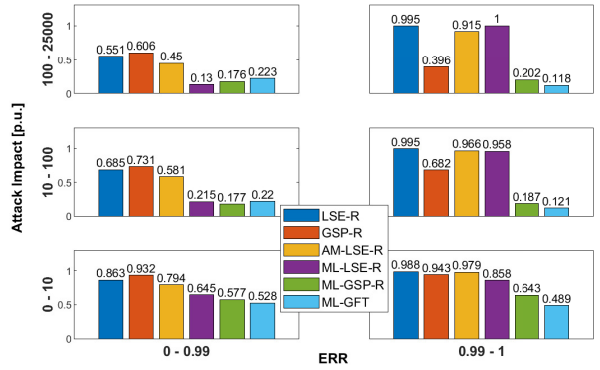


Fig. 9: The relation between the FNR (at 5% False positive rate) for the considered TSA localization methods, the vulnerability of the attacked PMUs, and the attack impact, for the IEEE 39-bus system. The proposed localization methods detect more attacks than the state-of-the-art, especially for higher ERR values.

PMUs will be correctly detected (FNR $<$ 0.21) when the attack impact is at least 10 p.u. and ERR is high, when allowing a 5% false alarm rate. Compared to the baseline methods which can detect at most 9% (FNR $>$ 0.91) of attacked PMUs for the same category, the proposed methods can detect at least 70% more attacked PMUs.

Overall, we can conclude that the proposed ML methods based on GSP show excellent performance in detecting TSAs, and they perform significantly better in localization of the TSAs compared to state-of-the-art methods based on residuals.

E. Evaluation on Large Benchmark Systems

To investigate the scalability of the proposed detection and localization methods, we conducted experiments on three large systems, namely the IEEE 118-bus, the IEEE 300-bus, and the 2383-bus with winter peak (Polish) benchmark systems. We placed the minimum number of PMUs needed for observability by following the optimal placement procedure proposed in [50], as opposed to the random placement used for the smaller systems. For the IEEE 118-bus system, $T = 32$ PMUs were deployed making a total of $M = 158$ measurements for observability. Each PMU measures the bus voltage phasor as well as all incident branch current phasors. We then used $k = 95$ to compute the GSP residuals, and the value $j = 10$ for the high-pass graph filter. The results in Figure 10a show excellent detection performance for the proposed TSA detection methods.

For TSA localization (Figure 10b), even though the proposed localization methods clearly outperform the state-of-the-art methods, the results suggest that the performance of the proposed methods might be less satisfactory for large power systems, especially for high ERR. This is due to that the number of choices involved in localization increases linearly with the size of the system, making it challenging to create appropriate training data sets for large systems. To alleviate this problem, one could partition the system into multiple interconnected areas, as is done for example in distributed state estimation [51], [52], and implement the GSP-based localization methods in each partition. For each area, one needs to consider the corresponding part of the state vector, and use an appropriate GSO for the sub-graph. We intend to explore this approach in the future, including the development of an appropriate GSO. Therefore, for the IEEE 300-bus system as well as the 2383-bus Polish system, we do not show results for localization.

For the IEEE 300-bus system, the deployment strategy placed $T = 99$ PMUs making a total of $M = 419$ measurements for observability. Using $k = 260$ and $j = 10$, the detection results in Figure 11 show that our proposed detection methods can significantly improve the TSA detection performance compared to the LSE-based methods. Finally, for the 2383-bus polish system, the deployment strategy placed $T = 811$ PMUs making a total of $M = 3079$ measurements for observability. Using $k = 1500$ and $j = 10$, the detection

results in Figure 12 show that our proposed detection methods can accurately detect TSAs for this large system.

F. Execution Time

Finally, we evaluate the execution time of the proposed TSA detection and localization methods. Figure 13 shows the results for the IEEE 14, 39, 118, 300, and 2383-bus systems. Note that due to the very high execution time of the baseline AM-LSE-R method for the 2383-bus system, its the execution time is not reported. Observe also that the reported time for the ML methods is the testing time and does not include the time to train the ML model, since training can be done offline. The figure shows that the proposed methods take at most a few milliseconds, and can thus support real-time detection and localization of TSAs. Interestingly, the inference time of the proposed localization methods is at least one order of magnitude less than that of the AM-LSE-R state-of-the-art method [33]. Furthermore, our results show that the computation times increase linearly with the size of the system.

VII. Conclusion

In this paper we proposed methods for the detection and localization of TSAs against PMUs using tools from graph signal processing. Our analytical results indicate that modeling the power system using GSP has great potential in improving its security against TSAs, which we leveraged by combining GSP and machine learning algorithms. We evaluated the proposed methods on the IEEE 14, 39, 118, and 300-bus systems as well as the 2383-bus polish system. Our results showed that among TSAs of non-negligible impact, the proposed methods can outperform state-of-the-art methods based on LSE residuals. An interesting direction of future work would be to extend the detection and localization methods to the time domain, by incorporating appropriate dynamical models of the power system, e.g., based on Kalman filter.

Acknowledgement

The work was partly funded by the Swedish Civil Contingencies Agency (MSB) through the CERCES2 project and by the Swedish Research Council through project 2020-03860.

References

[1] R. S. Singh, H. Hooshyar, and L. Vanfretti, "Assessment of time synchronization requirements for phasor measurement units," in *2015 IEEE Eindhoven PowerTech*, 2015, pp. 1–6.

[2] Z. Zhang, S. Gong, A. D. Dimitrovski, and H. Li, "Time synchronization attack in smart grid: Impact and analysis," *IEEE Trans. on Smart Grid*, vol. 4, no. 1, pp. 87–98, 2013.

[3] X. Jiang, J. Zhang, B. J. Harding, J. J. Makela, and A. D. Domı´nguez-Garcı´a, "Spoofing GPS receiver clock offset of phasor measurement units," *IEEE Trans. on Power Systems*, vol. 28, no. 3, pp. 3253–3262, 2013.

[4] E. Shereen, F. Bitard, G. Dán, T. Sel, and S. Fries, "Next steps in security for time synchronization: Experiences from implementing IEEE 1588 v2.1," in *Proc. of IEEE ISPCS*, 2019, pp. 1–6.

[5] S. Barreto, A. Suresh, and J. Le Boudec, "Cyber-attack on packet-based time synchronization protocols: The undetectable delay box," in *Proc. of IEEE Intl. Instrumentation and Measurement Technology Conf.*, 2016, pp. 1–6.

[6] S. V. S. Chauhan and G. X. Gao, "Synchrophasor data under GPS spoofing: Attack detection and mitigation using residuals," *IEEE Trans. on Smart Grid*, pp. 1–1, 2021.

[7] M. Sarailoo, N. E. Wu, and J. S. Bay, "SA–Based PMU network upgrade for detectability of GPS spoofing attacks," in *2019 IEEE Power Energy Society General Meeting (PESGM)*, 2019, pp. 1–5.

[8] S. Barreto, M. Pignati, G. Dán, J. Le Boudec, and M. Paolone, "Undetectable PMU timing-attack on linear state-estimation by using rank-1 approximation," *IEEE Trans. on Smart Grid*, vol. 9, no. 4, pp. 3530–3542, 2018.

[9] E. Shereen, M. Delcourt, S. Barreto, G. Dán, J. Le Boudec, and M. Paolone, "Feasibility of time-synchronization attacks against PMU-based state estimation," *IEEE Trans. on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3412–3427, 2020.

[10] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. on Signal Processing*, vol. 61, no. 7, pp. 1644–1656, 2013.

[11] R. Ramakrishna and A. Scaglione, "Detection of false data injection attack using graph signal processing for the power grid," in *IEEE Global Conf. on Signal and Information Processing (GlobalSIP)*, 2019, pp. 1–5.

[12] R. Ramakrishna and A. Scaglione, "Grid-graph signal processing (Grid-GSP): A graph signal processing framework for the power grid," *IEEE Trans. on Signal Processing*, 2021.

[13] M. Delcourt, E. Shereen, G. Dän, J.-Y. Le Boudec, and M. Paolone, "Time-synchronization attack detection in unbalanced three-phase systems," *IEEE Trans. on Smart Grid*, vol. 12, no. 5, pp. 4460–4470, 2021.

[14] A. S. Musleh, G. Chen, and Z. Y. Dong, "A survey on the detection algorithms for false data injection attacks in smart grids," *IEEE Trans. on Smart Grid*, vol. 11, no. 3, pp. 2218–2234, 2020.

[15] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. of ACM CCS*, 2009, p. 21–32.

[16] G. Dán and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *Proc. of IEEE SmartGridComm*, 2010, pp. 214–219.

[17] M. Jin, J. Lavaei, and K. H. Johansson, "Power grid AC-based state estimation: Vulnerability analysis against cyber attacks," *IEEE Trans. on Automatic Control*, vol. 64, no. 5, pp. 1784–1799, 2019.

[18] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Trans. on Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012.

[19] T. A. Alexopoulos, G. N. Korres, and N. M. Manousakis, "Complementarity reformulations for false data injection attacks on PMU-only state estimation," *Electric Power Systems Research*, vol. 189, 2020.

[20] 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. on Industrial Informatics*, vol. 11, no. 5, pp. 1–12, 2015.

[21] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Sparse attack construction and state estimation in the smart grid: Centralized and distributed models," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1306–1318, 2013.

[22] O. Vukovic, K. C. Sou, G. Dan, and H. Sandberg, "Network-aware mitigation of data integrity attacks on power system state estimation," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1108–1118, 2012.

[23] N. Živković and A. T. Sarić, "Detection of false data injection attacks using unscented Kalman filter," *Journal of Modern Power Systems and Clean Energy*, vol. 6, pp. 847–859, 2018.

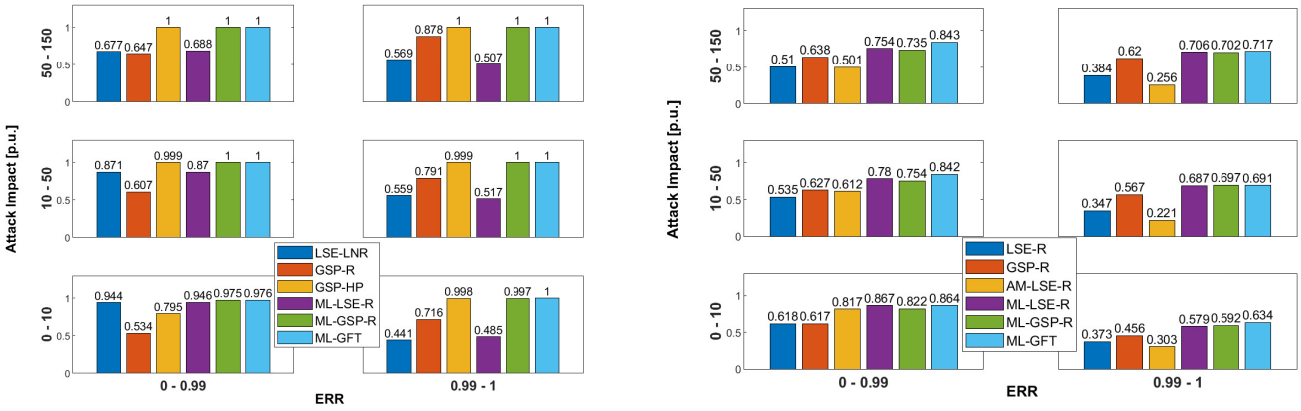
[24] H. Long, Z. Wu, C. Fang, W. Gu, X. Wei, and H. Zhan, "Cyber-attack detection strategy based on distribution system state estimation," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 4, pp. 669–678, 2020.

[25] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Trans. on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1773–1786, 2016.

[26] E. Drayer and T. Routtenberg, "Detection of false data injection attacks in power systems with graph fourier transform," in *IEEE Global Conf. on Signal and Information Processing (GlobalSIP)*, 2018, pp. 890–894.

[27] J. Shi, B. Foggo, X. Kong, Y. Cheng, N. Yu, and K. Yamashita, "Online event detection in synchrophasor data with graph signal processing," in *Proc. of IEEE SmartGridComm*, 2020, pp. 1–7.

[28] F. Zhu, A. Youssef, and W. Hamouda, "Detection techniques for data-level spoofing in GPS-based phasor measurement units," in *2016 Intl. Conf. on Selected Topics in Mobile Wireless Networking (MoWNeT)*, 2016, pp. 1–8.



(a) TSA Detection

(b) TSA Localization

Fig. 10: AUC achieved by the TSA (a) detection and (b) localization methods for the IEEE 118-bus system

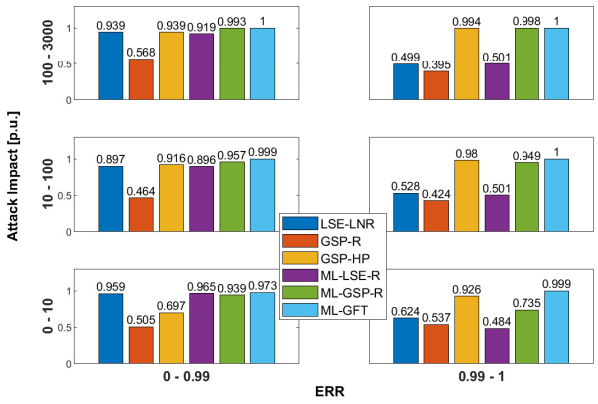


Fig. 11: AUC achieved by the TSA detection methods for the IEEE 300-bus system

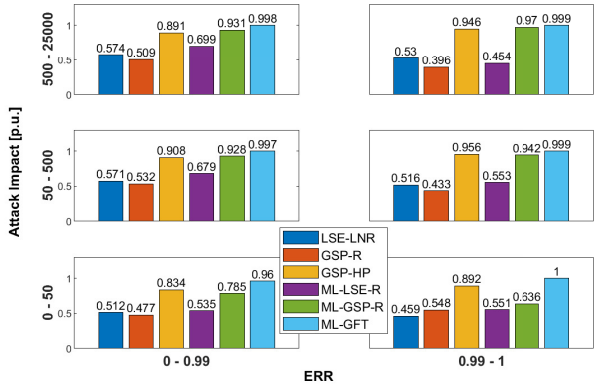


Fig. 12: AUC achieved by the TSA detection methods for the 2383-bus polish system

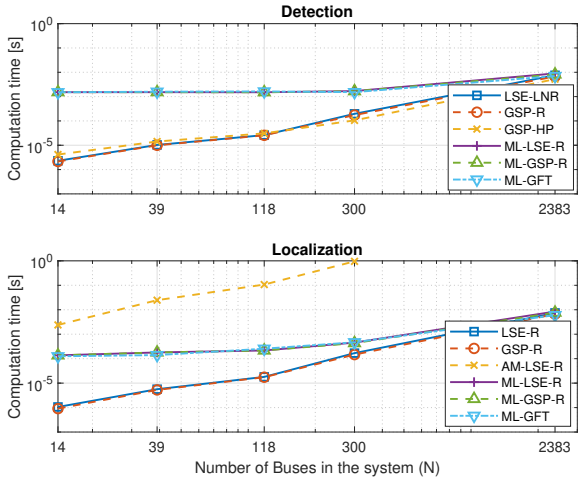


Fig. 13: Comparison of the mean computation time per sample for the TSA detection and localization methods

[29] B. Moussa, M. Debbabi, and C. Assi, "A detection and mitigation model for PTP delay attack in an IEC 61850 substation," *IEEE Trans. on Smart Grid*, vol. 9, no. 5, pp. 3954–3965, 2018.

[30] X. Fan, L. Du, and D. Duan, "Synchrophasor data correction under GPS spoofing attack: A state estimation-based approach," *IEEE Trans. on Smart Grid*, vol. 9, no. 5, pp. 4538–4546, 2018.

[31] Y. Zhang, J. Wang, and J. Liu, "Attack identification and correction for PMU GPS spoofing in unbalanced distribution systems," *IEEE Trans. on Smart Grid*, vol. 11, no. 1, pp. 762–773, 2020.

[32] S. Siamak, M. Dehghani, and M. Mohammadi, "Dynamic GPS spoofing attack detection, localization, and measurement correction exploiting PMU and SCADA," *IEEE Systems Journal*, pp. 1–10, 2020.

[33] P. Risbud, N. Gatsis, and A. Taha, "Vulnerability analysis of smart grids to GPS spoofing," *IEEE Trans. on Smart Grid*, vol. 10, no. 4, pp. 3535–3548, 2019.

[34] S. A. Desilva, J. Kim, E. Cotilla-Sanchez, and T. Hagan, "On PMU data integrity under GPS spoofing attacks: A sparse error correction framework," *IEEE Trans. on Power Systems*, pp. 1–1, 2021.

[35] S. De Silva, J. Kim, and E. Cotilla-Sanchez, "Data driven sparse error correction for PMU measurements under GPS spoofing attacks," in *2021 IEEE Power Energy Society Innovative Smart Grid Technologies Conf. (ISGT)*, 2021, pp. 1–5.

[36] A. Xue, F. Xu, J. Xu, J. H. Chow, S. Leng, and T. Bi, "Online pattern recognition and data correction of PMU data under GPS spoofing attack," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1240–1249, 2020.

[37] Y. Fan, Z. Zhang, M. Trinkle, A. D. Dimitrovski, J. B. Song, and H. Li, "A cross-layer defense mechanism against GPS spoofing attacks on PMUs in smart grids," *IEEE Trans. on Smart Grid*, vol. 6, no. 6, pp. 2659–2668, 2015.

[38] E. Shereen and G. Dán, "Model-based and data-driven detectors for time synchronization attacks against PMUs," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 169–179, 2020.

[39] A. Abur and A. G. Expósito, "Power system state estimation : Theory and implementation." Marcel Dekker, 2004.

[40] T. Humphreys, B. Ledvina, M. Psiaki, B. O'Hanlon, and J. Kintner, "Assessing the spoofing threat: Development of a portable GPS civilian spoofer," in *Proc. of the Institute of Navigation GNSS (ION GNSS)*, 2008, pp. 2314–2325.

[41] M. Delcourt and J.-Y. Leboudec, "Security measures for grids against rank-1 undetectable time-synchronization attacks," *IEEE Trans. on Control of Network Systems*, pp. 1–1, 2021.

[42] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Trans. on Power Systems*, vol. 31, no. 2, pp. 1006–1013, 2016.

[43] Y. Ge, A. J. Flueck, D.-K. Kim, J.-B. Ahn, J.-D. Lee, and D.-Y. Kwon, "Power system real-time event detection and associated data archival reduction based on synchrophasors," *IEEE Trans. on Smart Grid*, vol. 6, no. 4, pp. 2088–2097, 2015.

[44] K. Mahapatra and N. R. Chaudhuri, "Online robust PCA for malicious attack-resilience in wide-area mode metering application," *IEEE Trans. on Power Systems*, vol. 34, no. 4, pp. 2598–2610, 2019.

[45] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[46] N. Sato and W. Tinney, "Techniques for exploiting the sparsity or the network admittance matrix," *IEEE Transactions on Power Apparatus and Systems*, vol. 82, no. 69, pp. 944–950, 1963.

[47] T. K. Ho, "Random decision forests," in *Proceedings of 3rd Intl. Conf. on Document Analysis and Recognition*, vol. 1, Aug 1995, pp. 278–282.

[48] R. D. Zimmerman and C. E. Murillo-Sanchez, "(2019). MATPOWER (version 7.0) [Software]," available: https://matpower.org.

[49] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, pp. 8024–8035.

[50] B. Gou, "Generalized integer linear programming formulation for optimal pmu placement," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1099–1104, 2008.

[51] G. N. Korres, "A distributed multiarea state estimation," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 73–84, 2011.

[52] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617–1626, 2013.

Ezzeldin Shereen (Member, IEEE) received the B.Sc. and M.Sc. degrees in networking engineering from German University, Cairo, Egypt, in 2014 and 2015, respectively. He received the Ph.D. degree in 2021 at the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden, where he is currently working as a Postdoctoral Researcher. His current research interests include cyberphysical systems security with focus on power systems, machine learning, and performance evaluation of wireless networks.

Raksha Ramakrishna (Member, IEEE)received the B.E degree in electronics and communications engineering from the Rashtreeya Vidyalaya College of Engineering, Bangalore, India, in 2014, and the M.S. and Ph.D. degrees in electrical engineering from Arizona State University, in 2017 and 2020, respectively. She is currently a Postdoctoral Researcher with the Division of Network and Systems Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. Her research interests include the domain of statistical signal processing, data analytics for power systems, and recently in security and privacy in federated machine learning systems.

György Dán (Senior Member, IEEE) received the M.Sc. degree in computer engineering from the Budapest University of Technology and Economics, Budapest, Hungary, in 1999, the M.Sc. degree in business administration from the Corvinus University of Budapest, Budapest, in 2003, and the Ph.D. degree in telecommunications from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2006. From 1999 to 2001, he was a Consultant in the field of access networks, streaming media, and videoconferencing with BCN Ltd., Budapest. He was a Visiting Researcher with the Swedish Institute of Computer Science, Stockholm, in 2008, a Fulbright Research Scholar with the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 2012 and 2013, and an Invited Professor with the Swiss Federal Institute of Technology of Lausanne (EPFL), Lausanne, Switzerland, in 2014 and 2015. He is currently a Professor with the KTH Royal Institute of Technology. His current research interests include the design and analysis of content management and computing systems, game theoretical models of networked systems, and cyber-physical system security and resilience. Dr. Dán has been an Area Editor of Computer Communications since 2014 and the IEEE TRANSACTION ON MOBILE COMPUTING since 2019.