Characterization and Classification of Cyber Attacks in Smart Grids using Local Smoothness of Graph Signals

Md Abul Hasnat and Mahshid Rahnamay-Naeini
Electrical Engineering Department, University of South Florida, Tampa, Florida, USA
hasnat@usf.edu, mahshidr@usf.edu

2021 North American Power Symposium (NAPS) | 978-1-6654-2081-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/NAPS52732.2021.9654459

***Abstract*—Characterization and classification of cyber attacks in smart grids are crucial for situational awareness and mitigation of their effects. Graph signal processing (GSP) framework for the analysis of energy data, provides new perspectives and opportunities for such characterization by capturing topology, interconnections, and interactions among the components of smart grids. In this work, several forms of cyber stresses on power system's measurements and state estimation have been analyzed using the local smoothness of their graph signals. Using the local smoothness, characteristics of different cyber stresses are described analytically and evaluated by simulations. Moreover, the local smoothness features are used in machine learning models to classify multiple random and clustered cyber stresses and determine attack center and radius in case of clustered attacks.**

***Index Terms*—Graph Signal Processing, cyber-physical stress, smart grid, local smoothness, classification.**

I. Introduction

Situational awareness of large, complex, and dynamic cyber-physical infrastructures, such as smart grids, has been a salient concern for researchers and practitioners to ensure their reliable and efficient operation. Functions supporting situational awareness rely on the measurement data communicated from measurement devices and sensors through the communication system. However, the cyber layer of smart grids including the sensing, communication, and computing components, are vulnerable to various forms of cyber attacks and stresses.

Cyber attacks can be launched in the smart grid by hampering the availability (e.g., denial-of-service aka DoS attack), or the integrity (e. g., false data injection attack - FDIA [1]) of the measurement data, which are crucial for situational awareness and consequently for proper operation and timely maintenance of smart grids. Moreover, the cyber attacks can be coordinated or random depending on the intention and available resources of the attackers. As such, proper characterization of the nature and types of cyber attack is necessary.

Graph signal processing (GSP) [2], [3] is an emerging and fast-growing domain that enables analyzing structured data over graphs. The tools and techniques from GSP are particularly suitable for analyzing data that are topologically distributed as well as represent and embed the interactions and interconnections among the components of a system, such as smart grids. By extending the concepts and the tools from the classical signal processing to non-Eucleadian signals defined in the irregular graph domain, GSP provides an opportunity to incorporate the hidden information about the structures, interconnectivity, and interrelations among the components in the analysis of the system data. GSP tools have been recently adopted in various fields for analysis of structured and inter-correlated data including analysis of the brain data [4], structural health monitoring by the sensor networks [5], and crime data analysis [6]. The measurement data from power systems bear structures by nature due to the physics of electricity and the structural and operating conditions governing these systems. These properties suggest that GSP can be a suitable framework for the analysis of the energy data.

Graph signals in power systems can be defined over the structural topology of the system with buses as the vertices and the transmission lines as the edges of the graph. In our earlier works [7], [8], different types of cyber attacks and their signatures on the graph spectrum were analyzed using voltage angle measurements as the time-varying graph signals. In addition, a local smoothness-based technique was proposed to detect and locate a single cyber attack in the system, which showed superior performance. In the current paper, the study and techniques in [7], [8] are extended to detect and locate multiple and coordinated cyber attacks in smart grids with an analytical derivation of the conditions and simulations to show the performance. Specifically, it has been shown analytically that the attackers can bypass the detection technique by designing a coordinated cyber attack using the local smoothness properties of the vertices. Furthermore, graph signal smoothness features are used to classify random and clustered cyber attacks in the system.

The main contributions of this paper are as follows:

- The effects of a single cyber attack, multiple random cyber attacks, and multiple clustered/coordinated cyber attacks on the local smoothness of graph signals have been analyzed assuming no load demand changes in the system. By analytical evaluation of the effects of attacks, characterization of the stresses (i.e., determining the type of the attack, number of the attacked buses, size of the area under attack) has been enabled.
- A neural network-based technique for the classification of multiple random cyber attacks, and the clustered/coordinated cyber attacks has been proposed.
- The effectiveness and the limitation of the local

smoothness-based detection technique has been evaluated analytically based on the analysis of the attack strategies and the available information to the attackers.

II. Related Works

Cyber attacks in smart grids have been studied intensively over the past two decades. The majority of the works in this domain are on the detection of different types of cyber attacks in power systems using various techniques including dimensionality reduction methods and principal component analysis (PCA) [9], signal processing and statistical analysis-based methods [10]–[12], neural network and reinforcement learning-based methods [13]–[15], and graph neural network-based method [16]. In addition, some works, such as the ones in [16]–[18], addressed the problem of locating the attacks along with the detection. The work by Khalafi *et al.* [19] provides a comparative discussion and summary of the false data injection attack (FDIA) literature on the basis of detection techniques, types of data used, the ability to locate attacks, the ability to detect concurrent random attacks, and determining the number of attacks. Along with the detection and locating the cyber stresses, the current paper proposes a technique for determining the number of cyber attacks. While there is a large body of research pertaining to the security of smart grids, new perspectives provided by the GSP tools and techniques can present new opportunities for improving the security of these critical systems.

The application of GSP to smart grid problems is fairly recent. For instance, Ramakrishna and Scaglione in [20] developed a graph signal model for the time-varying voltage phasor measurements in the smart grid capturing the spatio-temporal dynamics among the grid's components. The model in [20] was proposed to be utilized for the problems of detecting stresses, recovering the measurement values, and optimal placement of the PMUs in the smart grid. Kroizer *et al.* in [21] proposed a graph filter model for the non-linear measurement functions in the power system and proposed to utilize the inverse system of the graph filter for the recovery of the grid signal. Drayer and Routtenberg [22] proposed an FDIA detection technique using graph Fourier transform based on the assumption that bus voltage angles in normal conditions are smooth over the grid and thereby do not contain high graph-frequency components. Shi *et al.* [23] also proposed a GSP-based anomaly detection technique in the smart grid. Although the global graph spectral quantities (e.g., GFT) are used widely for anomaly detection, the direct use of local smoothness values of graph signals is limited in the literature. Dwivedi and Tajer in [24] proposed a technique for detecting line outages in the smart grid based on the global smoothness of graph signals and locating the outages based on the local smoothness of the graph signals associated with the power grid. In the latter work, the conformity of the data with a model, which is based on a known topology, was used to detect the topological changes, i.e., line outages in the grid.

III. Cyber Attack Characterization

A. Power System Model and Graph Signals

The power system is modeled as an undirected weighted graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, v_2, ..., v_N\}$ is the set of all the vertices representing the buses and $\mathcal{E}$ is the set of links (i.e., $\mathcal{E} = \{e_{ij} : (i,j) \in \mathcal{V} \times \mathcal{V}\}$) representing the transmission lines connecting the buses. The cardinality of the sets are $|\mathcal{V}| = N$ and $|\mathcal{E}| = M$, respectively. The weight corresponding to the edge e_{ij} is defined as $w_{ij} = \frac{1}{d_{ij}}$, if there is an edge between node i and node j (i.e., $e_{ij} = 1$) and $w_{ij} = 0$, otherwise (if there is no edge between node i and node j, i.e., $e_{ij} = 0$). Here, d_{ij} is the geographical distance between the buses, i and j. The graph Laplacian matrix $\mathbf{L}$, with elements l_{ij}, is defined as $l_{ij} = \sum_{j=1}^{N} w_{ij}$ if $i = j$ and $l_{ij} = -w_{ij}$, otherwise.

In this work, the voltage angles of the buses at a particular time are considered as the graph signal $x(v_n) : \mathcal{V} \rightarrow \mathbb{R}$ defined over the vertices of the graph, $\mathcal{G}$. For simplicity, $x(v_n)$ will be further denoted as $x(n)$ and $\underline{\mathbf{x}}$ would be considered as a $N \times 1$ column vector containing the values of $x(n)$. Since the voltage angle of the buses changes with time, we denote the time-varying graph signal as $x(n,t)$.

B. Cyber Attack Model

In this work, the cyber attack models described in our earlier work [7] are adopted. Specifically, according to [7], the cyber attack on the measurement or the estimated value of the signal associated with the buses can be expressed by the following generalized model:

$$x(n_A, t) = c(t), \quad for \quad t_{start} \leq t \leq t_{end}, \quad and \quad n_A \in \mathcal{V}_{\mathcal{A}}, \tag{1}$$

where $\mathcal{V}_{\mathcal{A}} \subset \mathcal{V}$ is set of all buses under cyber attack from time t_{start} to t_{end}. Different expressions for the corrupted signal $c(t)$ has been used to model the effect of DoS attack, FDIA, data-replay attack, ramp attack, and delay attack on the time-varying graph signals associated with the bus voltage angles of the smart grid. A detailed description of the models can be found in [7].

1) Denial-of-service (DoS Attack): In this work, the DoS attack has been modeled as the suspension of updating the graph signal values corresponding to compromised buses. For DoS attack, $c(t) = x(n_A, t_{start}) + q(t)$, where $q(t)$ is the additive white Gaussian noise (AWGN) with zero mean and variance $\sigma^2_{n_A}$.

2) False Data Injection Attack (FDIA): In this paper, the FDIA is modeled according to the following expression of the corrupted signal: $c(t) = x(n_A, t) + (-1)^d x'$, where $d \in \{0, 1\}$, $|x'| \leq \tau$, and τ is the residue threshold used in the traditional bad data detector of the smart grid.

3) Ramp Attack: In the ramp attack, the attacker gradually inserts falsified measurements into the data corresponding to the compromised buses according to the equation: $c(t) = x(n_A, t_{start}) + m \times (t - t_{start}) + q(t)$, where m is the slope of the linear change.

4) Replay Attack: In general, the replay attack can be modeled using the following expression of the corrupted signal: $c(t) \in \{x(n_R, t_p)\}, \quad t_p < t_{start}, \quad n_R \in \mathcal{V}_\mathcal{R}$, where $\mathcal{V}_\mathcal{R} \subset \mathcal{V}$ is the set of all vertices from which the the attackers can record data at $t < t_{start}$. In this work, a special type of replay attack has been proposed according to the expression: $c(t) = x(n_A, -t)$.

5) Delay Attack: The delay attack involves loss of synchronization in the time stamped measurements and modeled as: $c(t) = x(n_A, t - t_d)$, where t_d is the amount of delay.

C. Effects of Cyber Attacks on Local Smoothness of Graph Signals

The local smoothness [25] of the graph signal $x(n)$ associated with the bus voltage angles is described by:

$$s(n,t) = \frac{l_x(n,t)}{x(n,t)}, \quad x(n,t) \neq 0, \tag{2}$$

where $l_x(n,t)$ is the $n-$th element of the vector $L\underline{x}$ and $\underline{x}$ is the vector form of the graph signal $x(n,t)$. In our previous works [7], [8], it has been shown that cyber stresses in the smart grid can be detected and located using GSP techniques based on the local smoothness of the graph signals. In this approach, the probability distributions of the second time-derivative of the local smoothness values associated with the voltage angle measurements of each bus $p_{s''_n}(\zeta)$ are estimated from the past data. If the likelihood of the second time-derivative of the local smoothness value at any time instant at bus/vertex n falls below the threshold $\theta_{s''_n}$ (i.e., $p_{s''_n}(s''(n,t)) < \theta_{s''_n}$), a cyber stress is declared at bus n at that time. The time instant at which the attack is detected is denoted as t_{detect}. It is worth mentioning that the cyber stresses affect the local smoothness values significantly. The second time-derivative of the local smoothness is considered to cancel the effect of non-stationarity in $s(n,t)$ originated from the non-stationarity in the graph signal $x(n,t)$ due to the variation of the load demand in the grid. Our experiments showed that the proposed local smoothness-based methods outperform other GSP-based detection and locating methods in terms of accuracy, especially for detecting cyber attacks with no sharp changes of values at the attack onset. In the current paper, the effects of cyber attacks on the local smoothness values are inspected analytically to evaluate the effectiveness of the detection and locating method in case of multiple, clustered, and coordinated cyber attacks.

In this paper, by utilizing the sparse nature of the graph Laplacian matrix, $\mathbf{L}$, we express $l_x(n,t)$ alternatively as:

$$l_x(n,t) = \sum_{i \in \{n\} \cup \mathcal{N}_1(n)} l_{ni} x(i,t), \tag{3}$$

where $\mathcal{N}_k(n)$ is the set of all vertices in the $k-$hop neighborhood of n. Next, the effects of cyber attacks on the local smoothness $s(n)$ will be illustrated.

a) Single Cyber Attack Case with Fixed Load: Let us consider a single cyber attack at bus $n_A \in \mathcal{V}_\mathcal{A}$. According to equation (3), this single attack will affect $l_x(n, t_{detect})$ for $n \in \{n_A\} \cup \mathcal{N}_1(n_A)$. According to equation (2), the local smoothness values for these vertices are affected because of the single cyber attack at the vertex n_A. For instance, Fig.1(a) illustrates the difference between the local smoothness values of each vertex before and after a single cyber attack at bus 100.

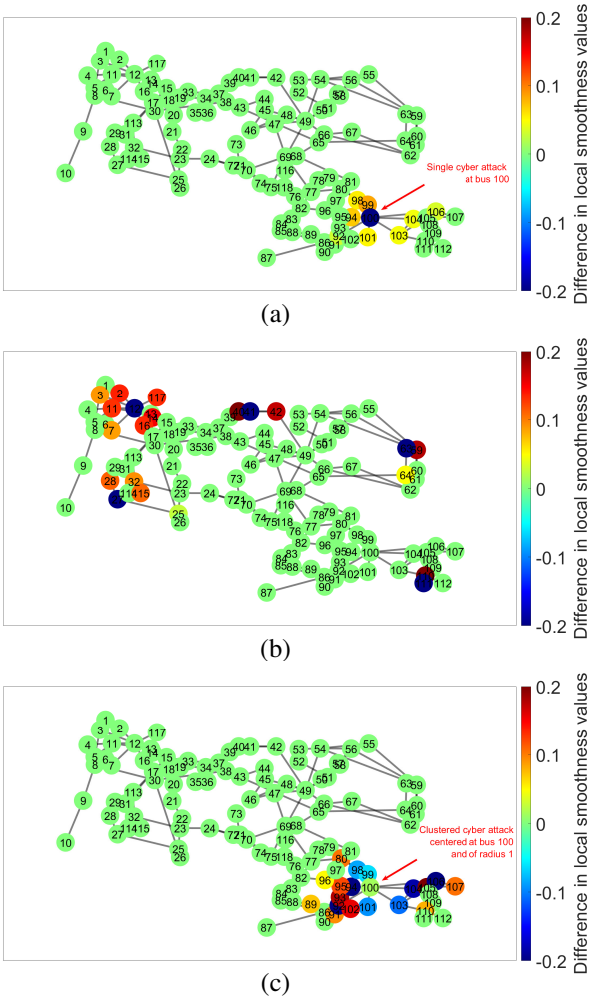

Fig. 1. Changes in the local smoothness values due to different types of false data injection attack assuming no load changes in the system: (a) single attack at bus 100 affects bus 100 and its $1-$hop neighbors i. e. $\{100\} \cup \mathcal{N}_1(100)$, (b) multiple attacks at bus 12, 27, 41, 63, 111 affect those buses and their $1-$hop neighbors i.e. $\{12, 27, 41, 63, 111\} \cup \mathcal{N}_1(12) \cup \mathcal{N}_1(27) \cup \mathcal{N}_1(41) \cup \mathcal{N}_1(63) \cup \mathcal{N}_1(111)$ (c) clustered attack centered at bus 100 and radius 1 affects $\{100\} \cup \mathcal{N}_1(100) \cup \mathcal{N}_2(100)$.

b) Multiple Random Cyber Attack Case with Fixed Load: Let us consider p number of cyber attacks at buses $n_{A_1}, n_{A_2}, \ldots n_{A_p} \in \mathcal{V}$. According to equation (3), this multiple attack will affect $l_x(n, t_{detect})$ and thereby the local smoothness $s(n, t_{detect})$ for $n \in \{n_{A_1}, n_{A_2}, \ldots n_{A_p}\} \cup \mathcal{N}_1(n_{A_1}) \cup \mathcal{N}_1(n_{A_2}) \ldots \mathcal{N}_1(n_{Ap})$. For instance, Fig. 1(b) illustrates the effect of FDIA at buses 12, 27, 41, 63, 111.

c) Clustered Cyber Attack Case with Fixed Load: In a clustered cyber attack case, we assume the attacker attacks a central node and its $K-$hop neighbors. The parameter K is called the *radius* of the attack. In the clustered cyber attacks, the attacker can inject false data at any vertex within the radius

K. For example, in a clustered cyber attack with attack center $n_C \in \mathcal{V}$ and radius 1, the attacker changes the graph signal $x(n)$ for $n \in \{n_C\} \cup \mathcal{N}_1(n_{A_C})$. According to equation (3), due to the changes of the signal values in the attack center n_C, the value of $l_x(n)$ would change for $n \in \{n_C\} \cup \mathcal{N}_1(n_C)$ and due to the changes of the value at each of the vertices $n'_C \in \mathcal{N}_1(n_C)$, the value of $l_x(n, t_{detect})$ would change for $n \in \mathcal{N}_1(n'_C)$, $\forall n'_C \in \mathcal{N}_1(n'_C)$. Therefore, a clustered attack centered at n_C and radius 1 would affect the values of $l_x(n, t_{detect})$ for the vertices $n \in \{n_C\} \cup \mathcal{N}_1(n_C) \cup \mathcal{N}_2(n_C)$. According to equation (2), the local smoothness $s(n, t_{detect})$ changes for these vertices. A clustered attack centering at bus 100 and radius 1 has been considered as an example at Fig. 1(c). In general, a clustered cyber attack, with attack center n_C and radius K, can affect the local smoothness values of the vertices: $\{n_C\} \cup \{\bigcup_{j=1}^{K+1} \mathcal{N}_j(n_C)\}$.

The above discussion provides an insight for detecting, locating, and characterizing cyber attacks in power grids based on the local smoothness of the associated graph signals. However, the assumption of no-load change does not hold in real-life scenarios. Therefore, due to the perpetual changes of load demands, the graph signal $x(n, t)$ and thereby the local smoothness $s(n, t)$ change continuously over time. It is a challenge to distinguish the changes in local smoothness due to the cyber attack from the regular changes in local smoothness due to the load changes. To overcome this problem, we propose to estimate the probability distribution of the second time-derivative of the local smoothness values for each of the buses under the load changes from the past data. If the likelihood of the second time-derivative of the local smoothness at any bus falls below a certain threshold, an attack is declared at that vertex at that time instant. This technique is able to detect and locate sophistically designed cyber attacks very accurately. However, the main focus of the current article is to characterize the attacks based on the local smoothness features.

D. Classification between Multiple Random Attack and Clustered Attack

Based on the discussions in the previous subsection, we can conclude that single random attacks, multiple random attacks, and clustered cyber attacks have distinctive signatures in the pattern of local smoothness values at the time of detecting the attack, t_{detect}. However, under the load change at different buses, the voltage angle graph signals and thereby the local smoothness values associated with the signals vary in time. As a result, the rule-based decision-making from the signatures of multiple random cyber attacks and the clustered cyber attacks becomes difficult. Therefore, we propose a neural network-based classification between the two types of attacks.

A total of $N+1$ input features, $f_1, f_2, \ldots f_{N+1}$, are considered for the deep learning model. Among them, the first N features are binary, indicating whether the likelihood of the second time-derivative of the local smoothness value of a particular bus at the detection instant is less than a predefined threshold $\theta_{s''_n}$ (i.e., $f_i = 1$ if $p_{s''_n}(s''(n, t_{detect})) < \theta_{s''_n}$ and 0 otherwise, for $i = 1, 2, \ldots N$). The last feature, f_{N+1}, is a real-valued feature representing the global smoothness of the graph signal, $x(n, t_{detect})$, which is expressed as $\frac{\underline{\mathbf{x}}^T(n,t_{detect})\mathbf{L}\underline{\mathbf{x}}(n,t_{detect})}{\underline{\mathbf{x}}^T(n,t_{detect})\underline{\mathbf{x}}(n,t_{detect})}$.

E. Determining Attack Center and Attack Radius in Clustered Cyber Attacks

After the identification of a cyber attack as a clustered one, it is crucial to determine the center of the attack and the attack radius to enhance the situational awareness and mitigate the effect of the attack. In this work, the goal is to identify the center of the attack, n_C and the attack radius, K to help with situational awareness.

In this work, the problem of locating the attack center is formulated as an $N-$class classification problem considering each bus as a class. The first N features used in the classification between multiple random and clustered attack are proposed to be taken as the input features. A multi-class $k-$nearest neighbor (kNN) technique is used to solve the classification to obtain the location of the center of the attack.

Once the center of the attack, n_C, is detected, the radius, K of the attack can be estimated using the fact that a clustered attack of radius K, affects the local smoothness values of the vertices within the $K+1$ neighborhood of the the attack center. The radius K can be specified as $K = \max\{\mathbb{D}(n_C, n_P)\} - 1$, where $\forall n_P \in \mathcal{V}$ such that $p_{s''_{n_P}}(s''(n_P, t_{detect})) < \theta_{s''_{n_P}}$ and $\mathbb{D}(n_1, n_2)$ is the hop-distance between the vertices n_1 and n_2, within the graph, $\mathcal{G}$.

IV. Simulation and Performance Analysis

The analyses and the proposed techniques described in the previous section have been evaluated using simulations on the *IEEE* 118 *bus system* [26]. The power flow calculations are performed in *MATPOWER 6.0* [27]. The time-varying bus voltage angle signals are simulated by introducing time-varying load demand. The patterns of the time variation of load demand throughout the day have been collected from the New York Independent System Operator (NYISO) [28] and applied as described in [7]. Different types of cyber attacks are created according to the attack model as described in [7]. A detailed description of each experiment and their performances have been presented in the following subsections.

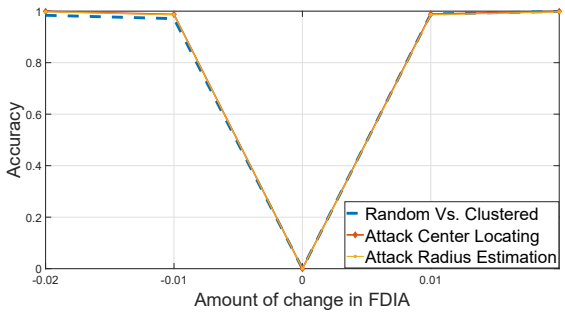

Fig. 2. Dependence of accuracies on attack intensities for FDIA.

TABLE I
PERFORMANCE EVALUATIONS FOR DIFFERNT TYPES OF CYBER ATTACKS.

Attack Type	Accuracy		
	Random Vs. Clustered Classification	Attack Center Locating (Clustered)	Attack Radius Estimation (Clustered)
DoS	0.851	0.886	0.978
Replay	0.922	0.883	0.955
FDIA	See Fig. 2	See Fig. 2	See Fig. 2
Ramp	0.858	0.849	0.973
Delay	See Fig. 3	See Fig.3	See Fig. 3

A. Classification between Multiple Random and Clustered Cyber Attacks

For the classification between multiple random and clustered cyber attacks, the performance is evaluated separately for each of the five types (i.e., DoS, FDIA, data-replay, ramp, and delay attacks) of attacks as well as for different levels of attack intensities. For each case, we create a data-set with $10,000$ scenarios. Whether a scenario corresponds to multiple random cyber attack or clustered cyber attack is chosen randomly with equal probabilities. For the multiple random attacks (e.g. p number of attacks), the attack locations (p locations) are chosen from the 118 buses, with uniform probabilities for all the buses. In the case of the clustered cyber attack, the attack-center (n_C) is chosen randomly from the 118 buses with equal probabilities.

The deep learning model for the classification between the random and the clustered attack consists of 3 hidden dense layers with $256, 128,$ and 32 neurons, respectively. With the binary cross-entropy loss function, ADAM optimizer, and an initial learning rate of 0.5 which decreases at a rate of 0.5 exponent of the time step. The model has been trained and tested in *Sci-kit learn* [29] with $10-$ fold cross-validation. The performance of the classification model for the sophistically designed cyber attacks described in equation (1) has been summarized in Table I. The accuracy of the classification signifies the rate of classifying in between the random attack and clustered attack correctly. Since the performance in the case of the FDIA and the delay attack is dependent on the attack intensities, their performances are illustrated separately in Fig. 2 and Fig. 3, respectively showing the variation with the amount of change in FDIA, x' and amount of delay (in samples) in delay attack.

B. Determining n_C and K in Clustered Cyber Attacks

For the determination of the attack center, n_C, and the attack radius K in case of clustered cyber attack, we have considered $10,000$ clustered cyber attack scenarios. In each scenario, the attack center, n_C, has been chosen from the 118 buses and the attack radius, K has been chosen from $\{1, 2, 3, 4\}$, with equal probability. Five nearest neighbors are considered in the $k-$NN method for the classification for determining the attack center. Performances are summarized in Table I, Fig. 2, and Fig. 3. The accuracy of determining the attack center and the attack radius imply the rates of correctly determining the location of the central bus of the clustered attack and the radius of the attack.

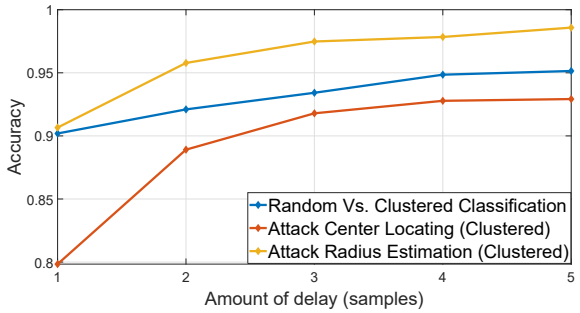

Fig. 3. Dependence of accuracies on amount of delay in delay attack.

V. STRENGTH OF THE LOCAL SMOOTHNESS-BASED DETECTION AGAINST COORDINATED ATTACK

In this section, the effectiveness of the local smoothness-based detection of cyber attacks has been analyzed from the attacker's point of view for bypassing the detection scheme. It has been shown that by launching a coordinated cyber attack in multiple vertices, the detection techniques can be bypassed by the attackers, although the coordination requires information about the grid topology inside the attack area as well as past data to estimate the distribution of the local smoothness values at different vertices of the grid.

To establish the coordination technique for the attack at multiple buses, let us consider that the attackers have access to all the buses of the set $\mathcal{V}_{\mathcal{A}}$ to alter the signal values associated with the buses. Let us denote the buses (vertices) in which attackers do not have access (uncompromised buses) as $\mathcal{V}_{\mathcal{U}}$, i.e., $\mathcal{V}_{\mathcal{U}} = \mathcal{V} \setminus \mathcal{V}_{\mathcal{A}}$. From equation (2) and equation (3) the local smoothness of the vertices can be written as:

$$s(n,t) = \frac{1}{x(n,t)} \sum_{i \in \{n\} \cup \mathcal{N}_1(n)} l_{ni} x(i,t), \tag{4}$$

which can be further decomposed into the following with respect to the compromised and uncompromised buses as:

$$\begin{aligned} s(n,t)x(n,t) = & \sum_{i \in [(\{n\} \cup \mathcal{N}_1(n)) \cap \mathcal{V}_{\mathcal{A}}]} l_{ni} x(i,t) \\ & + \sum_{j \in [(\{n\} \cup \mathcal{N}_1(n)) \cap \mathcal{V}_{\mathcal{U}}]} l_{nj} x(j,t), \end{aligned} \tag{5}$$

The relation among the signal values and local smoothness values of the compromised buses and the uncompromised buses at a particular time instant can be represented by a set of equations expressed in the following matrix form:

$$(\mathbf{L}_{\mathcal{A}} - \mathbf{S}_{\mathcal{A}})\underline{\mathbf{x}}_{\mathcal{A}} = \underline{\mathbf{l}}_{\underline{\mathbf{x}}_{\mathcal{U}}} - \mathbf{L}_{\mathcal{U}}\underline{\mathbf{x}}_{\mathcal{U}}, \tag{6}$$

where $\mathbf{L}_{\mathcal{A}}$ is a matrix containing the $a-$th columns of $\mathbf{L}$, $\forall a \in \mathcal{V}_{\mathcal{A}}$ and $\mathbf{L}_{\mathcal{U}}$ is a matrix containing the $u-$th columns of $\mathbf{L}$, $\forall u \in \mathcal{V} \setminus \mathcal{V}_{\mathcal{A}}$ in the sequence of the original Laplacian matrix $\mathbf{L}$. $\mathbf{S}_{\mathcal{A}}$ is a $N \times |\mathcal{V}_{\mathcal{A}}|$ matrix whose $(i,i)-$th element is $s(i,t)$, when $i \in \mathcal{V}_{\mathcal{A}}$, and 0 otherwise, and $t \in [t_{start}, t_{end}]$. $\underline{\mathbf{l}}_{\underline{\mathbf{x}}_{\mathcal{U}}}$ is a $N \times 1$ column vector whose $i-$th element is $l_x(i)$ if $i \notin \mathcal{A}$,

and 0 otherwise. $\underline{\mathbf{x}}_{\mathcal{U}}$ and $\underline{\mathbf{x}}_{\mathcal{A}}$ are the column vector containing the state variables i.e., the graph signal values corresponding to the compromised and uncompromised buses, respectively, at any $t \in [t_{start}, t_{end}]$.

Although the detection method is based on the likelihood of the second time-derivative of the local smoothness, $s''(n,t)$ rather than the local smoothness, $s(n,t)$, it can be justified that maximizing the likelihood of $s(n,t)$ would maximize the likelihood of $s''(n,t)$ and thereby have a greater probability to bypass the detection technique. Therefore, the coordinated attack design problem can be formulated as an optimization problem:

$$\begin{aligned} \underset{x(n),\forall n\in\mathcal{V}_{\mathcal{A}}}{\text{minimize}} \quad & \sum_{n\in\mathcal{V}} log(p_{s_n}(\zeta)) \\ \text{subject to} \quad & (\mathbf{L}_{\mathcal{A}} - \mathbf{S}_{\mathcal{A}})\underline{\mathbf{x}}_{\mathcal{A}} = \underline{\mathbf{l}}_{\underline{\mathbf{x}}_{\mathcal{U}}} - \mathbf{L}_{\mathcal{U}}\underline{\mathbf{x}}_{\mathcal{U}}. \end{aligned}$$

Solving the optimization problem to obtain the set of falsified measurement $\underline{\mathbf{x}}_{\mathcal{A}}$ requires the knowledge of the topology of the grid inside the attack area through $\mathbf{L}_{\mathcal{A}}$ and $\mathbf{L}_{\mathcal{U}}$ and the past data to estimate the probability distribution of the local smoothness which is in general challenging for the attackers to acquire.

VI. Conclusion

Once an event of cyber attack is detected and located in the smart grid, its characterization and classification are crucial for the prompt mitigation of the damage of the attack as well as for perceiving the intention and strategy of the attacker. In this paper, it has been shown that along with the detection and locating of the cyber stresses in the grid, the features extracted from the local smoothness of the graph signal associated with the electrical attributes of the buses are effective for their characterization and classification. The strength of the local smoothness-based detection technique has been assessed analytically considering the attackers' availability of resources and data.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 2118510.

References

[1] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng and Z. Han, "Detecting Stealthy False Data Injection Using Machine Learning in Smart Grid," in IEEE Systems Journal, vol. 11, no. 3, pp. 1644-1652, Sept. 2017.

[2] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura and P. Vandergheynst, "Graph Signal Processing: Overview, Challenges, and Applications," in Proc. of the IEEE, vol. 106, no. 5, pp. 808-828, May 2018.

[3] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.

[4] S. Itani and D. Thanou, "A Graph Signal Processing Framework for the Classification of Temporal Brain Data," 28th European Signal Processing Conference (EUSIPCO), 2021, pp. 1180-1184.

[5] S. Bloemheuvel, J. V. D. Hoogen and M. Atzmueller, "Graph signal processing on complex networks for structural health monitoring," in Proc. International Conference on Complex Networks and Their Applications, Springer, pp 249–261, 2020.

[6] J. Moreno, S. Quintero, A. Riascos, L. G. Nonato and C. Sanchez, "Homicide Prediction Using Sequential Features from Graph Signal Processing," in Arai K. (eds) Intelligent Computing, Lecture Notes in Networks and Systems, vol 285. Springer, Cham., 2021.

[7] M. A. Hasnat and M. Rahnamay-Naeini, "Reflection of Cyber and Physical Stresses in Smart Grids on their Graph Signals," accepted in ISGT Europe, 2021.

[8] M. A. Hasnat and M. Rahnamay-Naeini, "Detection and locating cyber and physical stresses in smart grids using graph signal processing," arXivpreprint arXiv:2006.06095, 2020.

[9] K. Mahapatra and N. R. Chaudhuri, "Online Robust PCA for Malicious Attack-Resilience in Wide-Area Mode Metering Application," in IEEE Trans. on Power Systems, vol. 34, no. 4, pp. 2598-2610, July 2019.

[10] M. N. Kurt, Y. Yılmaz and X. Wang, "Real-Time Detection of Hybrid and Stealthy cyber attacks in Smart Grid," in IEEE Trans. on Information Forensics and Security, vol. 14, no. 2, pp. 498-513, Feb. 2019.

[11] Md Abul Hasnat and Mahshid Rahnamay-Naeini, "Detecting and locating cyber and physical stresses in smart grids using the k-nearest neighbour analysis of instantaneous correlation of states," IET Smart Grid, Vol. 4, Issue 3, 2021.

[12] A. Xue, F. Xu, J. H. Chow, S. Leng, H. Kong, J. Xu and T. Bi, "Data-driven detection for GPS spoofing attack using phasor measurements in smart grid," International Journal of Electrical Power and Energy Systems, Volume 129, 106883, July 2021.

[13] M. Ganjkhani, S. Fallah, S. Badakhshan, S. Shamshirband, and Kwok-wing Chau. "A Novel Detection Algorithm to Identify False Data Injection Attacks on Power System State Estimation." Energies 12, no. 11, 2019.

[14] M. N. Kurt, O. Ogundijo, C. Li and X. Wang, "Online cyber attack Detection in Smart Grid: A Reinforcement Learning Approach," in IEEE Transactions on Smart Grid, vol. 10, no. 5, pp. 5174-5185, Sept. 2019.

[15] W. Qiu, Q. Tang, Y. Wang, L. Zhan, Y. Liu and W. Yao, "Multi-View Convolutional Neural Network for Data Spoofing cyber attack Detection in Distribution Synchrophasors," in IEEE Transactions on Smart Grid, vol. 11, no. 4, pp. 3457-3468, July 2020.

[16] O. Boyaci, M. R. Narimani, K. Davis, M. Ismail, T. J. Overbye and E. Serpedin, "Joint Detection and Localization of Stealth False Data Injection Attacks in Smart Grids using Graph Neural Networks", arXiv preprint arXiv:2104.11846.

[17] T. R. Nudell, S. Nabavi and A. Chakrabortty, "A Real-Time Attack Localization Algorithm for Large Power System Networks Using Graph-Theoretic Techniques," in IEEE Transactions on Smart Grid, vol. 6, no. 5, pp. 2551-2559, Sept. 2015.

[18] L. Dan, N. Gebraeel, K. Paynabar and A. P. Meliopoulos, "An Online Approach to Cyberattack Detection and Localization in Smart Grid," arXiv preprint arXiv:2102.11401, 2021.

[19] Z. s. Khalafi, M. Dehghani, A. Khalili, A. Sami, N. Vafamand and T. Dragicevic, "Intrusion Detection, Measurement Correction, and Attack Localization of PMU Networks," in IEEE Transactions on Industrial Electronics, 2021.

[20] R. Ramakrishna and A. Scaglione, "Grid-Graph Signal Processing (Grid-GSP): A Graph Signal Processing Framework for the Power Grid," in IEEE Transactions on Signal Processing, vol. 69, pp. 2725-2739, 2021.

[21] A. Kroizer, Y. C. Eldar and T. Routtenberg, "Modeling and Recovery of Graph Signals and Difference-Based Signals," IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2019, pp. 1-5.

[22] E. Drayer and T. Routtenberg, "Detection of False Data Injection Attacks in Smart Grids Based on Graph Signal Processing," in IEEE Systems Journal, vol. 14, no. 2, pp. 1886-1896, June 2020.

[23] J. Shi, B. Foggo, X. Kong, Y. Cheng, N. Yu and K. Yamashita, "Online Event Detection in Synchrophasor Data with Graph Signal Processing," IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), 2020, pp. 1-7.

[24] A. Dwivedi and A. Tajer, "Scalable Quickest Line Outage Detection and Localization via Graph Spectral Analysis," in IEEE Transactions on Power Systems, 2021.

[25] M. Daković, L. Stanković and E. Sejdić. "Local smoothness of graph signals," Mathematical Problems in Engineering 2019.

[26] Electrical and Computer Engineering Department, IEEE 118-Bus,54 Unit, 24-Hour System Unit and Network Data, Illinois Institute of Tech.

[27] R. D. Zimmerman, C. E. Murillo-Sánchez and R. J. Thomas, "MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education," in IEEE Transactions on Power Systems, vol. 26, no. 1, pp. 12-19, Feb. 2011

[28] The New York Independent System Operator, Inc[US], https://www.nyiso.com/.

[29] SciKit-learn, https://scikit-learn.org/stable/.