

Model-Agnostic Algorithm for Real-Time Attack Identification in Power Grid using Koopman Modes

Sai Pushpak Nandanoori*, Soumya Kundu*, Seemita Pal[†], Khushbu Agarwal[‡] and Sutanay Choudhury[‡]

*Optimization and Control Group

[†]Electricity Security Group

[‡]Data Sciences Group

Pacific Northwest National Laboratory, USA

Emails: {saipushpak.n, soumya.kundu, seemita.pal, khushbu.agarwal, sutanay.choudhury}@pnnl.gov

Abstract—Malicious activities on measurements from sensors like Phasor Measurement Units (PMUs) can mislead the control center operator into taking wrong control actions resulting in disruption of operation, financial losses, and equipment damage. In particular, false data attacks initiated during power systems transients caused due to abrupt changes in load and generation can fool the conventional model-based detection methods relying on thresholds comparison to trigger an anomaly. In this paper, we propose a Koopman mode decomposition (KMD) based algorithm to detect and identify false data attacks in real-time. The Koopman modes (KMs) are capable of capturing the nonlinear modes of oscillation in the transient dynamics of the power networks and reveal the spatial embedding of both natural and anomalous modes of oscillations in the sensor measurements. The Koopman-based spatio-temporal nonlinear modal analysis is used to filter out the false data injected by an attacker. The performance of the algorithm is illustrated on the IEEE 68-bus test system using synthetic attack scenarios generated on GridSTAGE, a recently developed multivariate spatio-temporal data generation framework for simulation of adversarial scenarios in cyber-physical power systems.

Index Terms—cyber-physical security; Koopman mode decomposition; false data injection; online attack identification.

I. INTRODUCTION

Modern electrical grids are increasingly complex cyber-physical systems, with advanced sensing, control and communication layers overlaid on nonlinear dynamical networks. Real-time grid operations are aided by closed-loop feedback control decisions that rely on

advanced sensor measurements to dynamically secure the balance between supply and demand. For instance, the *automatic generation control* (AGC), illustrated in Fig. 1, is designed to process measurements from advanced sensors such as *phasor measurement units* (PMUs) to update the governor set-points of the generators every few seconds. Such tightly coupled cyber-physical control operation arguably presents an opportunity for malicious cyber agents to sneak into the system and cause disruptions, equipment damages and/or financial losses. Real-world events such as the *Stuxnet* [1], the *Dragonfly* [2], and the 2015 cyber-attacks on the Ukrainian power grid [3] and on a substation in the state of California in US [4], demonstrate the vulnerability of power systems worldwide to cyber attacks. Different, possibly overlapping, classes of cyber-attacks have been investigated in the power systems community, including *denial of service* (DoS) attacks [5], [6]; *gray-hole* or packet drop attacks [7]; *jamming* or link-failure attacks [5], [8]; and *false data injection* (FDI) or *data integrity* attacks [5], [6], [9]–[11].

Related Work Due to economic and other reasons, sensors are arguably among the least-protected components in a power network [12] and, therefore, particularly vulnerable to malicious cyber attacks. As such, we focus in this work on identification of FDI attacks on PMUs which can report GPS time-stamped measurements at typically 20-60 samples per second [7] and allows visibility into the system transient dynamics. A broad category of research efforts that considers power systems transient dynamics while detecting and identifying sensor attacks relies on physics-based models, coupled with a dynamic state estimator (e.g. Kalman filters), to detect anomalous sensors data via outlier detection methods such as *cosine similarity test* or statistical χ^2 -test [13]–[15]. In [15], a state-space transformation and

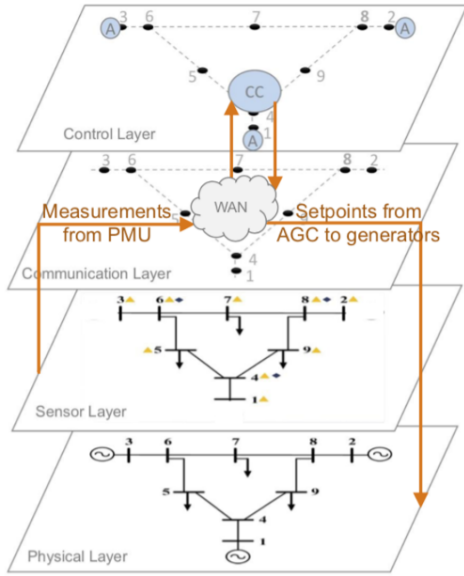


Fig. 1. Illustration of a closed-loop power system, including the controller (e.g. AGC), sensors (e.g. PMUs), actuators (e.g. generators) and the communication layer (via WAN).

decomposition technique was proposed to distinguish between the anomalies caused due to malicious attacks and physical changes (e.g. changes in topology, load, generation). In [16], the authors extended the conventional *static* outlier detection method of evaluating one sample of measurements at a time to a *dynamic* change-point detector based on the *cumulative sum* (CUSUM) test on a moving window of measurements. The work in [17] presents an attack detection algorithm that involves superimposing *watermarked* signals on the control inputs transmitted over secure channels. In addition to the physics-based methods mentioned above, machine learning-based online attack detection algorithms have been proposed in [18], [19].

Approach Most of the existing detection methods are often unreliable during system transients due to unforeseen changes or insufficient training data. In addition, the system dynamics is expected to display rich nonlinear behavior during the transients. In this work, we propose an online attack detection algorithm which uses the Koopman mode decomposition (KMD) to identify the spatial embedding of anomalous (and natural) modes in the sensors data. The *Koopman operator* is a linear, infinite-dimensional, composition operator that describes the temporal evolution of a set of observable functions along the trajectories of a finite-dimensional (nonlinear) dynamical system [20], [21]. Spectral analysis of the Koopman operator reveals that single-frequency modes,

henceforth termed as the *Koopman modes* (KMs), can be embedded in the spatio-temporal dynamics of the nonlinear systems [22]. The KMs are particularly promising in power systems applications, by way of extending the notions of linear modal analysis (e.g. participation factors) to the nonlinear regime in a computationally efficient manner in contrast to other methods such as normal forms [23], [24]. Efficient algorithms exist to compute KMs, as well as the Koopman operator from streaming data in power systems [25], [26]. Koopman modal analysis has found several applications in power systems, including coherency detection, stability, and partitioning (see [21] and the references therein for details). As a novel contribution of this work, we extend the Koopman modal analysis to the case of real-time identification (detection and localization) of malicious data attacks on power system sensors, in presence of transient fluctuations.

Contributions In this paper, we develop an on-line attack identification algorithm that performs a Koopman-based spatio-temporal nonlinear modal analysis on streaming data from the sensors to detect and localize malicious activities. Performance of the algorithm is illustrated on IEEE 68-bus test-case using synthetic attack data generated using GridSTAGE [27] which is a simulations-based framework (built upon [28] and [29]) to synthesize adversarial scenarios in a cyber-physical power grid. The rest of the article is structured as follows: in Sec. II, we present the attack identification problem in a closed-loop power systems; Sec. III briefly recaps the KMD, while the proposed attack identification algorithm is described in Sec. IV; Sec. V illustrates the working of the algorithm with numerical simulation results; before the article is concluded in Sec. VI with a summary and future plan.

II. PROBLEM DESCRIPTION

The power grid is a nonlinear electro-mechanical dynamical system primarily driven by the complex interactions of various synchronous generator models, excitation systems, power system stabilizers, governor controls, as well as dynamic load modulation [28], [29]. Since the attack identification algorithm does not require any knowledge of the system model, it suffices to consider only a compact abstraction of the system dynamics in the control-affine form of:

$$x_{k+1} = f(x_k) + h(x_k) u_k \quad (1a)$$

$$y_k = g(x_k) \quad (1b)$$

where x_k are the vector-valued internal states (e.g. voltage angles, voltage magnitudes, and frequencies); y_k are the vector-valued measurements from sensors (e.g. PMUs); u_k are the vector-valued control inputs to the generators (e.g. governor set-points); f , g and h are locally Lipschitz vector-valued functions of states representing the system dynamics, measurements, and control channels respectively. For notational convenience, we assume that the nominal (target) operating point is shifted to the origin (i.e. $f(0)=0$), and that $g(0)=0$.

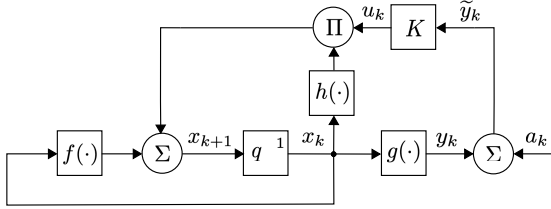


Fig. 2. Block diagram of the closed-loop system under attack.

Attack Model: The adversary is assumed to have compromised one or more of the PMUs, *phasor data concentrators*, network routers and/or communication links. Note that since IEEE Standard C37.118-2 places no restriction on the choice of the communication medium, portions of the PMU communication network may include the internet which is vulnerable to cyber-attacks. At each of the compromised nodes in the communication network, the adversary can manipulate PMU data which is typically not encrypted. The control center receives the manipulated measurements data \tilde{y}_k , with $a_k := \tilde{y}_k - y_k$ being the vector-valued *injected attack signal*. The control inputs are computed using a proportional output-feedback controller with control gain K , such that the closed-loop system becomes:

$$x_{k+1} = f_c(x_k) + h_c(x_k) a_k \quad (2a)$$

$$\tilde{y}_k = g(x_k) + a_k \quad (2b)$$

where $h_c(x_k) := h(x_k)K$ and $f_c(x_k) := f(x_k) + h_c(x_k)g(x_k)$. Fig.2 illustrates the closed-loop system including the control inputs and the attack signals. Note that we do not place any restriction on the type of attack, and merely interpret the attack signal (a_k) as any maliciously introduced anomaly between the actual measurements (y_k) and the measurements received at the control center (\tilde{y}_k).

Attack Identification: Attack identification algorithms involve construction of vector-valued *measurement residuals* (r_k) as a difference between the mea-

surements received at the control center (\tilde{y}_k) and the estimated measurements (\hat{y}_k):

$$r_k := \tilde{y}_k - \hat{y}_k. \quad (3)$$

The measurement residuals are fed into some anomaly detector, such as *cosine similarity* test, χ^2 -test, or CUSUM test [13]–[17]. However, the estimates of the measurements are often inaccurate (or, unavailable) during power system transients caused by unforeseen and/or unknown changes in load and generation - potentially misleading the attack detectors into *false positives* (natural events identified as attacks) or *false negatives* (labeling an attack event as benign). While the work in [15] proposes a method to distinguish between anomalies due to attacks and modeling errors, its reliance on physics-based linear system models and determination of multiple thresholds for statistical tests potentially limits its application to large-scale power systems displaying rich nonlinearity.

In this work, we propose a Koopman-based attack identification algorithm which uses a moving window of $n+1$ most recent snapshots of measurements $\{\tilde{y}_{k-n}, \dots, \tilde{y}_{k-2}, \tilde{y}_{k-1}, \tilde{y}_k\}$, $n \geq 2$, to identify in real-time the sensors under attack.

III. BACKGROUND: KOOPMAN MODE DECOMPOSITION

Consider the discrete-time closed-loop system (2) under the scenario when there is no attack, i.e. $a_k=0 \forall k$,

$$(\text{no attack}) \quad x_{k+1} = f_c(x_k) \quad (4)$$

The *Koopman operator* \mathcal{U} , [21], [23], is an infinite-dimensional linear operator that acts on any scalar-valued function of the state, ψ , and maps it into another scalar-valued function of the state $\mathcal{U}\psi$ defined as follows:

$$\mathcal{U}\psi := \psi \circ f_c. \quad (5)$$

The linearity of the operator is exemplified by the fact that for any two scalars α_1 and α_2 , and any two scalar-valued functions of state ψ_1 and ψ_2 , we have

$$\mathcal{U}(\alpha_1\psi_1 + \alpha_2\psi_2) = \alpha_1(\mathcal{U}\psi_1) + \alpha_2(\mathcal{U}\psi_2)$$

The Koopman operator \mathcal{U} admits (infinitely many) Koopman eigenfunctions ϕ_j (non-zero scalar-valued functions of state) and Koopman eigenvalues $\lambda_j \in \mathbb{C}$ defined as follows:

$$\mathcal{U}\phi_j = \lambda_j\phi_j, \quad j = 1, 2, 3, \dots \quad (6)$$

Note that iterative application of (5) and (6) yields

$$\phi_j(x_k) = \mathcal{U}^k \phi_j(x_0) = \lambda_j^k \phi_j(x_0). \quad (7)$$

The set of all Koopman eigenvalues (λ_j) forms the *discrete spectrum* of the Koopman operator \mathcal{U} and, along with its continuous spectra, define the spectral properties of \mathcal{U} [21].

Koopman Mode Decomposition: The Koopman operator translates the finite-dimensional nonlinear dynamics into an infinite-dimensional linear representation. In particular, let us consider the measurement function $g = (g_1, g_2, \dots, g_p)^T$ as a p -dimensional vector-valued function of observables. If each g_i lies within the span of the Koopman eigenfunctions ϕ_j , then the vector-valued observables g can be expanded as:

$$g(x_k) = \sum_{j=1}^{\infty} \phi_j(x_k) v_j = \sum_{j=1}^{\infty} \lambda_j^k \phi_j(x_0) v_j \quad (8)$$

where v_j are p -dimensional vector-valued coefficients of the expansion, called the *Koopman modes* (KMs) [21]. The expansion in (8) is referred to as the *Koopman Mode Decomposition* (KMD). Moreover, λ_j encode the temporal signatures in the dynamics, with associated spatial signatures encoded in $\phi_j(x_0) v_j$. In particular, the magnitude and phase of the Koopman eigenvalue λ_j are called, respectively, the *growth rate* and the *frequency* of the associated KM.

Empirical Computation: There exist efficient algorithms to compute finite-sum approximations of the KMD (8) based on, for example, the Arnoldi algorithm [21], [22] and the extended dynamic mode decomposition [23], [30]. While we avoid a description of the methods here for space constraint, it is sufficient to note that the algorithms use a sequence of $n+1$ vector-valued observations $\{g(x_0), g(x_1), \dots, g(x_{n-1}), g(x_n)\}$, $n \geq 2$, to obtain KMD as:

$$g(x_k) \approx \sum_{j=1}^n \hat{\lambda}_j^k \hat{v}_j \quad (9)$$

where $\hat{\lambda}_j$ are estimates of the Koopman eigenvalues (λ_j), while \hat{v}_j estimate the products of the Koopman eigenfunctions ($\phi_j(x_0)$) and the KMs (v_j). A finite dimensional approximation of the Koopman operator $K \in \mathbb{R}^{p \times p}$ is given as

$$K = K_1 K_2^+, \quad \text{where} \quad (10)$$

$$K_1 = \frac{1}{n} \sum_{k=0}^{n-1} g(x_{k+1}) g(x_k)^T \text{ and } K_2 = \frac{1}{n} \sum_{k=0}^{n-1} g(x_k) g(x_k)^T,$$

and the superscript “ $+$ ” denotes the *pseudo-inverse*, [30]. Notice that the empirical estimate K of

Koopman operator, learnt using $\{g(x_0), \dots, g(x_n)\}$ can be used to predict future observations $\{\hat{g}(x_{n+1}), \hat{g}(x_{n+2}), \hat{g}(x_{n+3}), \dots\}$ as

$$\hat{g}(x_{n+m}) = K^m g(x_n) \quad \forall m = 1, 2, 3, \dots \quad (11)$$

The empirical computational steps (9)-(11) will be used to design the attack identification algorithm in Sec. IV.

IV. PROPOSED ATTACK IDENTIFICATION ALGORITHM

Fig. 3 depicts key steps involved in the proposed on-line attack identification algorithm combining KMD with spectral clustering, as explained below. For identifying an attack at time k , a moving window sequence of $n+1$ observations are used, which is split into two sub-sequences: a (longer) *learning window* $\mathcal{L} \in \{k-n, \dots, k-\tilde{n}-1\}$, and a (shorter) *prediction window* $\mathcal{P} \in \{k-\tilde{n}, \dots, k\}$, for some $\tilde{n} < n$.

Step 1 Prediction: We identify the finite-dimensional empirical Koopman operator K on the *learning window* \mathcal{L} from (10) and using the empirical Koopman operator to predict the observations (\hat{g}) within the *prediction window* \mathcal{P} from (11).

Step 2 Koopman Mode Analysis: We compute the *error sequence* (\tilde{g}) by subtracting the predicted sequence (\hat{g}) from the actually observed sequence during the *prediction window* \mathcal{P} . Next the empirical KMs \hat{v}_j are computed using (9) from the *error sequence* (\tilde{g}).

Step 3 Spectral Clustering The KMs identified in Step 2 encode the latent spatio-temporal signatures within the error sequence required to identify the adversarial attack. The intuitive idea is that if there is an attack during the *prediction window* \mathcal{P} then it should leave certain anomalous signature within the KMs specific to the attacked sensor measurements. On the other hand, if there is a natural event (sudden load or generation change), that would excite similar modes in the physically neighboring nodes reflecting as non-anomalous deviations within the KMs. This allows differentiating between adversarial attacks and natural events. To facilitate the identification of anomalous sensor measurements, we perform a spatio-temporal normalization on the KMs \hat{v}_j . First, we stack up the absolute values of the KMs into a $p \times \tilde{n}$ matrix:

$$\hat{v} := [\|\hat{v}_1\| \quad \|\hat{v}_2\| \quad \dots \quad \|\hat{v}_{\tilde{n}}\|] \quad (12)$$

Then each column of \hat{v} is normalized to unity-sum, which effectively puts *equal* weight to each KM or, equivalently, disregards the impact of the associated Koopman eigenvalues. Next, each row in the resulting

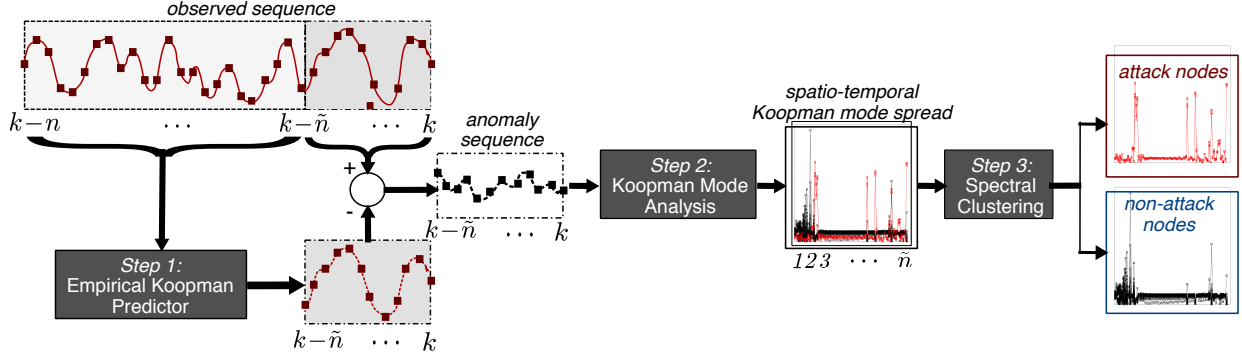


Fig. 3. Step-wise schematic of the proposed online attack identification algorithm.

Algorithm 1 Koopman-Based Online Attack Identification

```

Choose  $n$ 
Choose  $\tilde{n} < n$ 
Set  $k \geq n$ 
while true do
    Set learning window  $\mathcal{L} = \{k-n, \dots, k-\tilde{n}-1\}$ 
    Set prediction window  $\mathcal{P} = \{k-\tilde{n}, \dots, k\}$ 
    procedure STEP 1 PREDICTION
        Estimate  $K$  from (10) using the observation
        sequence  $\{g(x_l)\}_{l \in \mathcal{L}}$  (i.e. the learning window)
        Use  $K$  to predict, as per (11), the observation
        sequence  $\{\hat{g}(x_l)\}_{l \in \mathcal{P}}$  (i.e. the prediction window)
    end procedure
    procedure STEP 2 KOOPMAN MODE ANALYSIS
        Compute error sequence  $\tilde{g}(x_l) = g(x_l) - \hat{g}(x_l)$ 
        for every  $l \in \mathcal{P}$  (i.e. the prediction window)
        Compute the empirical KMs from the error
        sequence  $\{\tilde{g}(x_l)\}_{l \in \mathcal{P}}$  using (9)
    end procedure
    procedure STEP 3 SPECTRAL CLUSTERING
        Spatio-temporally normalize KMs
        Perform spectral clustering on the normalized
        KMs by using the weighted adjacency matrix
    end procedure
     $k \leftarrow k + 1$ 
end while

```

matrix is normalized to unity-sum, which effectively puts *equal* weight to each sensor irrespective of whether it is from a generator or a load. This results in the spatio-temporally normalized KMs stacked up into \bar{v} . Each row in \bar{v} sums up to unity, and represent a sensor.

Finally a spectral clustering is performed to identify the anomalous sensors from the spatio-temporally nor-

malized KMs. Since each row in \bar{v} , corresponding to a sensor, can be seen as a *probability mass function*, we use the KullbackLeibler divergence between the (sensor) rows to construct the weighted adjacency matrix which is then fed to *k-means clustering algorithm* [31] to identify the dominant (attacked and non-attacked) clusters in the normalized KMs.

The key algorithmic steps are summarized in Algorithm 1.

V. NUMERICAL RESULTS

We begin this section by describing the adversarial data generation framework GridSTAGE, and then present an adversarial use-case illustrating the application of proposed Koopman-based attack identification algorithm.

GridSTAGE (*Spatio-Temporal Adversarial scenario Generation*) To evaluate the performance of the closed-loop power system under several attack strategies and mitigation efforts, we developed *GridSTAGE* which is a multivariate spatio-temporal data generation framework for simulation of adversarial scenarios in cyber-physical systems [27]. There is no existing adversarial data-generation framework that can incorporate several attack characteristics and yield adversarial PMU data. GridSTAGE is available as open source from GitHub.

GridSTAGE models the cyber-physical system of the power grid, simulates adversarial scenarios in the system and generates multi-variate, spatio-temporal network data. GridSTAGE is developed based on Power System Toolbox (PST) on which nonlinear time-domain simulations can be generated for standard IEEE bus systems. Using GridSTAGE, one can create several event scenarios by enabling or disabling any of the following: faults, AGC control, PSS control, exciter control, load changes, and different types of cyber-attacks. IEEE bus system data is used for defining the power system environment.

Sensors in the power system include both PMU and SCADA, and simulated data are generated for both type of sensors. Rate of frequency and location of the sensors can be adjusted as well. Detailed instructions on generating several scenarios with different attack characteristics, load characteristics, sensor configuration, control parameters can be found in the github repository [27]. The GridSTAGE framework currently supports simulation of *FDI* attacks (such as ramp, step, random, trapezoidal, multiplicative, replay) and *DoS* attacks (such as time-delay, packet-loss, freezing) on PMU data [32]. We use one of these attacks to demonstrate the proposed attack identification algorithm, while more extensive analysis on different types of attacks will be performed in future.

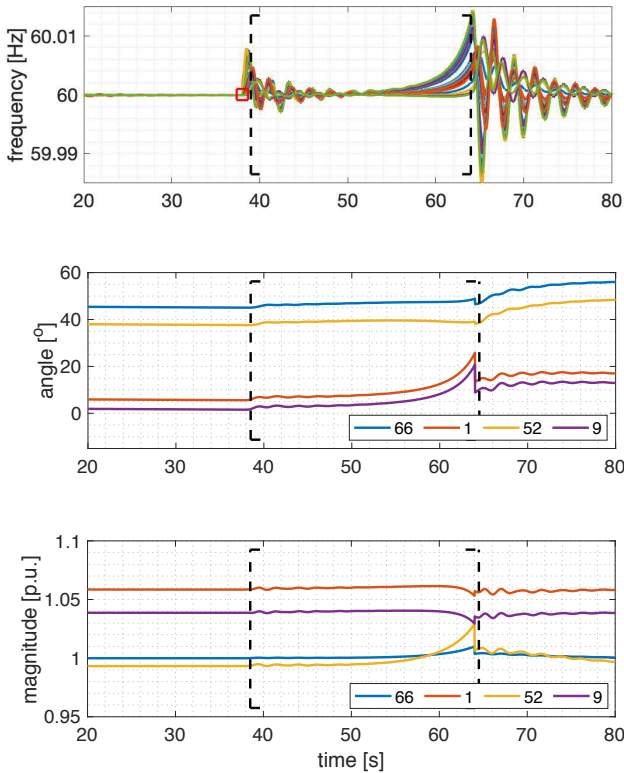


Fig. 4. Time-series closed-loop simulation (using GridSTAGE) of an adversarial attack scenario. Bottom two plots show the *attacked* PMU measurements of voltage angle and magnitude at the buses 1, 9, 52 and 66 in IEEE 68-bus network. Top plot shows the (delayed) impact of the attack on the resulting system frequencies. The attack duration (shown in black dashed lines) are between 39s and 64s, while a natural load change happens in bus 23 (marked by red square) at 38s.

Identifying Multiplicative Attack on PMUs: In this subsection, we briefly discuss a multiplicative attack on the PMU data using the IEEE 68-bus as example. In this attack scenario, the attacker strategically initiates an *FDI* attack *soon after* there is a load change in the system,

by injecting a signal that grows over time in proportion to the disturbance being measured at the sensor. The attacker's goal is to *hide behind* the existing transient disturbances in the network (caused due to load change). Based on the attack characteristics, this introduces a delayed impact on the system measurements and especially when the attack is over, the system sees a sudden change in measurements which engages control resources to counter the disturbance. Ideally, these resources should not have been engaged. The motivation of this work is to identify such attacks in their early stage such that disturbance seen by the system is deemed as attack and additional resources are not engaged. In Fig. 4, the attack scenario is illustrated, which is initiated at 39s right after a natural load change which happens at 38s. Finally Fig. 5 illustrates a successful application of the online attack identification algorithm which is able to correctly identify all the attack locations (buses 1, 9, 52 and 66) within 1s of the attack initiation, by performing a spectral clustering on the normalized Koopman mode spreads of all the sensors (top plot). It is evident that the Koopman mode spread for the attacked sensors (middle plot) are distinctively different from those at the other locations (bottom plot), and hence successfully picked out by the clustering algorithm.

VI. CONCLUSION

This work introduced a novel real-time data-driven attack detection algorithm based on the Koopman mode decomposition and spectral clustering. Koopman modes embed the time-series streaming data (such as from PMUs) from the power grid network into spatio-temporal modes. A *k-means* based spectral clustering on the spatio-temporally normalized KMs identifies the attacked and non-attacked clusters of sensors. The proposed attack detection algorithm is illustrated on a multiplicative attack on the PMU measurements representing the bus voltage angle and magnitudes generated using the GridSTAGE, a spatio-temporal multivariate adversarial data generation platform that is developed in the scope of this work. This induced attack if allowed to propagate for a few seconds, introduces unintended frequency changes in the network and engages resources without the need. However, the proposed algorithm successfully detects the induced attack within 1 second of attack initiation in the presence of load changes in the network. Future scope of this work will explore the application of the proposed attack detection algorithm on wider attack scenarios, including various strategic and stealthy attack types, as well as applicability of the method in

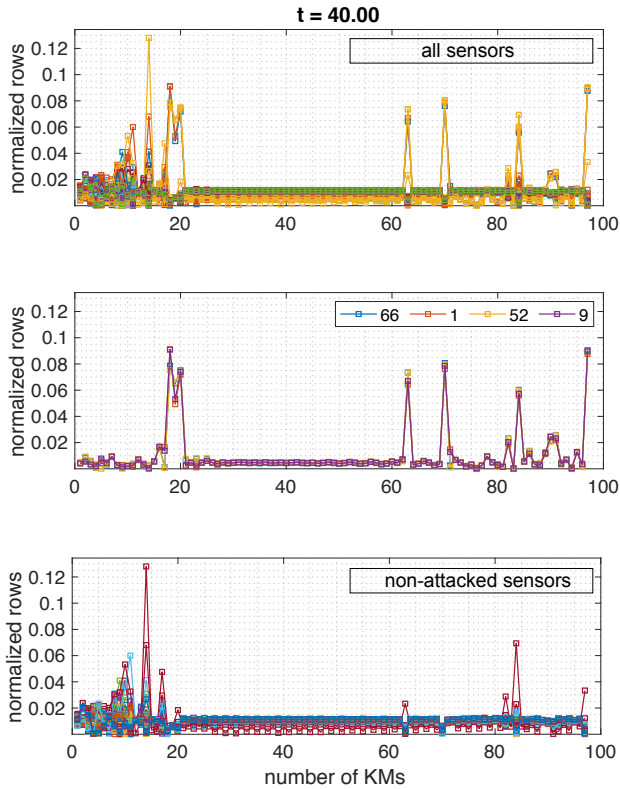


Fig. 5. Koopman-based modal analysis followed by spectral clustering separates out the attacked sensors from the non-attacked one, even during transients. The attack is correctly identified at 40s, within 1s of the attack initiation.

differentiating between naturally caused events such as faults and component failures.

ACKNOWLEDGMENT

The authors thank Jiangmeng Zhang and Prof. Alejandro D. Domnguez-Garca at UIUC for providing the AGC code.

REFERENCES

- [1] S. Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security," in *IECON 2011-37th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2011, pp. 4490–4494.
- [2] S. Team, "Dragonfly: Western energy sector targeted by sophisticated attack group," 2017.
- [3] D. U. Case, "Analysis of the cyber attack on the ukrainian power grid," *Electricity Information Sharing and Analysis Center*, vol. 388, 2016.
- [4] S. Reilly and R. Sabalow, "Bracing for a big power grid attack:one is too many," *USA Today*, vol. 24, 2015.
- [5] S. Pushpak, A. Diwadkar, M. Fardad, and U. Vaidya, "Vulnerability analysis of large-scale dynamical networks to coordinated attacks," in *2014 4th Australian Control Conference*. IEEE, 2014, pp. 89–94.
- [6] M. N. Kurt, Y. Yilmaz, and X. Wang, "Distributed quickest detection of cyber-attacks in smart grid," *IEEE Trans. Information Forensics and Security*, vol. 13, no. 8, pp. 2015–2030, 2018.
- [7] S. Pal, B. Sikdar, and J. H. Chow, "An Online Mechanism for Detection of Gray-Hole Attacks on PMU Data," *IEEE Trans. Smart Grid*, vol. 9, pp. 2498–2507, 2018.
- [8] M. N. Kurt, Y. Yilmaz, and X. Wang, "Real-time detection of hybrid and stealthy cyber-attacks in smart grid," *IEEE Trans. Information Forensics and Security*, vol. 14, no. 2, pp. 498–513, 2018.
- [9] R. B. Bobba, K. M. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T. J. Overbye, "Detecting false data injection attacks on DC state estimation," in *Workshop on Secure Control Systems, CPSWEEK*, 2010.
- [10] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Information and System Security (TISSEC)*, vol. 14, no. 1, pp. 1–33, 2011.
- [11] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, 2011.
- [12] E. M. Ferragut, J. Laska, M. M. Olama, and O. Ozmen, "Real-time cyber-physical false data attack detection in smart grids using neural networks," in *2017 International Conference on Computational Science and Computational Intelligence (CSCI)*. IEEE, 2017, pp. 1–6.
- [13] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using Kalman filter," *IEEE Trans. Control of Network Systems*, 2014.
- [14] D. B. Rawat and C. Bajracharya, "Detection of false data injection attacks in smart grid communication systems," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1652–1656, 2015.
- [15] M. Ghosal, "Diagnosis of anomaly in the dynamic state estimator of a power system using system decomposition," in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 1–6.
- [16] C. Murguia and J. Ruths, "Characterization of a cusum model-based sensor attack detector," in *IEEE 55th Conference on Decision and Control*. IEEE, 2016, pp. 1303–1309.
- [17] T. Huang, B. Satchidanandan, P. Kumar, and L. Xie, "An online detection framework for cyber attacks on automatic generation control," *IEEE Trans. Power Systems*, vol. 33, no. 6, pp. 6816–6827, 2018.
- [18] P.-Y. Chen, S. Choudhury, and A. O. Hero, "Multi-centrality graph spectral decompositions and their application to cyber intrusion detection," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016.
- [19] M. N. Kurt, O. Ogundijo, C. Li, and X. Wang, "Online cyber-attack detection in smart grid: A reinforcement learning approach," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5174–5185, 2018.
- [20] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," *Proceedings of the national academy of sciences of the united states of america*, vol. 17, no. 5, p. 315, 1931.
- [21] Y. Susuki, I. Mezic, F. Raak, and T. Hikiyara, "Applied Koopman operator theory for power systems technology," *Nonlinear Theory and Its Applications, IEICE*, vol. 7, no. 4, pp. 430–459, 2016.
- [22] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," *Journal of fluid mechanics*, vol. 641, pp. 115–127, 2009.
- [23] M. Netto, Y. Susuki, and L. Mili, "Data-driven participation factors for nonlinear systems based on Koopman mode decom-

- position,” *IEEE control systems letters*, vol. 3, no. 1, pp. 198–203, 2018.
- [24] J. J. Sanchez-Gasca, V. Vittal, M. Gibbard, A. Messina, D. Vowles, S. Liu, and U. Annakkage, “Inclusion of higher order terms for small-signal (modal) analysis: committee report-task force on assessing the need to include higher order terms for small-signal (modal) analysis,” *IEEE Trans. Power Systems*, vol. 20, no. 4, pp. 1886–1904, 2005.
 - [25] E. Barocio, B. C. Pal, N. F. Thornhill, and A. R. Messina, “A dynamic mode decomposition framework for global power system oscillation analysis,” *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 2902–2912, 2014.
 - [26] S. Sinha, S. P. Nandanoori, and E. Yeung, “Data driven online learning of power system dynamics,” *arXiv preprint arXiv:2003.05068*, 2020.
 - [27] “GridSTAGE: SpatioTemporal Adversarial Scenario Generation framework,” Mar. 2020. [Online]. Available: <https://github.com/pnnl/GridSTAGE>
 - [28] J. Zhang and A. D. Domínguez-García, “Augmenting the power system toolbox: Enabling automatic generation control and providing a platform for cyber security analysis,” in *2016 North American Power Symposium (NAPS)*. IEEE, 2016, pp. 1–5.
 - [29] J. H. Chow and K. W. Cheung, “A toolbox for power system dynamics and control engineering education and research,” *IEEE Trans. Power Systems*, vol. 7, no. 4, pp. 1559–1564, 1992.
 - [30] S. Klus, P. Koltai, and C. Schütte, “On the numerical approximation of the Perron-Frobenius and Koopman operator,” *arXiv preprint arXiv:1512.05997*, 2015.
 - [31] A. Y. Ng, M. I. Jordan, and Y. Weiss, “On spectral clustering: Analysis and an algorithm,” in *Advances in neural information processing systems*, 2002, pp. 849–856.
 - [32] Y.-L. Huang, A. A. Cárdenas, S. Amin, Z.-S. Lin, H.-Y. Tsai, and S. Sastry, “Understanding the physical and economic consequences of attacks on control systems,” *International Journal of Critical Infrastructure Protection*, vol. 2, no. 3, pp. 73–83, 2009.