Identification of Destabilizing Attacks in Power Systems

Mike Izbicki, Sajjad Amini, Christian R. Shelton, and Hamed Mohsenian-Rad

***Abstract*— In a destabilizing attack against a power system, the adversary hacks into generators or load control mechanisms to insert positive feedback into the power system dynamics. The implementation of destabilizing attacks, both on the generation and load sides, have recently been studied. There are also recent advances on how to detect, i.e., realize the presence of, destabilizing attacks in power systems. However, identifying the location(s) of the compromised buses is still an open problem. This is particularly challenging if, as in practice, one does not even know the number of compromised buses. Another challenge is to keep the computational complexity low to allow fast attack identification with high accuracy. To address these various issues, we observe in this paper that destabilizing attacks can be modeled as a reparameterization of the power system's dynamical model. Therefore, we propose an attack detection method that uses the unscented Kalman filter to jointly estimate both the system states and parameters of the attack. We also propose a low-rank modification to the Kalman filter that improves computational efficiency while maintaining the detection accuracy. We show empirically that this method successfully identifies complex attacks involving many buses.**

I. Introduction

In this paper, we consider attacks against power system stability. Such attacks can be conducted in different ways, either on the generation side [3], [10] or on the load side [1], [7], [8]. Either way they essentially involve inserting positive feedback into various power system control mechanisms.

Recent work has made advances on how to *detect* destabilizing attacks in power systems. Specifically, [2] applies the fast Fourier transform (FFT) to system measurements. It detects frequencies that a destabilizing attack adds that are not present during normal operation. The presence of such new frequencies beyond certain pre-specified magnitude thresholds indicate the presence of a destabilizing attack.

In this paper, we move one step ahead and address the open problem of *identifying* destabilizing attacks against power systems by only monitoring state variables. That is, we devise a method that examines the state-variable data from power system sensors such as phasor measurement units (PMUs), to indicate at which exact power system buses (i.e., nodes) the load and/or generation are compromised.

Our proposed method has three main properties:

1) It does not require prior knowledge of the number of buses that are compromised. That is, as in practice, we assume that the grid operator is not initially aware of how many buses are compromised. Nevertheless, our method can identify which buses are compromised.
2) Prior methods, e.g., in [2], often analyze the attack at each individual bus, which would allow attack identification only if the method is applied several times separately on every bus. In contrast, our method naturally identifies attacks on the entire system considered as a whole. This reduces the computation time.
3) It is capable of distinguishing destabilizing attacks, i.e., load or generation control loops that are malicious and based on positive feedback, from the many load and generation control loops that exist in a power system that are benign and based on negative feedback.

The main tool that we use in this paper is the unscented Kalman filter (UKF) to perform dual state estimation to estimate an unknown attack matrix. We then identify the attack location(s) through a proper thresholding mechanism applied to the entries of the estimated attack matrix.

A. Related work

This paper belongs to the family of studies that address destabilizing attacks against power systems. While the focus so far has been mostly on the definition and implementation of such attacks [1], [3], [8], on the methods to protect the power system against such attacks [7], [10], and occasionally on methods to detect such attacks [2], the focus in this paper is on the less explored topic of identifying the attack by finding the location(s) of the compromised buses.

In terms of the methodology used in this paper, the UKF has been used before for power system problems, e.g., to estimate the rotor angle and speed in synchronous generators [4]. However, no joint estimation is used, and the system is not under attack. A more recent analysis estimates the parameters of the motor controller and bus loads [5]. Again, the system is not under attack.

This paper shows through numerical results that the current power system monitoring systems would require calculation of a large Jacobian matrix if one wants to apply the UKF without modification for the purpose of identifying destabilizing attacks. In this regard, the current study is related also to a thread of work, such as in [11], [14], that similarly have to deal with the computational issues that arise when applying the UKF to power systems. Again, these studies do not address the power system under attack and the details of the analysis are different from those in this paper.

II. Problem statement

We consider a power system with g generator buses and ℓ load buses. An example is shown in Figure 1. We model the dynamics of this system using the standard linear power

This work was supported in part by the National Science Foundation (NSF) grant ECCS/AFOSR 1462530. The authors are with the Bourns College of Engineering, University of California, Riverside, CA, USA. Emails: mike@izbicki.me, saamini@ece.ucr.edu, cshelton@cs.ucr.edu, hamed@ece.ucr.edu.

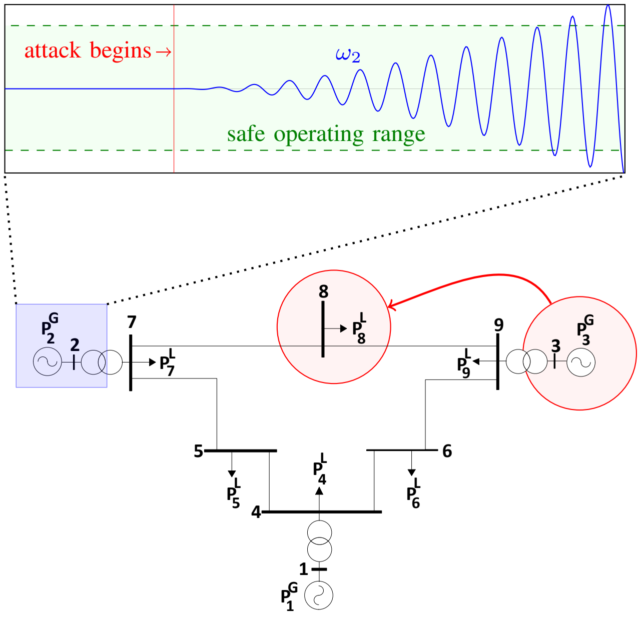

Fig. 1. The standard IEEE 9 bus power system under a destabilizing attack. The attacker has added a load to bus 8 that gets positive feedback from generator bus 3. This causes instabilities throughout the power system. The inset shows that generator 2's rotar angle frequency deviation (ω_2). During normal operations, ω_2 remains close to 0 because the system is stable. The attack adds instability to the system, which causes ω_2 to exceed safe operating levels unless corrective action is taken. Our goal is to identify which system buses are compromised so that we can take this corrective action.

system state space equation

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}, \tag{1}$$

where the state variables and inputs are

$$\mathbf{x} = \begin{bmatrix} \delta \\ \theta \\ \omega \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} P^G \\ P^L \end{bmatrix}. \tag{2}$$

Here, δ is the vector of voltage phase angles at all generator buses, ω is the vector of rotor angular frequency deviations at all generator buses, θ is the vector of voltage phase angles at all load buses, P^L is the vector of power consumption at all load buses, and P^G is the vector of power generation at all generator buses. Many generators are equipped with Automatic Generation Control (AGC), which is a system for automatically adjusting the power output in response to the load. Entries of P^G associated with generators that have AGC are zero. All other entries are nonzero. The system dynamics matrices are

$$A = \begin{bmatrix} 0 & 0 & I \\ (D^L)^{-1}H^{LG} & (D^L)^{-1}H^{LL} & 0 \\ -M^{-1}(K^I + H^{GG}) & -M^{-1}H^{GL} & -M^{-1}(K^P + D^G) \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 0 \\ 0 & (D^L)^{-1} \\ -M^{-1} & 0 \end{bmatrix},$$

where I is the identity matrix of appropriate dimension, and M, D^G, and D^L are diagonal matrices with diagonal entries equal to the inertia, damping coefficients of generators, and damping coefficients of loads, respectively. Matrices H^{LG}, H^{GG}, H^{GL}, and H^{LL} are the imaginary components of the standard system admittance matrix H^{bus}. Specifically,

$$H^{bus} = \begin{bmatrix} H^{GG} & H^{GL} \\ H^{LG} & H^{LL} \end{bmatrix}.$$

Matrices K^I and K^P are diagonal matrices with diagonal entries equal to the integral and proportional coefficients of the generator controllers with AGC capability. Note that the coefficients corresponding to generators without AGC are zero. This system incorporates the swing equations for generators, power flow equations for the transmission network, and the governor and load frequency controller for generators with AGC.

The focus in this paper is attacks against power system stability. As shown in Figure 1, an attack in one portion of the power system can cause instability throughout the power system. These destabilizing attacks can affect both generation and load sectors. On the generation side, the attack might compromise the generator governor controller directly or the generator's sensors and communications systems. (See [3], [10] for more details.) On the load side, the attack might compromise the direct or indirect load control mechanisms, e.g., in demand response programs, or their associated sensors, or command signals. (See [1], [7], [8] for more details.) In either case, the attacker's goal is to insert positive feedback into the system.

We model a destabilizing attack by decomposing the control input vector $\mathbf{u}$ as

$$\mathbf{u} = \mathbf{u}^n + \mathbf{u}^a. \tag{3}$$

Here, the vector $\mathbf{u}^n$ denotes the control input under normal operation and the vector $\mathbf{u}^a$ represents the control input added by the attacker. We consider the case where the attacker uses a proportional controller to dynamically determine the value of $\mathbf{u}^a$ with feedback taken from the demand side. Loads with this feedback are called *frequency-responsive controllable loads*, and are widely used in practice [9], [12], [16]. We model the proportional controller input vector as

$$\mathbf{u}^a = A^p\mathbf{x} + \mathbf{u}^p, \tag{4}$$

where $\mathbf{u}^p$ is the constant term for the proportional controller and

$$A^p = \begin{bmatrix} 0 & 0 & -(D^L)^{-1}K^{LG} \\ 0 & 0 & 0 \end{bmatrix}. \tag{5}$$

The entry in row i column j of matrix K^{LG} contains the proportional gain for load bus i getting feedback from the frequency of generator bus j. If the entry is positive, then bus i is under attack and the system may destabilize. If the entry is negative, then there is a benign frequency-responsive load. These loads will never destabilize the system. If the entry is zero, there are no frequency-responsive loads. Substituting (4) and (3) into (1) gives us our final system dynamics under attack:

$$\dot{\mathbf{x}} = (A + BA^p)\mathbf{x} + B(\mathbf{u}^n + \mathbf{u}^p). \tag{6}$$

The attack will be destabilizing if the matrix $(A + BA^p)$ has an eigenvalue whose absolute value is greater than 1.

Detecting the presence of an attack is not difficult. It can be done by monitoring only a few state variables [1]. However, identifying which load bus is compromised is an open problem. In this paper, we propose an identification method that directly estimates the K^{LG} matrix. Our method automatically determines which load buses are compromised and can distinguish between destabilizing and benign loads. Our method requires access only to synchronized measurements of the state vector $\mathbf{x}$, and does not require access to the control input $\mathbf{u}$. These state measurements are widely available in existing modern power systems through Phasor Measurement Units (PMUs).

III. Attack Identification Method

Our attack identification procedure has two steps. First we estimate the K^{LG} matrix using *dual state estimation*. This is a standard technique that applies the unscented Kalman filter (UKF) [13] to simultaneously estimate the entries of matrix K^{LG} and the system states $\mathbf{x}$. Unfortunately, the standard application of this technique does not work well for our problem. It is too slow computationally and has poor accuracy. So we introduce a novel rank-1 approximation which lets us effectively apply dual state estimation to our problem. Finally, once K_{LG} is estimated, we apply a thresholding procedure to identify the attacked buses.

A. Standard dual state estimation and its limitations

Dual state estimation is traditionally described using the system's discrete state equations, so we begin our presentation by discretizing (6) as

$$\mathbf{x}_{t+1} = (sA + sBA^p_t + I)\mathbf{x}_t + sB(\mathbf{u}^n_t + \mathbf{u}^p_t) + \epsilon. \quad (7)$$

The subscripts indicate the timestep, s is a scalar that represents the length of a time step, and $\epsilon \sim \mathcal{N}(0, Q^\epsilon)$ is an error term capturing both modeling and observation errors.

Now we describe how to estimate the A^p_t matrix. Recall that in the definition of A^p_t, the K^{LG}_t matrix is unknown and determined by the attacker; all other elements are statically known. In dual state estimation, we augment the original dynamical system's state variables to also include the elements of K^{LG}_t. The resulting augmented system is

$$\begin{bmatrix}\mathbf{x}_{t+1}\\ \operatorname{vec} K^{LG}_{t+1}\end{bmatrix} = \begin{bmatrix} sA + sBA^p_t + I & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}\mathbf{x}_t\\ \operatorname{vec} K^{LG}_t\end{bmatrix} + \begin{bmatrix} sB & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}\mathbf{u}^n_t + \mathbf{u}^p_t\\ \mathbf{u}^m_t\end{bmatrix} + \begin{bmatrix}\epsilon\\ \epsilon^m\end{bmatrix}, \quad (8)$$

where

$$\epsilon \sim \mathcal{N}(0, Q^\epsilon), \quad \epsilon^m \sim \mathcal{N}\left(0, Q^{\epsilon^m}\right) \quad (9)$$

Here we have also introduced a new control input $\mathbf{u}^m_t$ with error ϵ_m. It is unobserved and controlled by the attacker. Specifically, the attacker uses $\mathbf{u}^m_t$ to manipulate the entries of K^{LG}_t, and hence A^p_t. The notation $\operatorname{vec} K^{LG}_t$ refers to the column vector constructed by stacking the columns of K^{LG}_t on top of each other.

Next we note that the control inputs $\mathbf{u}^n_t$, $\mathbf{u}^a_t$, and $\mathbf{u}^m_t$ are unobserved. A standard technique for modeling unobserved inputs is to replace them with random error terms. The true distribution of these random errors is unknown, but for computational convenience we assume they are normally distributed. In particular, we assume the control inputs are zero-mean Gaussians with covariance Q^n, Q^a, and Q^m respectively. Under these assumptions, we can rewrite the dualized system dynamics described in (9) as

$$\begin{bmatrix}\mathbf{x}_{t+1}\\ \operatorname{vec} K^{LG}_{t+1}\end{bmatrix} = \begin{bmatrix} sA + sBA^p_t + I & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}\mathbf{x}_t\\ \operatorname{vec} K^{LG}_t\end{bmatrix} + \begin{bmatrix}\epsilon\\ \epsilon^{KL}\end{bmatrix}, \quad (10)$$

where

$$\epsilon \sim \mathcal{N}(0, B(Q^n + Q^p) + Q^\epsilon),$$
$$\epsilon^{KL} \sim \mathcal{N}(0, Q^m).$$

Observe that the dynamical system described by (10) is nonlinear because the K^{LG} term appears in the definition of A^p_t. It is standard to solve systems of this form using the UKF [13]. We defer to the cited paper for details.

The UKF encounters two problems when run on (10). The first is that the problem is *underspecified*. The number of parameters we are trying to estimate (i.e. the number of entries in K^{LG}) grows as $O(\ell g)$, but the size of the observed data (i.e. the size of $\mathbf{x}$) grows at the slower rate of $O(\ell + g)$. In general, underspecified problems are difficult to solve without introducing additional statistical assumptions. As the size of the power grid increases, the degree of underspecification increases, so we would expect this method to have low accuracy on large grids. The second problem is computational. At each time step, the UKF takes the inverse of a matrix whose dimension depends on the number of state variables. There are $O(\ell g)$ states in (10), and so the runtime of this inversion is $O((\ell g)^3)$. This poor scaling makes the standard method impractical to run on power systems with more than about 50 buses. These limitations of the standard method motivate our proposed rank-1 method, which we describe next.

B. The rank-1 method

In this method, we assume that the K^{LG} matrix has rank 1. We justify this assumption as follows. In a typical destabilizing attack, only a small number of buses are compromised and subject to positive feedback. For each of these compromised buses, there is a corresponding nonzero entry in the K^{LG}_t matrix. A basic fact of linear algebra is that the rank of a matrix is less than or equal to the number of nonzero entries in the matrix. Specifically, we have

$$\text{Rank}\{K^{LG}_t\} \leq \text{Non-zero entries in } K^{LG}_t. \quad (11)$$

Therefore, assuming that there are a small number of compromised buses is equivalent to assuming that K^{LG} has low rank.

Specifically, we assume that

$$K^{LG}_t = \mathbf{k}^L_t {\mathbf{k}^G}^\mathsf{T}_t, \quad (12)$$

where $\mathbf{k}_t^L$ and $\mathbf{k}_t^G$ are column vectors. Under this assumption, we can rewrite the standard method's dynamics from (10) as

$$\begin{bmatrix}\mathbf{x}_{t+1}\\ \mathbf{k}_{t+1}^L\\ \mathbf{k}_{t+1}^G\end{bmatrix} = \begin{bmatrix} sA + sBA_t^p + I & 0 & 0\\ 0 & I & 0\\ 0 & 0 & I\end{bmatrix}\begin{bmatrix}\mathbf{x}_t\\ \mathbf{k}_t^L\\ \mathbf{k}_t^G\end{bmatrix} + \begin{bmatrix}\epsilon\\ \epsilon_1\\ \epsilon_2\end{bmatrix}, \quad (13)$$

where

$$\begin{aligned}\epsilon &\sim \mathcal{N}\left(0, B(Q^n + Q^p) + Q^\epsilon\right),\\ \epsilon_1 &\sim \mathcal{N}\left(0, Q^m + Q^{\epsilon_1}\right),\\ \epsilon_2 &\sim \mathcal{N}\left(0, Q^m + Q^{\epsilon_2}\right),\end{aligned}$$

and the new attack matrix is

$$A_t^p = \begin{bmatrix} 0 & 0 & -(D^L)^{-1}{\mathbf{k}^L}_t^{\mathsf{T}}\mathbf{k}_t^G\\ 0 & 0 & 0\end{bmatrix}.$$

This system remains nonlinear and is solved using the UKF.

The rank-1 method has improved statistical and computational performance. Statistically, there are only $O(\ell + g)$ parameters to estimate in the rank-1 method. This matches the size of the state vector $\mathbf{x}$, so the problem is no longer underspecified. We no longer expect statistical performance to degrade as the problem size increases. Computationally, the run time of each iteration of the UKF is only $O((\ell+g)^3)$. This is much faster than the $O((\ell g)^3)$ required for the standard method.

C. Thresholding

Once the matrix K^{LG} is estimated, we apply a thresholding procedure to identify the attack. Define the function

$$f_t(i) = \sum_{j=1}^{n} K_t^{LG}(i,j) \quad (14)$$

to be the sum of the entries in the ith row of the K_t^{LG} matrix. This value is the total predicted attack on the ith bus in the power grid. Also define

$$\alpha_t = \arg\max_i |f_t(i)| \quad (15)$$

to be the bus we predict has the most compromised load and so is under the heaviest attack. If $f_t(\alpha_t)$ is greater than some threshold τ, then we declare that the system is under attack at bus α_t. At this point, the system operator can take defensive measures such as isolating the bus from the system.

IV. Connection to previous methods

Recall from Section I that destabilizing attacks in power systems can be detected using an FFT-based method applied to the system state variables, as explained in [2]. A destabilizing attack will introduce new frequencies that are not present during normal operation. Whenever these frequencies exceed a pre-specified threshold, we say an attack is happening. One can apply the same type of FFT-based method directly to the input measurements in a power system, i.e., to $\mathbf{u}_t = \mathbf{u}_t^n + A_t^p\mathbf{x}_t + \mathbf{u}_t^p$, to also identify the attack. However, this is not a desirable approach in practice due to several reasons, as we explain next.

First, unlike the state variables that are measured using advanced sensors such as PMUs at high resolutions, e.g., at 60 to 120 samples per second, the input signals, i.e., the load and generation levels, are metered at low resolutions, e.g., once every one to 15 minutes. These low resolution measurements do not allow observing the new frequencies that would be present, e.g., at 0.26 Hz [2], under an attack.

Second, even if we use the high resolution state variable data from PMUs, combine it with an unknown input observer (UIO) method, and then estimate the input signals at high resolutions, we would still have three problems: (i) the high computational cost of the UIO method; (ii) the high computational cost of applying the FFT method to each entry of the estimated input signal; (iii) the inherent weakness of FFT-based methods in distinguishing destabilizing attacks, i.e., load or generation control loops that are malicious and based on positive feedback, from the many load and generation control loops that exist in a power system that are benign and based on negative feedback, c.f. [2].

One of our contributions in this paper is to observe that the A_t^p matrix contains all the information we need to determine whether an attack is occurring, so there is no need to perform the subsequent FFT step. In fact, the FFT step necessarily loses some of the information contained within A_t^p. The FFT method is only able to detect the presence of frequency dependent loads or generators; it cannot identify either the sign or the magnitude of the feedback. Direct inspection of the A_t^p method, on the other hand, gives us both.

V. Simulation Results

In this section we compare the performance of the proposed method in Section III with a baseline approach that does not apply the rank-1 assumption from Section III-B. We refer to the latter method as the *standard* method. In this section, we show that compared to the standard method, our proposed method can:

1) significantly lower the computation time;
2) significantly lower the identification error; and
3) better distinguish positive and negative feedback.

We begin with a qualitative demonstration of these facts, and then conclude with a quantitative demonstration.

A. Test Setup and Qualitative Results

All experiments in this section use a single randomly generated power grid with 20 generator and 20 load buses. We test on this relatively small grid size because the standard method that estimates a full rank K^{LG} matrix cannot scale to larger problems. On this size problem, a single iteration of our rank-1 method takes about 1 second, and a single iteration of the standard method takes about 1 minute. On a problem with 100 generators and 100 loads, a single iteration of our rank-1 method takes about 5 seconds, and a single iteration of the standard method takes over an hour. The computation advantage of our proposed method is evident.

We follow the *clusterSmallWorld* procedure for generating the power grid [15]. Note that, standard methods for generating random graphs do not exhibit the topological

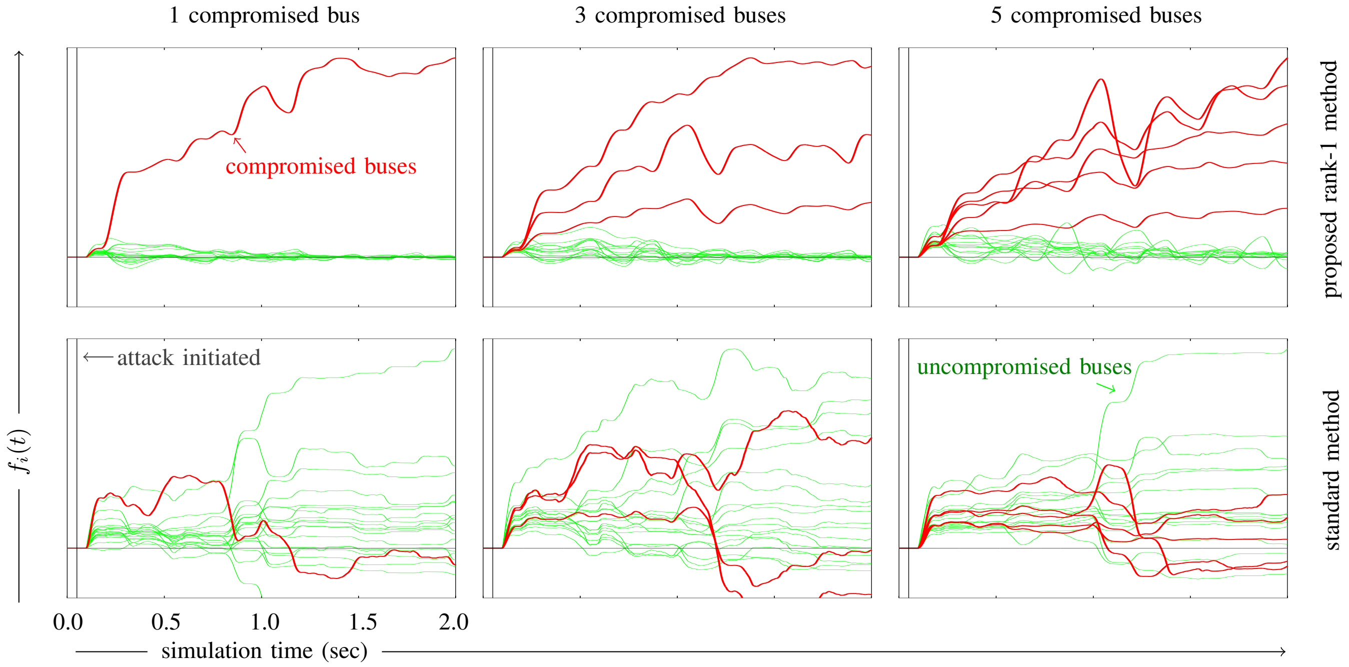

Fig. 2. Each line in the figures above represents the predicted positive feedback of a particular load bus. Compromised buses are drawn in bold red, and uncompromised buses are drawn in thin green. For all times t before the attack begins, each bus i has $f_i(t)$ near zero. After the attack begins, the $f_i(t)$ deviate from zero. Our method is correctly identifying the attacked buses whenever the red lines are above the green lines. In the top row, we see that our rank-1 approximation of K^{LG} provides relatively accurate predictions even when the number of attacks increases and the rank-1 approximation is no longer true. In the bottom row, we see that the standard method has poor accuracy.

and electrical properties of real world power grids [6], but *clusterSmallWorld* was designed specifically for modeling real world power grids. An outline of the procedure is: First generate a random number of ring shaped grids with fewer than 10 buses each; Then randomly add connections between the buses until the average degree of each node is 4. To ensure the stability of the resulting system, scale matrix A so that its maximum eigenvalue is no greater than 0.999. This model generates realistically shaped power grids up to about 300 buses. Once the power grid has been generated, a load input vector, i.e., $\mathbf{u}_t$, is sampled from a Gaussian process truncated so that values are always non-negative.

The first experiment has 6 separate scenarios that test how the proposed method and the standard method perform in identifying 1, 3, and 5 compromised buses. In each case, the attack begins at time 0.1 seconds. Matrix A_t^p is selected such that $(A+BA_t^p)$ has maximum eigenvalue 1.05, ensuring that the attack destabilizes the system. Figure 2 shows the results. The proposed method clearly has better qualitative performance on this particular problem. Specifically, it identifies the compromised buses faster and more accurately.

To look carefully into how our proposed method can differentiate between benign and malicious loads, next we randomly selected a load i and generator j, then set the ith row and jth column of K^{LG} to -10. Recall that negative values of the K^{LG} matrix correspond to benign loads. The results are shown in Figure 3. Negative feedback does not destabilize the system, yet we are able to detect the feedback. The standard method (not shown) has difficulty with this problem as it takes much longer for the standard method to converge.

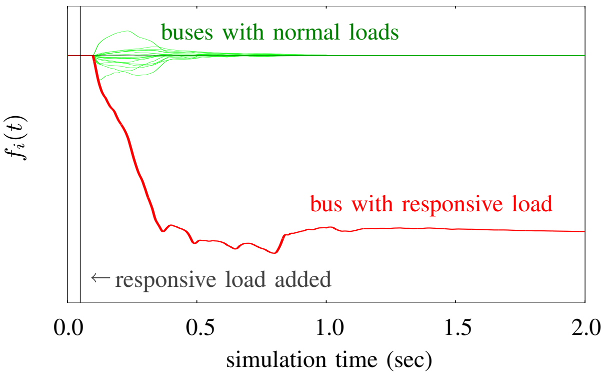

Fig. 3. In this simulation, we added a benign frequency responsive load. Our rank-1 method is able to quickly identify the load with the responsive feedback. The corresponding value of $f_i(t)$ is negative because the feedback is negative. Previous work cannot distinguish these benign frequency responsive loads from malicious loads, whereas ours can.

B. Quantitative Results

We now explore the quantitative performance of our methods by measuring its performance on several power systems. We generated two sets of power grids, one with 20 generators and 20 loads (as in the previous section), and the other one with 100 generators and 100 loads. The standard method was run only on the smaller grid, again because it is computationally infeasible to run it on the larger one, and

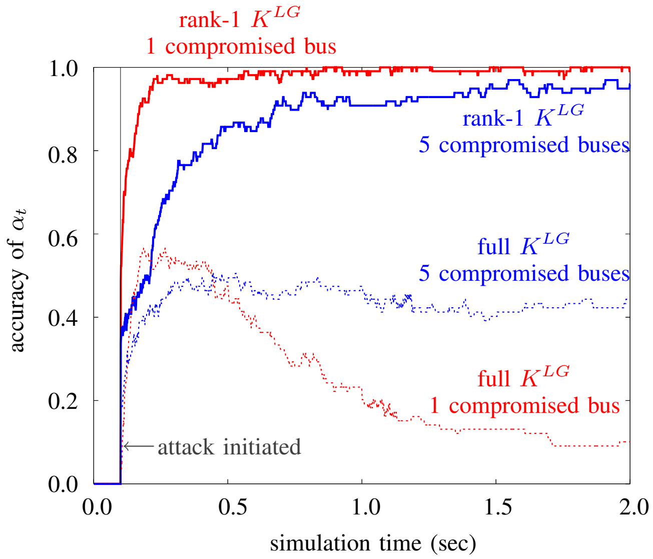

Fig. 4. It takes only about a quarter of a second for our rank-1 method to detect a single attack with 99% accuracy. As the number of attacks increases, our rank-1 method takes longer to achieve high accuracy. With 5 compromised buses, the attack is detected with 95% accuracy by two seconds. This is fast enough to implement corrective actions. The full K^{LG} method has much worse accuracy no matter how many buses are compromised.

the proposed method was run only on both grids.

A major strength of both methods is that they experienced no false positives. We define a false positive to be the detection of an attack when no attack occurred. It does not matter if the value of α_t is correct. When no attack is underway, the largest entries of the estimated K_t^{LG} are typically less than 10^{-6}. When an attack is underway, the largest values of the estimated K^{LG} skyrocket to well above 10^{-1}. Therefore, it is easy to set the threshold τ to avoid false positives.

Finally, we evaluate the method's *accuracy of identification*. We define the accuracy at time point t to be the fraction of α_t values that correctly predict the attacked bus. Figure 4 shows that the longer we wait to declare an attack occurs (i.e. the larger we set τ), the higher our accuracy is. In the case of the rank-1 method detecting a single attack, we observed 99% accuracy in under one second. The rank-1 method operating on the 200 bus system has much higher accuracy than the standard method operating on the significantly easier 40 bus system. The standard method's accuracy is little better than random guessing after two seconds.

VI. Conclusions

In this paper we addressed the open problem of detecting a destabilizing attack against the power system, i.e., identifying which buses are compromised through a possible positive feedback. Our method does not require prior knowledge on the number of buses that are compromised. It also does not require conducting a separate analysis at each bus. Instead, it naturally identifies attacks on the entire system considered as a whole. Therefore, it has low computational complexity. Furthermore, it is capable of distinguishing destabilizing attacks, i.e., load or generation control loops that are malicious and based on positive feedback, from the many load and generation control loops that exist in a power system that are benign and based on negative feedback. Numerical results show that this method successfully identifies complex attacks involving many buses. The detection is accurate and fast.

References

[1] Sajjad Amini, Hamed Mohsenian-Rad, and Fabio Pasqualetti. Dynamic load altering attacks in smart grid. In *Innovative Smart Grid Technologies Conference (ISGT), 2015 IEEE Power & Energy Society*, pages 1–5. IEEE, 2015.

[2] Sajjad Amini, Fabio Pasqualetti, and Hamed Mohsenian-Rad. Detecting dynamic load altering attacks: A data-driven time-frequency analysis. In *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 503–508. IEEE, 2015.

[3] Christopher L. DeMarco, J.V. Sariashkar, and Fernando Alvarado. The potential for malicious control in a competitive power systems environment. In *Control Applications, 1996., Proceedings of the 1996 IEEE International Conference on*, pages 462–467. IEEE, 1996.

[4] Esmaeil Ghahremani and Innocent Kamwa. Online state estimation of a synchronous generator using unscented Kalman filter from phasor measurements units. *IEEE Transactions on Energy Conversion*, 26(4):1099–1108, 2011.

[5] Esmaeil Ghahremani and Innocent Kamwa. Local and wide-area pmu-based decentralized dynamic state estimation in multi-machine power systems. *IEEE Transactions on Power Systems*, 31(1):547–562, 2016.

[6] Paul Hines, Seth Blumsack, E Cotilla Sanchez, and Clayton Barrows. The topological and electrical structure of power grids. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*, pages 1–10. IEEE, 2010.

[7] Angelos K. Marnerides, Paul Smith, Alberto Schaeffer-Filho, and Andreas Mauthe. Power consumption profiling using energy time-frequency distributions in smart grids. *IEEE Communications Letters*, 19(1):46–49, 2015.

[8] Amir-Hamed Mohsenian-Rad and Alberto Leon-Garcia. Distributed internet-based load altering attacks against smart power grids. *IEEE Transactions on Smart Grid*, 2(4):667–674, 2011.

[9] Angel Molina-Garcia, Franois Bouffard, and Daniel S. Kirschen. Decentralized demand-side contribution to primary frequency control. *IEEE Trans. on Power Systems*, 26(1):411–419, February 2011.

[10] Fabio Pasqualetti, Florian Dörfler, and Francesco Bullo. Cyber-physical security via geometric control: Distributed monitoring and malicious attacks. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 3418–3425. IEEE, 2012.

[11] Xiangyun Qing, Hamid Reza Karimi, Yugang Niu, and Xingyu Wang. Decentralized unscented Kalman filter based on a consensus algorithm for multi-area dynamic state estimation in power systems. *International Journal of Electrical Power & Energy Systems*, 65:26–33, 2015.

[12] Joe Short, David Infield, and Leon L. Freris. Stabilization of grid frequency through dynamic demand control. *IEEE Trans. on Power Systems*, 22(3):1284–1293, August 2007.

[13] Eric A Wan and Rudolph Van Der Merwe. The unscented Kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158. Ieee, 2000.

[14] Shaobu Wang, Wenzhong Gao, and A.P. Sakis Meliopoulos. An alternative method for power system dynamic state estimation based on unscented transform. *IEEE Transactions on Power Systems*, 27(2):942–950, 2012.

[15] Zhifang Wang, Anna Scaglione, and Robert J Thomas. Generating statistically correct random topologies for testing smart grid communication and control networks. *IEEE transactions on Smart Grid*, 1(1):28–39, 2010.

[16] Changhong Zhao, Ufuk Topcu, and Steven H. Low. Optimal load control via frequency measurement and neighborhood area communication. *IEEE Trans. on Power Systems*, 28(4):3576–3587, November 2013.