2018 IEEE Conference on Decision and Control (CDC)
Miami Beach, FL, USA, Dec. 17-19, 2018

Secure Control of Wide-Area Power Systems: Confidentiality and Integrity Threats

S. M. Dibaji, M. Pirani, A. M. Annaswamy, K. H. Johansson, and A. Chakrabortty

Abstract— **A cyber-physical model for wide-area control of power systems is considered, where the state variables of each generator are measured and sent to the cyber-network and the corresponding control inputs are computed distributively. The secure control of such wide-area power systems is considered in the presence of cyber attacks that introduce threats that compromise their integrity and confidentiality. Detection, prevention, and resilience for these attacks and algorithms for accomplishing these goals are proposed. In particular, an algorithm to overcome confidentiality attacks of the underlying control gains is presented. Also proposed is an algorithm for defense against integrity attacks that might take place on the cyber-network. For this purpose, a resilient information retrieval approach is leveraged which recovers the true state variables despite the malicious attacks on both virtual machines and communication links. The retrieved states are then used to detect possible attacks on phasor measurement units (PMU) in the next time-step. Simulation studies are included to validate our proposed approaches.**

I. Introduction

Secure control of cyber-physical systems has been a focus area of systems and control community in the last decade. Specifically, after the announcement of catastrophic events, e.g., power system attack in Ukraine [1], resilient control, measurement, and monitoring of power systems became a key concern [2]. Security goals of each system can be categorized into three parts, which include *Confidentiality*, *Integrity* and *Availability* [3]. Most cyber-attacks typically target one of these security goals. A secure control is one that defends against these cyber-attacks and typically includes one or more defense strategies such as *Prevention*, *Resilience*, and *Detection* [4].

Wide-area monitoring and control of power systems using real-time measurements from Phasor Measurement Units (PMUs) is a topic that has gained much importance lately [5]. In this topic, each generator sends its measurements every 33 msec to its rented cloud, denoted as Virtual Machine (VM)s, and each VM, through communication with other VMs in a cyber layer, computes its own control input. The control inputs are sent back to the generators in order to damp the wide-area oscillations. As power systems are large-scale networks, this architecture introduces a lot of computation and security concerns. Reference [6], for instance, discussed secure state estimation approaches under measurement attacks. However, the computational costs of this method are high for large-scale networks. Attack-resilient algorithms in cyber-physical systems have been discussed in [7]. A replay attack and its detection have been discussed in [8]. An event-triggered control design was suggested in [9] for availability threats. Denial of Service (DoS) attacks, which can be viewed as a special case of integrity attacks by zeroing out the unavailable data, has been addressed in [10].

The major issue with the above-mentioned literature is that they do not introduce a network-theoretic condition for the resilience of the cyber-physical system, as it is vital for designing safe interconnections within the cyber layer [11]. They also process the detection phases coincidentally with control, which means that the attacker may lead the system to a vulnerable region by the time they get detected. In this paper, we address such network-theoretic conditions, specifically in a wide-area power system, when subjected to confidentiality and integrity threats. In summary, one confidentiality attack and three integrity attacks (cyber intra-layer links, node attacks, and inter-layer attacks between the cyber and physical layers) are addressed and secure control solutions for defending against these attacks are proposed:

- First, we introduce an algorithm for selecting the control gains that protects against confidentiality attacks. With such an algorithm, the closed loop performance of the power system is preserved, while each generator keeps its control action private.
- Then, we apply a distributed retrieval technique for the power system which is resilient against integrity attacks on both communication links and VMs.
- Finally, we propose how the retrieved states can be used to detect possible faulty measurements from PMUs in the next time-step.

In Section II, preliminary notions in graph theory and linear algebra are introduced. In Section III, a cyber-physical model for the problem of wide-area control is described. Section IV introduces a confidentiality threat and a randomized approach to guarantee the privacy of the control gains. Section V is devoted to the integrity attacks on the cyber layer amongst the VMs. The information retrieval algorithm is given to avoid such threats. In Section VI, a case study is studied. Finally, Section VII summarizes the paper.

II. Graph Theory and Linear Algebra Notions

1) Graph Theory: A simple graph $\mathcal{H}$ is a pair of $(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ denotes the set of nodes, i.e. $\{v_1, v_2, \ldots, v_n\}$, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. The adjacency matrix is also defined as $H = \{h_{ij}\} \in \mathbb{R}^{n \times n}$, where $h_{ij} = 1$ if $(i, j) \in \mathcal{E}$.

S. M. Dibaji and A. Annaswamy are with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA, E-mails: (dibaji/aanna)@mit.edu, M. Pirani and K. H. Johansson are with the Department of Automatic Control, KTH Royal Institute of Technology, Sweden, E-mails: mpirani@uwaterloo.ca and kallej@kth.se, A. Chakrabortty is with the Department of Electrical Engineering, North Carolina State University, Raleigh, NC, USA, E-mail: achakra2@ncsu.edu.

978-1-5386-1395-5/18/$31.00 ©2018 IEEE

Our graphs are bidirectional, i.e. $h_{ij} = h_{ji}$. The terms edge and link are used, interchangeably. If all entries of H is 1, $\mathcal{H}$ is called a complete graph. The number of links connected to node i is called degree of i and denoted by d_i. Likewise, σ_i is the minimum degree of all nodes of the graph, i.e. $\sigma_i = \min(d_i), i = 1, \ldots, n$. A vertex cut is a set $\mathcal{S} \subset \mathcal{V}$ such that removing the vertices of $\mathcal{S}$ (and the associated edges) causes the graph to be disconnected. The connectivity of a graph is the smallest size of a vertex cut. A path, with length of q, is a sequence of nodes like $v_{i_1}, v_{i_2}, \ \ldots, \ v_{i_q}$, where $(v_{i_m}, v_{i_{m+1}}) \in \mathcal{E}$. The diameter of a graph is the length of the longest path in the graph [12].

2) Linear Algebra: We denote the $(p \times q)$-dimensional matrices on the set $\mathbb{S}$ by $M = \{m_{ij}\} \in \mathbb{S}^{p \times q}$. $\mathbb{I}_p$ is used for the identity $p \times p$ matrix. $\rho(M)$ shows the column rank of M. Left pseudo-inverse of M is denoted by $M_\ell^\dagger$, where $M_\ell^\dagger M = \mathbb{I}_q$. Likewise, right pseudo-inverse $M_r^\dagger$ is defined. Left null space of M is the space built by all vectors y such that $My^T = \mathbf{0}_{p\times 1}$ and is denoted by $\mathbf{N}_\ell(M)$.

III. A Cyber-Physical Model for Wide-Area Control

A. Power System Model

We consider a power system with a total of n generators. Assuming that each generator G_i has n_i state variables, which include rotor phase angle and frequency, excitation voltage, d-axis sub-transient flux, exciter states, power system stabilizer states, turbine/governor states, active-and reactive load modulation states, and states of Static Var Compensators, and FACTS devices, and has a scalar input which corresponds to the field excitation voltage. Aggregating all generators' state variables, the network model, as shown in [13], using Kron reduction, can be compactly written as

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= A_c\, \boldsymbol{x}(t) + B_c\, \boldsymbol{u}(t) \\ \boldsymbol{y}(t) &:= C\, \boldsymbol{x}(t), \end{aligned} \tag{1}$$

where $B_c = [B_1^T, \cdots, \cdots, B_n^T]^T \in \mathbb{R}^{N\times n}$. In this model $\boldsymbol{x}(t) = [x_1(t), \ldots, x_n(t)]^T \in \mathbb{R}^N$ which corresponds to state variables of i^{th} generator, and $\boldsymbol{u}(t) \in \mathbb{R}^n$ which corresponds to the deviation of u_i, $i = 1, \ldots, n$ around their equilibrium values, and $N = \sum_{i=1}^n n_i$. Let the power system network be divided into m non-overlapping areas. We assume that $\boldsymbol{y}$ is a vector of the small-signal generator frequency $\omega(t)$ which often is measurable, and is an effective indicator of the underlying damping in the dynamic system. The input $\boldsymbol{u}(t)$ is commonly used for designing feedback controllers such as Power System Stabilizers (PSS), which takes local feedback from the generator speed, and passes it through a lead-lag controller for producing damping effects on the oscillations in phase angle and frequency. PSSs, however, are most effective in adding damping to the fast oscillation modes in the system, and perform poorly in adding damping to the slow or inter-area oscillation modes [5]. It is assumed that $\boldsymbol{x}(t)$ in the linearized model (9) is available, by using a state-estimator as in [14]. Our goal is to design a supplementary resilient controller $\boldsymbol{u}(t)$ for the model (1) in addition to the local PSS by using a state-feedback from selected sets of other generators.

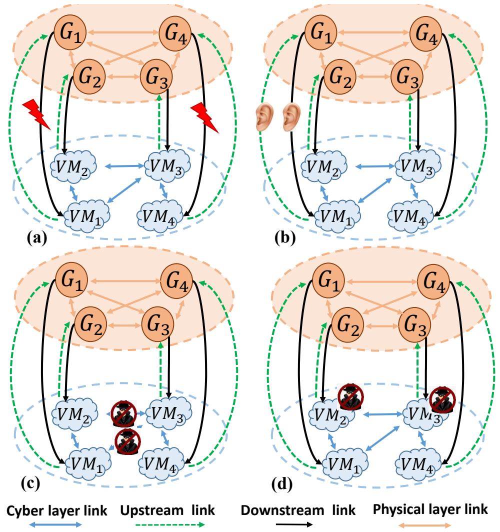

Fig. 1. Wide-Area control via Internet of clouds. (a) Attacks on downstream links,(b) Confidentiality attacks on inter-layer links to obtain the control gains , (c) Cyber-attacks on cyber-links, (d) Cyber-attacks on VMs.

The wide-area control architecture is realized using a cloud-in-the-loop cyber-physical architecture that was recently proposed in [15]. This architecture is briefly shown in Fig. 1. In this architecture every generator is assumed to own a local cloud network, where virtual computer, referred to as Virtual Machines (VMs) can be created on-demand for computing control signals. Every generator sends its estimated state vector at each T_p msec to its designated VM. Upon receiving these estimates, VMs communicate with each other, through a communication graph $\mathcal{H}$ and share their respective state vectors. After that, each VM computes the control input for its respective generator by using the block row of the state-feedback gain matrix K which is pre-embedded in that VM. One problem with the gain matrix K being dense is that every VM must communicate with every other VM, or equivalently $\mathcal{H}$ is a complete graph, which may lead to unnecessary data flooding and immediate cyber-attack propagation despite heavy communication cost. One of our goals in this paper is to circumvent this problem by finding a tractable way for introducing algorithms and sparser $\mathcal{H}$ topologies such that the closed-loop system remains resilient against different cyber attacks.

B. A Delay-Aware Control Architecture

The control design that we work on is based on a sampled-data representation of (1). The overall goal of the control input is damping the power oscillations using a delay-aware and implementable design and its details are described in [15]. Having effect of delays, it turns into the following sampled-data equation

$$\boldsymbol{x}[k+1] = A\boldsymbol{x}[k] + B_1\boldsymbol{U}[k] + B_2\boldsymbol{U}[k-1], \tag{2}$$

where $A = e^{A_c T_p}$ and B_1 and B_2.The state space form of (2) can be re-written in the following extended form

$$\boldsymbol{W}_h[k+1] = A_h\boldsymbol{W}_h[k] + B_h\boldsymbol{U}[k], \tag{3}$$

where $\boldsymbol{W}_h[k] = \begin{bmatrix} \boldsymbol{x}[k] \\ \boldsymbol{U}[k-1] \end{bmatrix}$, $A_h = \begin{bmatrix} A & B_2 \\ O_{N\times N} & O_{n\times n} \end{bmatrix}$ and $B_h = \begin{bmatrix} B_1 \\ I_n \end{bmatrix}$. An LQR state-feedback control is employed through minimizing the cost function

$$\mathcal{J} = \sum_{0}^{\infty} \boldsymbol{W}_h[k]^T Q \boldsymbol{W}_h[k] + \boldsymbol{U}[k]^T R \boldsymbol{U}[k], \tag{4}$$

where the weights are chosen as described in [16]. The optimal state-feedback gains K and G will be

$$\boldsymbol{U}[k] = K_h \boldsymbol{W}_h[k], \tag{5}$$

where $K_h \triangleq \begin{bmatrix} K & G \end{bmatrix} \in \mathbb{R}^{n\times(n+N)}$. Note that matrices K and G are computed once (offline) in an Independent System Operator (ISO) and saved in the memory of each VM. Then the closed-loop system becomes

$$\boldsymbol{W}_h[k+1] = (A_h + B_h K_h)\boldsymbol{W}_h[k] = \mathcal{A}_{cl}\boldsymbol{W}_h[k]. \tag{6}$$

IV. Secure Control against Confidentiality Attacks

In this section, we present a prevention technique based on randomization which each VM keeps its control gain private and protected from confidentiality attacks. This is accomplished by assigning control gains to VMs (in random) from a large number of possible control gains all of which ensure the desired closed-loop performance. Furthermore, by taking advantage of a wide possibility of choosing control gains, we adopt a strategy for choosing the gains which further strengthens vulnerable sub-spaces of the overall closed-loop controllability space.

A. Confidentiality Attacks on Inter-Layer Links

One of the confidentiality goals in wide-area control design is that the gains of each VM must be kept private, i.e., each VM should only have access to its own gain value. This is because the control gains determine how much each entity contributes in damping. The eavesdroppers can attack the upstream/downstream links and obtain the information of $\boldsymbol{U}[k]$ and $x[k]$ for $(N+n)$ time-steps. Since

$$U_i[k] = K_i x[k] + G_i U[k-1], \tag{7}$$

the $(N+n)$ entries of K_i and G_i can be obtained by the eavesdropper in $(N+n)$ time-steps, assuming that it has access to $\boldsymbol{x}[k]$ and $\boldsymbol{U}[k-1]$ through downstream and upstream links, respectively, as shown in Fig. 1 (b).

In the following subsection, we present a procedure for ensuring that the control gains are not accessed by the attacker and remain private.

B. Randomization against Eavesdroppers

Randomization has proved to be a useful approach in cyber-security problems, e.g. [17]. Our approach here is that instead of just using K_i and G_i, each VM picks the gains from a set of control gains, randomly, such that the eigenvalues and eigenvectors of the closed-loop form $\mathcal{A}_{cl}$ are preserved. The following proposition ensures that the size of such a set of gains is large enough.

Proposition 1: The set of control gains K_i and G_i in (7), which preserve the poles of the closed-loop system fix, is not unique.

Proof: We denote $V = [v_1^T, \cdots, v_n^T]^T$, where v_i^T is the i^{th} eigenvector of the closed-loop system $\mathcal{A}_{cl}$ in (6). We must show that

$$FV = 0, \tag{8}$$

has non-zero solutions, where $F \in \mathbb{R}^{(N+n)\times(N+n)}$. Here we should note that $\mathbf{N}_\ell(V)$ is not an empty set. This is due to the fact that the closed-loop matrix $\mathcal{A}_{cl}$ does not have linearly independent eigenvectors (verified via physical values of dynamic matrices and the control gain). Hence, (8) has non-trivial solutions. Based on this, i^{th} VM chooses its control gain from the following set (similarly for G_i)

$$\mathcal{K}_i = \{\overline{K}_i = K_i + F_a \ | B_h F_a = \ F_{is} \text{ and } F_{is} \text{ solves (8)}\}. \tag{9}$$

With this, we have $(A_h + B_h(K_h + F_a))V = \Lambda V$, where Λ is the diagonal matrix of the eigenvalues of $A_h + B_h K_h$. This is because

$$(A_h + B_h(K_h + F_a))V = (A_h + B_h(K_h))V + B_h F_a V = \\ (A_h + B_h K_h)V + F_{is}V = (A_h + B_h(K_h))V = \Lambda V$$

and this means that control gain $\overline{K}_i$ preserves the closed-loop poles and eigenvectors of the system. ■

We should note that the above confidentiality procedure is conducted offline. In particular, before the overall power-plant starts working, the ISO saves γ distinct solutions of (8) and produces different control gain based on those. Those distinct solutions need not be linearly independent but have to be different from each other (they can be even multiplications of each other). The above procedure provides sub-optimal gains for (4) which is the cost paid for increasing the confidentiality of the system.

V. Secure Control against Integrity Attacks

In this section, we propose several detection/resilience mechanisms against the attacks on the integrity of data. The integrity attacks can be on the cyber layer or on the inter-layer communications.

A. Integrity Attacks on Cyber Layer

The cyber layer attacks are categorized into those on links and VMs. The main defense idea for both, however, is that information of each VM, either faulty or non-faulty, is propagated in the network. Each VM receives data of all VMs from sufficient number of paths in the graph so that via comparison, it can recognize which paths include a faulty link and which paths do not. In this subsection, we present a distributed algorithm to implement this idea and also retrieve the true state variables, even if they were not originally faulty. These states are then used for state-feedback control as in (5). Although such a control needs all state variables that usually are provided through a complete graph $\mathcal{H}$, the information retrievals are carried out over a sparser graph, which is preferred due to lower communication cost and less vulnerable points.

Here we describe new update rules among virtual machines to be used later for retrieval procedures. We assume that the state $(x_i[k], U_i[k-1])$ is taken as the initial state $\chi_i^k[0]$ in its corresponding VM in the cyber layer. The final purpose of each VM is to retrieve, at each time-step kT_p the $\chi^k[0] = [x[k]^T, U[k-1]^T]^T$ to be used in state-feedback, in LT_c time, where $LT_c < T_p$ (see Fig. 4 (b)). We use the matrix $\Phi = [\phi_{ij}]$ with the same zero and non-zero structural pattern as $\mathcal{H}$. For example, we simply can set $H = \Phi$, for the retrieval process. Then each VM in the cyber layer updates its states, every T_c msec, based on the following linear iterative rule

$$\chi_i^k[(\ell+1)T_c] = \left(\phi_{ii}\chi_i^k[\ell T_c] + \sum_{j\in\mathcal{N}_i} \phi_{ij}\chi_j^k[\ell T_c]\right), \quad (10)$$

where $T_c \ll T_p$. At each time-step ℓ, virtual machine v_i has access to its own value, as well as the values of its neighbors, i.e.,

$$\psi_i^k[\ell T_c] = \Gamma_i \chi^k[\ell T_c], \quad (11)$$

where Γ_i is a $(d_i+1) \times n$ matrix with a single 1 in each row that denotes the positions of the state-vector $\chi^k[\ell T_c]$ available to i^{th} virtual machine v_i (i.e., these positions correspond to virtual machines that are neighbors of i^{th} VM, along with itself). Hereafter, we drop the sampling time T_c from the equations, unless it is necessary to be indicated.

Remark 1: Note that since each communication link conveys *all* state variables of each generator j along with the control inputs $U_j[k-1]$ at each time-step k, the matrices Φ and Γ_i, using Kronecker product, should be replaced, for example, with $\Phi \otimes \mathbb{I}_{n_i+1}$ and $\Gamma_i \otimes \mathbb{I}_{n_i+1}$, respectively, if n_i is identical for all generators. □

1) Integrity Attacks on Cyber Intra-Layer Links: We first discuss the case where the communication between two VMs is impaired by an external attacker, as shown in Fig. 1, and those VMs try to retrieve information despite the existence of such malicious action. Now suppose that link (p, i) in the network modifies the value that virtual machine v_i receives from v_p at time-step ℓ to be $\chi_i^k[\ell] + s_{pi}[\ell]$, where $s_{pi}[\ell]$ is an (arbitrary) additive error. Hence, each virtual machine v_i performs a linear iterative policy using the following time invariant updating rule

$$\chi_i^k[\ell+1] = \left(\phi_{ii}\chi_i^k[\ell] + \sum_{j\in\mathcal{N}_i} \phi_{ij}\chi_j^k[\ell]\right) + \phi_{ip}s_{pi}[\ell]. \quad (12)$$

When there are multiple unreliable links entering a node, there is the possibility that the errors that they introduce cancel each other out in expression (10). Hence, we have the following definition.

Definition 1: A set of unreliable links $\mathcal{L}$ into the same node i is malicious if $\sum_{p\in\mathcal{L}} \phi_{ip}s_{pi}[\ell]$ is nonzero for at least one time-step ℓ. □

Furthermore, let $\mathcal{S} \subset \mathcal{E}$ be any set of links in the network. Define the following set $\mathcal{L}_i(\mathcal{S}) \subset \mathcal{E}$

$$\mathcal{L}_i(\mathcal{S}) = \{(u, v) \in \mathcal{S} | v = v_i\}. \quad (13)$$

In words, the set $\mathcal{L}_i(\mathcal{S})$ contains all links in $\mathcal{S}$ that end at node v_i. Since the information coming from the neighbors of virtual machine v_i may be corrupted by the set of unreliable links $\mathcal{F}$, the visible states (11) will turn to

$$\psi_i^k[\ell] = \Gamma_i\chi^k[\ell] + E_{\mathcal{L}_i(\mathcal{F})}s_i[\ell], \quad (14)$$

where matrix $E_{\mathcal{L}_i(\mathcal{F})}$ is $(deg(v_i)+1) \times |\mathcal{L}_i|$ binary matrix where a 1 in row j and column m of $E_{\mathcal{L}_i(\mathcal{F})}$ indicates that element j of $\psi_i^k[\ell]$ is affected by element m of $s_i[\ell]$, i.e., the m^{th} unreliable link in $\mathcal{L}_i(\mathcal{F})$ affects the value received by v_i from its j^{th} neighbor. The following theorem, from [18], discusses the ability of a network with sufficiently large connectivity to perform the information retrieval algorithm despite of the existence of unreliable links. The details of the resilient distributed information retrieval will be presented later.

*Theorem 1 (**Faulty Communications Between VMs**):* Let the graph of the given network $\mathcal{H}$ have connectivity κ. If $\kappa \geq 2f+1$, then there exists a positive integer $L \leq N$ such that, for almost any choice of weights, every node in the network can correctly determine $\chi^k[0]$ after running the linear iteration for at most L time-steps, despite the existence of any f unreliable links in the network. □

2) Integrity Attacks on Virtual Machines: In this subsection, instead of unreliable links, we consider the case where the virtual machines are hijacked. In particular, there exists at least one virtual machine, v_i, which does not obey the predefined updating rule, as shown in Fig. 1. More formally, it performs a linear iterative policy using the following time invariant updating rule

$$\chi_i^k[(\ell+1)] = \phi_{ii}\chi_i^k[\ell] + \sum_{j\in\mathcal{N}_i} \phi_{ij}\chi_j^k[\ell] + a_i[\ell], \quad (15)$$

where $\phi_{ij} > 0$ are some predefined weights and $a_i[\ell]$ is the additive error (fault) injected by v_i. Dynamics (15) can be written in the vector form as

$$\chi^k[\ell+1] = \Phi\chi^k[\ell] + \underbrace{[\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_f]}_{\mathcal{B}}\boldsymbol{a}[\ell], \quad (16)$$

where $\Phi_{n\times n}$ is a matrix which captures the communication between virtual machines in the network and $\boldsymbol{a}[\ell] = [a_1[\ell], a_2[\ell], \dots, a_f[\ell]]^T$ is the vector of f faulty inputs and $\mathbf{e}_j$ denotes an $n \times 1$ unit vector with a single nonzero entry with value 1 at its j^{th} position. Similar to the case of faulty links, the visible states for v_i are defined as (11). The following theorem from [19] provides a sufficient condition of state retrieval by each virtual machine, in the presence of some malicious VMs in the network[1].

*Theorem 2 (**Hijacked Virtual Machines**):* Let the graph of the given network $\mathcal{H}$ have connectivity κ. If $\kappa \geq 2f+1$, then there exists a positive integer $L \leq N$ such that, for almost any choice of weights, every node in the network can correctly determine $\chi^k[0]$ after running the linear iteration for

[1] In [19], this technique is utilized for average consensus in multi-agent systems in the presence of malicious agents. We take advantage of it for online recovery of the state variables, in a faster period than the sensor measurements, so that the performance of the physical plant is not affected while the attacks are getting detected.

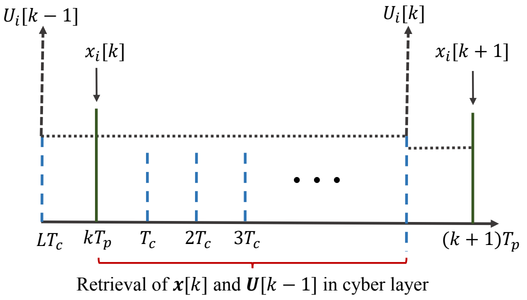

Fig. 2. Resilient architecture of cyber layer via information retrieval.

at most L time-steps, despite the existence of any f hijacked virtual machine in the network. □

3) Retrieval Procedure: In this subsection, we present an algorithm to retrieve the state variables despite attacks on VMs. The cyber-link-attack detection can be written in a similar way and is omitted due to space limitations. In order to retrieve all true state variables, *observability* matrices $\mathcal{O}_{i,L}$ are defined as

$$\mathcal{O}_{i,L} := \begin{bmatrix} \Gamma_i \\ \mathcal{O}_{i,L-1}\Phi \end{bmatrix}, \tag{17}$$

where $\mathcal{O}_{i,0} = \Gamma_i$. Moreover, since the attack set at the beginning of the process of retrieval is not known, we combinatorially start searching in potential attacked set of nodes $\mathcal{I}_j$, i.e $|\mathcal{I}_j| = f$, where j denotes for the j^{th} potential subset of $\mathcal{V}$ with f members. Algorithm 1 reveals correctness of our geuss on $\mathcal{I}_j$. Based on $\mathcal{I}_j$, we require also to define the *attack* matrices

$$\mathcal{M}_{i,L}^{\mathcal{I}_j} = \begin{bmatrix} O_{d_i\times f} & O_{d_i\times(L-1)f} \\ \mathcal{O}_{i,L-1}\mathcal{B}_{\mathcal{I}_j} & \mathcal{M}_{i,L-1}^{\mathcal{I}_j} \end{bmatrix}, \tag{18}$$

where $\mathcal{M}_{i,0}^{\mathcal{I}_j} = O_{d_i\times 1}$. Note that $\mathcal{O}_{i,L} \in \mathbb{R}^{d_i(L+1)\times n}$ and $\mathcal{M}_{i,L}^{\mathcal{I}_j} \in \mathbb{R}^{(L+1)d_i\times Lf}$.

Now, we re-write the cyber visible states of node i in a concatenated form as below

$$\psi_i^k[0:L] = \mathcal{O}_{i,L}\chi^k[0] + \mathcal{M}_{i,L}^{\mathcal{I}_j}a[0:L-1]. \tag{19}$$

Note that each VM need not compute all $\mathcal{M}_{i,L}^{\mathcal{I}_j}$ and $\mathcal{O}_{i,L}$ at each time-step kT_p and the ISO is responsible of saving safely these matrices and control gains of i^{th} generator on the processor of each VM. Since, this process is offline and before putting the cloud in the loop, the computational burden of each VM is, at most, searching among $\binom{n}{f}$ column spaces of $[\psi_i^k[0:L]\ \mathcal{O}_{i,L}\ \mathcal{M}_{i,L}^{\mathcal{I}_j}]$. The attack detection in Algorithm 1 is carried out in the while loop and the last $\mathcal{I}_j$ includes all attacked nodes. It is also noteworthy that the attacks can be different at each time interval T_p, but during each retrieval process, at most f attacked nodes/links, can be tolerated and detected.

B. Integrity Attacks on PMU Data

It is also noteworthy that another strategy for using a non-complete graph for state-feedback control is discussed in [15] where a participation factor (PF) based approach is proposed to design sparse control gains. Although all retrieval processes can be done in sparse control gain designs, the VMs require to only retrieve the states of their neighbors or the VMs in their vicinity. We did not follow this research line as (i) they give us sub-optimal control gains with computational complexities and (ii) by retrieval of all state variables, i^{th} VM has access to full state $\boldsymbol{x}[k-1]$; therefore, it can compute the residual

Algorithm 1 Distributed Retrieval Procedure

```
1: $j \leftarrow 1$.
2: while $\rho([\psi_i^k[0:L]\ \mathcal{O}_{i,L}\ \mathcal{M}_{i,L}^{\mathcal{I}_j}]) \neq \rho([\mathcal{O}_{i,L}\ \mathcal{M}_{i,L}^{\mathcal{I}_j}])$ do
3:    $j \leftarrow j+1$.
4:    Repeat 1.
5: end while
6: Find an $N_{i,L}^j \in \mathbf{N}_\ell(\mathcal{M}_{i,L}^{\mathcal{I}_j})$.
7: $P_{i,L}^j = (N_{i,L}^j\mathcal{O}_{i,L})_\ell^\dagger N_{i,L}^j$.
8: $\chi^k[0] = P_{i,L}^j\psi_i^k[0:L]$.
```

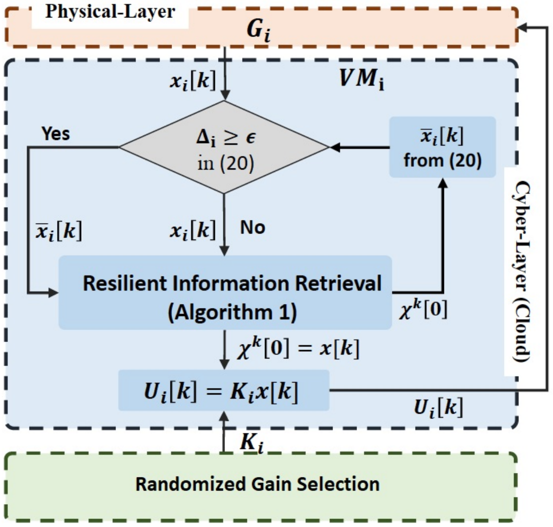

Fig. 3. Schematic diagram of the overall algorithm.

$$\Delta_i = x_i[k] - \underbrace{A_i\boldsymbol{x}[k-1] + B_{1i}\boldsymbol{U}[k-1] + B_{2i}\boldsymbol{U}[k-2]}_{\triangleq \bar{x}_i[k]}, \tag{20}$$

where A_i, B_{1i}, and B_{2i} are the i^{th} row of the matrices A, B_1, and B_2, respectively. If this residual exceeds a threshold, it means that the information coming from the downstream link (PMU) is corrupted. The comparison of $x_i[k]$ and its estimate in Fig. 3 cannot be carried out in [15] to catch the attacks on down-stream links.

The overall procedure in the paper is schematically depicted in Fig. 3. The randomized gain selection, which were discussed by detail in Section IV, is shown in the green box in Fig. 3. This procedure is performed offline and confidential control gains are prepared before the overall algorithm starts operating. The comparison, to make sure that the state $x_i[k]$ which is received by i^{th} VM is not affected by an intruder, is shown in a grey diamond. After that, either of the states $x_i[k]$ or $\bar{x}_i[k]$ (if PMU is faulty) is considered as the true state of i^{th} VM. Then each VM performs the resilient retrieval

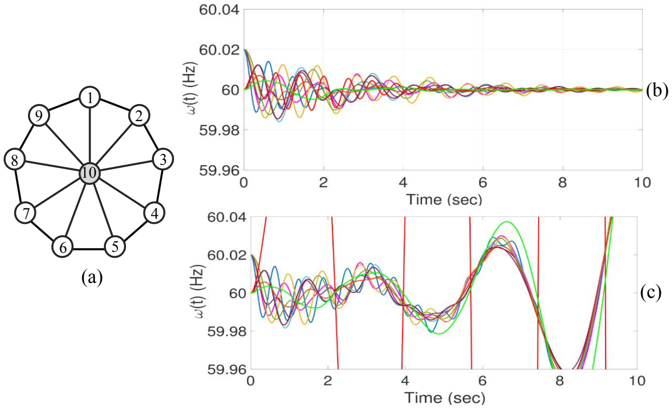

Fig. 4. (a) Communication graphs between VMs (wheel graph) (b) Frequency of all generators after applying the resilient retrieval procedure, and (c) Frequency of all generators when VM 10 in Fig. (a) is attacked such that all generators become unstable.

algorithm to obtain the states of all other VMs, $\chi^k[0]$, which is the state of all VMs at time step k and is used to produce the feedback control.

VI. Simulation Study

We conducted our simulations on an IEEE 39-Bus New England system with 10 generators ($n = 10$), where each generator has $r = 13$ state variables. The PMU measurements are sent to the VMs every 33 msec, i.e. $T_p = 33$ msec. A disturbance, initially, is injected to the system as a variation in initial values ($\omega_i(0) = 60.02$ Hz for $i = 1, \ldots, 5$). We carry out simulations on two cases: (i) when the cyber-layer, with the wheel topology shown in Fig. 4 (a) is attacked on one of its VMs and a resilient retrieval algorithm is performed, and (ii) when the cyber-layer (with the same size but complete graph) is under attack and the resilient retrieval procedure is not performed. The 10^{th} VM is hijacked such that the additive term $a_{10}[\ell]$ cancels all $\phi_{ij}[\ell]x_j[\ell]$ terms and replaces $(x[k], U[k-1])$ with $(-2K_r^\dagger(B_1)_r^\dagger Ax[k], -G_r^\dagger(B_1)_r^\dagger B_2U[k-1])$. Without retrieval, this ends up in the closed-loop equation $x_{10}[k+1] = -A_{10}x[k]$. We imposed the attack for all $t > 0$.

*1) **Wheel Graph with Retrieval Procedure**:* Here we consider the wheel topology, well-known as the Tutte's synthesis of a 3-connected network [12], and the center node is considered as an attacker. For the retrieval, we set $L = n - \sigma_i = 8$ and $T_c = 1$ msec; Thus, $\tau = 8$ msec. We ran the retrieval procedure on a wheel graph. The algorithm shows an acceptable performance as it is shown in Fig. 4 (b).

*2) **Complete Graph without Retrieval Procedure**:* Here, with the same initial conditions and control gains as the previous case, the topology is a complete graph (which has more than twice edges than that of the wheel graph); however, the resilient retrieval algorithm is not performed. As it is seen in Figure 4 (c), the attacker by injecting false data on a generator could affect badly and all generators became unstable. This is despite the fact that $f = 1$, A is Schur stable, and the communication graph is complete.

VII. Summary and Conclusions

We investigated a confidentiality attack and three integrity attacks on a cyber-physical model for wide-area control of power systems. To defend against confidentiality attacks, we proposed an algorithm for randomly switching the control gains. Furthermore, against integrity attacks in the cyber layer, we introduced a resilient information retrieval approach for recovering the true state variables despite the malicious attacks on both virtual machines and communication links. Finally, by making use of the retrieved state variables, we suggested a simple detection mechanism against integrity attacks on PMUs. Simulation studies verified our proposed approaches. In future, we will extend our results and examine the suggested techniques on much more realistic case studies.

References

[1] "Analysis of the cyber attack on the Ukrainian power grid: Defense use case," SANS Institute and Electricity Information Sharing and Analysis Center (E-ISAC), Tech. Rep., March 2016.

[2] A. Ashok, M. Govindarasu, and J. Wang, "Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid," *Proc. IEEE*, vol. 105, no. 7, pp. 1389–1407, 2017.

[3] M. Bishop, *Introduction to Computer Security*. Addison-Wesley, 2005.

[4] S. M. Dibaji, M. Pirani, D. Flamholz, A. M. Annaswamy, K. H. Johansson, A. Chakrabortty, and E. Lavretsky, "A Systems and Control Perspective of CPS Security," *Submitted for Journal Publication*.

[5] A. Chakrabortty and P. P. Khargonekar, "Introduction to wide-area control of power systems," in *Proc. American Control Conf.*, 2013, pp. 6758–6770.

[6] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S. S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *Proc. IEEE Conf. on Decision and Control*, 2010, pp. 5991–5998.

[7] F. Pasqualetti, F. Dorfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," in *IEEE Trans. Autom. Control*, vol. 58, no. 11, 2013, pp. 2715–2729.

[8] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *47th Annual Allerton Conf.*, 2009, pp. 91–918.

[9] C. De Persis and P. Tesi, "Input-to-state stabilizing control under Denial-of-Service," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 2930–2944, 2015.

[10] A. Jain, A. Chakrabortty, and E. Biyik, "Structurally constrained ℓ_1-sparse control of power systems: Online design and resiliency analysis," in *Proc. American Control Conf.*, 2018, pp. 4195–4200.

[11] E. Moradi Shahrivar, M. Pirani, and S. Sundaram, "Robustness and algebraic connectivity of random interdependent networks," in *5th IFAC Workshop on Distributed Estimation and Control of Networked Systems*, 2015, pp. 252–257.

[12] J. L. Gross and J. Yellen, *Graph Theory and its Applications*. CRC Press, 2005.

[13] P. W. Sauer and M. A. Pai, *Power System Dynamics and Stability*. Prentice Hall, 1998.

[14] A. K. Singh and B. C. Pal, "Decentralized dynamic state estimation in power systems using unscented transformation," *IEEE Trans. on Power Syst.*, vol. 29, no. 2, pp. 794–804, 2014.

[15] S. M. Dibaji, A. Annaswamy, A. Chakrabortty, and A. Hussain, "Sparse and distributed control of wide-area power systems with large communication delays," in *Proc. American Control Conf.*, 2018, pp. 3822–3827.

[16] S. M. Dibaji, Y. Yildiz, A. Annaswamy, A. Chakrabortty, and D. Soudbakhsh, "Delay-aware control designs of wide-area power networks," in *Proc. World Congress of the International Federation of Automatic Control*, 2017, pp. 79–84.

[17] S. M. Dibaji, H. Ishii, and R. Tempo, "Resilient randomized quantized consensus," *IEEE Trans. Autom. Control*, vol. 63, pp. 2508–2522, 2018.

[18] S. Sundaram, S. Revzen, and G. Pappas, "A control-theoretic approach to disseminating values and overcoming malicious links in wireless networks," *Automatica*, vol. 48, pp. 2894–2901, 2012.

[19] S. Sundaram and C. N. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Trans. Autom. Control*, vol. 56, no. 7, pp. 1495–1508, 2011.