Assessing and Mitigating Impact of Time Delay Attack against Cyber-Physical Systems

Xin Lou* Cuong Tran Huu† Rui Tan‡ David K.Y. Yau*† Zbigniew T. Kalbarczyk§
*Illinois at Singapore ‡Nanyang Technological University, Singapore
†Singapore University of Technology and Design §University of Illinois at Urbana-Champaign, USA

***Abstract*—Recent attacks against cyber-physical systems (CPSes) show that traditional reliance on isolation for security is insufficient. This paper develops real-time assessment and mitigation of an attack's impact as a system's built-in mechanisms. We focus on a general class of attacks, which we call *time delay attack*, that delays the transmissions of control data packets in a linear CPS control. Our attack impact assessment, which is based on a joint stability-safety criterion, consists of (i) a machine learning (ML) based safety classification, and (ii) a tandem stability-safety classification that exploits a basic relationship between stability and safety, namely that an unstable system must be unsafe whereas a stable system may not be safe. The ML addresses a state explosion problem in the safety classification, whereas the tandem structure reduces false negatives in detecting unsafety arising from imperfect ML. We apply our approach to assess the impact of the attack on power grid automatic generation control, and accordingly develop a two-tiered mitigation that tunes the control gain automatically to restore safety where necessary and shed load only if the tuning is insufficient. Extensive simulations based on a 37-bus system model are conducted to evaluate the effectiveness of our assessment and mitigation approaches. We also apply our attack impact assessment approach to a thermal power plant control system consisting of two PID control loops.**

I. Introduction

By integrating modern information and communication technologies (ICTs), critical systems (e.g., power grids and advanced manufacturing facilities) are transforming into cyber-physical systems (CPSes). However, whereas the use of ICT can improve system performance, it also incurs cybersecurity risks. To date, the security of these systems has largely relied on isolation from public networks through air gaps and firewalls. However, the isolation is questionable, due to insiders [1] and stepping stone attacks [2]. For instance, the Dragonfly attack against power grids [3] compromised a third-party virtual private network (VPN) software vendor first, and then used the result as a stepping stone to intrude into the grids. Once the attackers breach the isolation, they can launch powerful data integrity attacks similar to Stuxnet [4]. They can also build a botnet that exploits proliferating industrial Internet-of-Things (IoT) devices to launch distributed denial-of-service (DDoS) attacks. A prominent example is the 2016 Dyn attack launched from a massive Mirai-infected IoT botnet [5].

Motivated by the aforementioned security incidents, this paper studies the assessment and mitigation of the impact of an important and general class of attacks, which we call the *delay attack*, on a CPS that employs closed-loop control [6]–[8]. The attack maliciously delays the transmissions of control data packets without having to temper with the data content. Since CPS control often requires timely execution of control actions, the attack can undermine system performance severely and even cause catastrophic safety incidents. Compared with data tampering that needs to break non-trivial cryptographic protection, the delay attack can be implemented more simply using compromised routers or jamming communication channels through an IoT botnet to increase the communication latency. Hence, it is an important threat that requires immediate attention. However, whereas the attack can be readily detected by trustworthy synchronization of the clocks of coordinating CPS devices and subsequent verification of packet timestamps, assessing and mitigating its impact in real time are challenging due to the complexity of typical real-world CPS control systems.

In this paper, we propose to use a joint stability-safety criterion for the attack impact assessment and mitigation. *Stability* and *safety* concern a system's ability to keep its state fluctuations *bounded* and moreover *within a specified safe range*, respectively, in the presence of exogenous disturbances that are *of bounded magnitudes*. As the disturbances (e.g., sensor noises and system input changes) are inevitable, stability is a basic requirement that must be met by any CPS. Otherwise, the system may experience unacceptable state divergence given a possible disturbance. Besides stability, however, the CPS must further operate within its engineered safety limits. For instance, a 60 Hz power grid must maintain its frequency within a tight range of about 59.5 Hz to 60.5 Hz; otherwise, generators/loads may trip automatically causing blackouts. Thus, real-time knowledge of the system's stability and safety is critical. Based on this knowledge, if a delay attack is assessed to destabilize the system or push it into an unsafe region, attack mitigation must be initiated to regain the system's stability and safety.

This paper considers linear time-invariant (LTI) systems that can characterize a wide range of real-world CPS systems. From control theory, a LTI system's stability depends on the system model only. Simple LTI systems can have closed-form stability conditions. The stability of more complex LTI systems can be assessed numerically by simulations. In contrast, the safety of a system depends on its future transient trajectory, which presents various challenges. Particularly, simulating the transient trajectory of a complex system may be too slow for detecting and reacting to its impending unsafety. An alternative approach is to run offline simulations comprehen-

sively to understand the system's safety proactively, ahead of actual operation. However, as the trajectory depends on the initial system state, enumerating all the possible states in a continuous value domain is generally impossible. On the other hand, discretization of the value domain, even if it is sufficiently accurate, will lead to an enumeration of exponential complexity with respect to the dimension of the system state. For instance, for an n-bus power system, whose system state dimension is n, its total number of discretized states is m^n, where m is the number of quantization steps for the state variable corresponding to each bus. The value of n for practical systems can be in the hundreds, making the enumeration computationally infeasible.

To address the above challenge, we propose a novel delay attack impact assessment that features (i) a machine learning (ML) based safety classification, and (ii) a tandem stability-safety classification structure. First, to avoid the exponential complexity of enumerating the system states, we adopt a Monte Carlo method to randomly sample the state space and run offline transient simulations to generate safety labels for the sampled states. These states and their labels are used to train an ML model that can classify the safety of a current system based on its real-time state, even if this state is most likely not in the training set. The online safety classification based on the trained model will be fast enough to ensure the timeliness of the impact assessment. Second, we leverage a basic relationship between the stability and safety, namely that an unstable system must be unsafe whereas a stable system can be either safe or not. Based on this relationship, we design the tandem structure to classify the system's stability first and then its safety only if stability is indicated. As the stability classification is simpler, faster, and more accurate than the safety assessment, the tandem structure can reduce (i) false negatives in the unsafety detection due to the ML's inaccuracy, and (ii) average execution time for the overall attack impact assessment since the safety classification can be skipped for a system determined to be unstable.

This paper applies the proposed assessment approach to two real-world CPSes: power grid automatic generation control (AGC) [9] and power plant control (PPC). AGC which is a critical component of existing power grids and whose complexity is representative of actual CPSes. The goal of the AGC is to maintain the grid frequency at a standard nominal value (e.g., $60\,\text{Hz}$) in the presence of load changes as primary exogenous disturbances. As the AGC's control signals are transmitted over communication networks, the delay attack is an important concern for AGC. We report extensive simulations using PowerWorld [10], an industry-strength power system simulator used by actual grid operators. The results show that AGC's stability depends on the delay and the total load only, whereas its safety additionally depends on the load changes and detailed distribution of the load among the load buses. While the boundary of the stable region can be obtained easily via a small set of offline simulations, enumerating all possible load distributions to characterize the safety boundary faces the aforementioned state explosion problem with respect to the number of buses. Instead, we apply the Monte Carlo method and use the extreme learning machine (ELM) [11] to learn the safety boundary. Furthermore, we use the achieved stability-safety classification to develop a two-tiered mitigation of the attack's impact. The mitigation regains the stability and safety of the AGC whenever needed, by tuning the AGC gain whenever possible and resorting to shedding load whenever the gain tuning is insufficient. We also apply our assessment approach to a PPC system modeled in Modelica, which consists of two PID control loops. Specifically, we generate offline training data using a Modelica simulator to learn the stability and safety boundaries by ELM and also evaluate the accuracy of the trained ELMs.

The rest of this paper is organized as follows. §II reviews related work. §III presents preliminaries and a motivating example. §IV overviews our approach. §V and §VI present the attack impact assessment and mitigation approaches, respectively. §VII and§VIII present the evaluation results for AGC and PPC, respectively. §IX concludes.

II. Related work

Power system stability and safety classifications are often studied separately in the literature. In [12], a Lyapunov method is used to classify a nonlinear system's stability without solving the differential equations. In [13], [14], the stability of a system is classified based on its energy accumulated during a certain time period. Traditional safety classification methods often analyze post-contingency power flows [15]. They use active power [15], [16] or composite indices based on various physical parameters [16] to classify the safety. However, the high computational overhead of these approaches makes them unsuitable for real-time classification [17], [18].

To reduce the computational overhead of real-time classification, recent studies apply ML (e.g., decision tree [19], support vector machine (SVM) [18], and artificial neural network (ANN) [17], [20], [21]) to classify a power system's stability [18], [19] and safety with respect to certain contingencies [20], based on measured physical conditions of the system. In [18], a trained SVM classifies the power system's stability by using phasor measurement unit data. The SVM will be retrained if the system condition has changed significantly. An ANN model used in [20] takes the system loading as input to rank the severity of the contingency in question, in terms of a composite performance index. However, all these studies do not address the emergent concern of cybersecurity.

Power grid cybersecurity has received increasing research. Chen et al. [6] study the impact on voltage and angle transient stability of data tampering attacks against voltage support devices. They do not study how to mitigate attack's impact. Farraj et al. [8] study how the delay attack on power grid sensor measurements may affect the system's transient stability, and design a parametric feedback controller to adaptively enhance the system's tolerance to the attack. Their work is based on a closed-form analytic model of the system, which is often unavailable for complex systems in practice.

Existing research on AGC's cybersecurity only focuses on false data injection (FDI) attacks [22]–[24], where the attacker tampers with sensor and/or control data in the AGC control loop. Specifically, the studies [22], [23] analyze the impacts of cyber-attacks on a two-area system's safety by reachability analysis. In [24], instead of qualitative reachability analysis, a quantitative analysis on the minimum time until the system is unsafe is proposed. Different from these studies on FDI that assume the adversary's non-trivial ability to corrupt data, this paper considers delaying the data packets between communicating system components, which is easier to implement and thus represents an even more present danger. Rahimi et al. [25] study the impact of the delay attack on AGC's stability through simulations of a three-area power system model. They do not consider the more subtle but equally critical property of safety, however. Moreover, they do not provide attack mitigation.

III. Stability and Safety under Delay Attack

This section defines stability and safety, as well as our threat model. Then, we use a simple control system to illustrate the impacts of the delay attack on the stability and safety.

A. System Model and Definitions of Stability and Safety

We consider a discrete-time CPS control system. Time is divided into slots. A *controller* collects measurements by the *sensors* in a *plant* and sends control commands to the *actuators*, which may change the state of the plant to maintain it at a given setpoint. The system is subjected to various disturbances, such as measurement noises, actuation biases, setpoint changes, etc. We adopt a bounded-input, bounded-output (BIBO) stability criterion:

Definition 1. *A system is* BIBO-stable *if its state remains bounded while it experiences bounded disturbances.*

We note that there are other stability definitions, e.g., asymptotic stability [26]. A system is asymptotically stable if for any positive ϵ, there exists a positive δ such that for any initial state of the system $x(0)$, the system's asymptotic equilibrium $\lim_{t\to\infty} x(t)$ satisfies $||x(t) - \lim_{t\to\infty} x(t)|| < \epsilon$, $\forall t \geq 0$, where $||x(0) - \lim_{t\to\infty} x(t)|| < \delta$. An asymptotically stable system is also BIBO-stable. Thus, BIBO stability is more basic and it is widely adopted in research on CPS control. For instance, the IEEE/CIGRE joint task force defines power system stability based on the BIBO concept [27]. In this paper, by *stability* we mean BIBO stability unless otherwise stated. Stability is a mandatory property for CPS design and operations.

We adopt the following safety definition.

Definition 2. *A system is* safe *if its state remains within a specified range while it experiences disturbances of magnitudes no larger than specified values.*

Safety is naturally a key concern of system operators, because devices are designed to function properly only within specified ranges. Crossing these ranges may damage the devices or cause system failures. From Definitions 1 and 2, note that stability describes a *qualitative* "bounded" nature of the system state, whereas safety additionally imposes a *quantitative* range of the bounds. Thus, stability is a more basic requirement in that an unstable system must be unsafe, but a stable system may not be safe. This relationship between the two different properties of a system will be exploited in §V to improve the performance (e.g., accuracy and timeliness) of the attack impact assessment for both the properties.

B. Threat Model

The delay attack is formally described as follows. Let $w(t)$ denote packetized control data generated and transmitted by the controller in the t^{th} time slot. The transmissions of the packets are maliciously delayed by τ time slots. Thus, in the $(t+\tau)^{\text{th}}$ time slot, the data $w(t)$ arrives at the actuator. Note that τ is an integer since the actuator operates in discrete time. The delay attack does not tamper with the content of the transmitted data. As §I discusses, it can be launched through a compromised router or by jamming communication channels using an industrial IoT botnet. Note that the delay τ can also include the natural communication latency.

In this paper, we assume that the clocks of the controller and the actuator are synchronized. Thus, if the controller adds a timestamp t to the transmitted data $w(t)$, the actuator can easily measure the delay τ introduced by the attack. The measured τ is used as an input to the attack impact assessment and mitigation. We note that secure clock synchronization techniques [1] can be used to ensure trustworthy measurements of τ. The scenario in which τ is unknown to the actuator (e.g., due to disrupted clock synchronization between the controller and actuator) is left to future work.

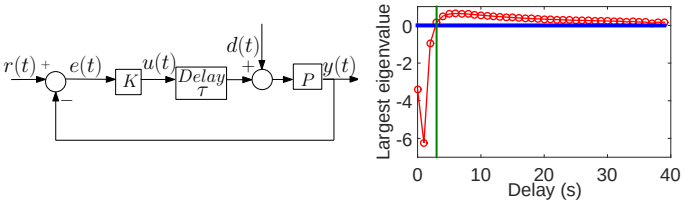
C. Stability and Safety of a Control System under Delay Attack

We use the feedback control system in Fig. 1(a) to illustrate the impacts of the attack on stability and safety. The results provide important observations that motivate the design of the attack impact assessment and mitigation approaches. In the absence of the attack, the system dynamics is

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}(\mathbf{u}(t) + \mathbf{d}(t)), \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t), \quad \mathbf{u}(t) = \mathbf{K}(\mathbf{r}(t) - \mathbf{y}(t)), \end{aligned} \tag{1}$$

where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{d}$, $\mathbf{r}$, and $\mathbf{u}$ are the system state, sensor measurement, disturbance, setpoint, and control signal, respectively; $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are system-specific matrices; $\mathbf{K}$ is a matrix characterizing the control law. Thus, the system employs proportional control. Note that the attack impact assessment and mitigation developed later in this paper do not depend on the control law. In particular, the AGC and PPC case studies employ proportional-integral (PI) and proportional-integral-derivative (PID) control. We consider the delay attack on $\mathbf{u}$, as illustrated in Fig. 1(a). Because of the attack, the $\mathbf{u}$ in Eq. (1) will be a delayed version $\mathbf{u}(t-\tau)$, which is given by $\mathbf{u}(t-\tau) = \mathbf{K}(\mathbf{r}(t-\tau) - \mathbf{y}(t-\tau) = \mathbf{K}(\mathbf{r}(t-\tau) - \mathbf{C}\mathbf{x}(t-\tau))$. Thus, Eq. (1) becomes

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) - \mathbf{B}\mathbf{K}\mathbf{C}\mathbf{x}(t-\tau) + \mathbf{B}\mathbf{K}\mathbf{r}(t-\tau) + \mathbf{B}\mathbf{d}(t). \tag{2}$$



(a) (b)

Fig. 1. A closed-loop control system under the delay attack: (a) System block diagram; (b) Largest eigenvalue vs. the delay.

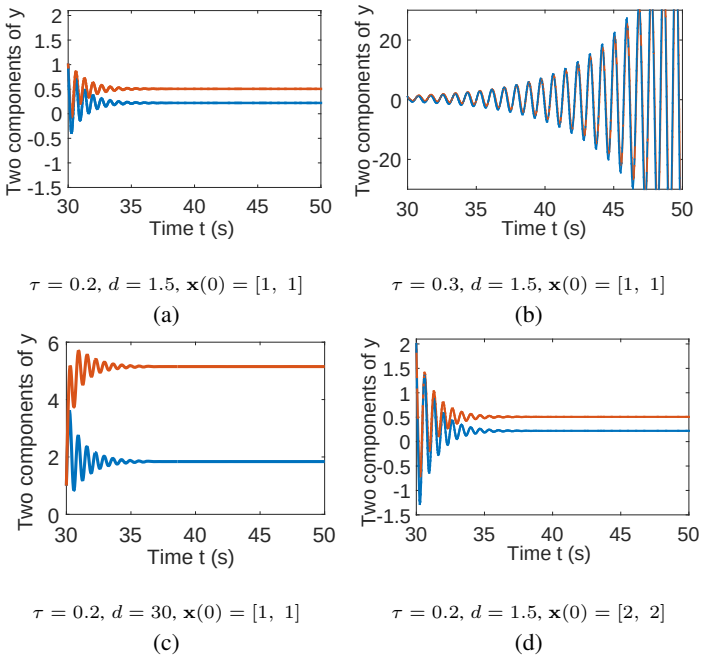
We now analyze the stability of the system in Eq. (2). By extending a result in [28], a necessary and sufficient condition for the stability of the system $\dot{\mathbf{x}}(t) = \hat{\mathbf{A}}\mathbf{x}(t) + \hat{\mathbf{B}}\mathbf{x}(t-\tau)$ is $\text{Re}\{\lambda_i(\hat{\mathbf{A}}+\mathbf{Q}(0))\} < 0$, where $\mathbf{Q}(0)$ is a non-singular solution of $e^{(\hat{\mathbf{A}}+\mathbf{Q}(0))\tau}\mathbf{Q}(0) = \hat{\mathbf{B}}$ and λ_i represents the ith eigenvalue of $\hat{\mathbf{A}} + \mathbf{Q}(0)$. By replacing $\hat{\mathbf{A}} = \mathbf{A}$ and $\hat{\mathbf{B}} = -\mathbf{BKC}$, we can analyze the system in Eq. (2). The numeric results in the rest of this section are based on the following settings: $\mathbf{A} = [-1\ -3;\ 3\ -5]$, $\mathbf{B} = [2\ -1;\ 1\ 0]$, $\mathbf{C} = [0.8\ \ 2.4;\ 1.6\ \ 0.8]$, $\mathbf{K} = 2$.

1) Impacts of delay on stability and safety: First, we analyze the impact of τ on the stability. The real component of the largest eigenvalue $\text{Re}\{\lambda_i(\hat{\mathbf{A}}+\mathbf{Q}(0))\}$ versus τ is shown in Fig. 1(b). When τ is larger than 0.2 s, the largest eigenvalue has a positive real component and so the system is unstable.

Second, we run time-domain simulations to understand the system's safety. The system output $\mathbf{y}$ over time under different settings is shown in Fig. 2. Both the delay against $\mathbf{u}$ and the step-change disturbance $\mathbf{d}$ of magnitude of 1.5 are introduced at $t = 30$ s. In Figs. 2(a) and 2(b), where $\tau = 0.2$ s and $\tau = 0.3$ s, the system is convergent and divergent, respectively. This result is consistent with the stability condition obtained from Fig. 1(b). The safety classification depends on how we define the safe range. For example, if we define the safe deviation range of $\mathbf{y}$'s components to be $[-1,\ 1]$, the system in Fig. 2(a) is safe. However, if the safe range is defined to be $[-0.4,\ 0.4]$, the system is unsafe. Thus, even if the system is stable, it can be either safe or unsafe, depending on the given safety conditions and the system's state trajectory.

2) Impacts of disturbance on stability and safety: Since stability is determined by the system's eigenvalues only, it is not affected by the disturbance $\mathbf{d}$, so that $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ do not include $\mathbf{d}$. In contrast, as safety depends on the trajectory of $\mathbf{y}$, which depends on $\mathbf{d}$, the magnitude of $\mathbf{d}$ can significantly affect the safety. We now illustrate this observation using Fig. 2(c) that has the same setting as Fig. 2(a) except that the disturbances in Fig. 2(a) and Fig. 2(c) are 1.5 and 30, respectively. Fig. 2(c) shows larger output deviations, which may violate the safety requirement.

3) Impacts of initial state on stability and safety: As the system's eigenvalues do not depend on the initial system state, the stability does not depend on the initial state. In contrast, since the initial state affects the system trajectory, it affects the system's safety. For instance, Fig. 2(d) has the same setting as Fig. 2(a) except that they have different



$\tau = 0.2, d = 1.5, \mathbf{x}(0) = [1,\ 1]$ (a)

$\tau = 0.3, d = 1.5, \mathbf{x}(0) = [1,\ 1]$ (b)

$\tau = 0.2, d = 30, \mathbf{x}(0) = [1,\ 1]$ (c)

$\tau = 0.2, d = 1.5, \mathbf{x}(0) = [2,\ 2]$ (d)

Fig. 2. The system output $\mathbf{y}$ under different settings.

initial states. The system remains convergent in this case, which generally implies a stable system. However, the output deviation is doubled compared with that of Fig. 2(a), and the larger deviation may violate safety.

In summary, we have these two observations: (i) the delay τ affects both stability and safety, (ii) the safety depends on the disturbance and the system's initial state, while the stability does not. These observations will guide the design of the proposed tandem stability-safety assessment method.

IV. Objective and Approach Overview

A. Objective and Challenges

We aim to develop delay attack impact assessment and mitigation for the CPS control. The input for the assessment includes the measured delay τ and the measurements of sensors monitoring the system state. If the system is classified unsafe (i.e., it will enter an unsafe region), mitigation actions should be initiated to regain safety.

We face the following main challenges. First, although we can obtain an analytic stability condition for the simple system in Fig. 1(a), it is challenging to obtain similar conditions for real-world complex systems. Second, the safety classification needs the system's trajectory such as those shown in Fig. 2. Although we can use a high-fidelity simulator to predict the trajectory, the transient simulations for complex systems can be too slow for real-time online prediction and control. For instance, a transient simulation for the 37-bus power grid shown in Fig. 4 takes 138 s on a 28-core computing server, while the grid under attack takes less than two minutes to cross its safe range (cf. Table II). Thus, the system will have well entered the unsafe region by the time the transient simulation completes. Third, as locating and removing an

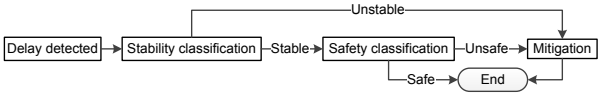


Fig. 3. Attack impact assessment and mitigation pipeline.

ongoing cyber-attack often takes significant time, before the attack is removed, it is critical to tolerate the attack and mitigate its impact by adapting tunable system parameters and settings. However, a model that characterizes the effects of the new parameters and settings on the safety will be needed to determine their suitable values. It is similarly challenging to obtain this model for complex systems.

B. Approach Overview

This section overviews our approach. In every time slot, if the measured total delay τ in transmitting sensor measurements and control commands exceeds a threshold (e.g., the typical communication delay), we execute the attack impact assessment and mitigation pipeline shown in Fig. 3. First, we classify system's stability. If the system is unstable, which implies that it is unsafe, we initiate mitigation to restore safety; otherwise, we classify system's safety. If and only if the system is classified unsafe, we initiate mitigation. We now discuss the design of the stability and safety classification, as well as the mitigation, that addresses the challenges described in §IV-A.

First, since it is difficult to analyze the stability and safety of complex systems, we use a simulation-based approach. We assume that a high-fidelity simulator that can accurately characterize the system dynamics is available. This assumption agrees with practice. For instance, power grid operators generally maintain high-fidelity simulators of their systems to guide design and operations. Using the simulator, we can explore key factors that affect the system's stability and safety.

Second, since the transient simulations, though accurate, are generally too slow for online use, we conduct offline simulations to generate extensive data with appropriate stability and safety labels. The labeled data will be used to characterize the stability and safety boundaries. However, the dependence of safety on the system's initial state, as illustrated in §III-C3, leads to state explosion if we were to enumerate all the initial states during the generation phase of training data. To deal with this issue, we apply a Monte Carlo method to generate the training data and train an ML model to characterize the safety boundary. The ML model can also be used to guide the search for suitable mitigation actions.

Third, the ML model may err occasionally in the safety classification. On the other hand, as observed from case studies in §III-C and §V, the stability classification is simpler, faster, and more accurate. Thus, we apply the stability classification first in the overall assessment, so that we can condition the safety classification on the more reliable and faster stability classification result. This conditional sequential strategy reduces the overall classification errors and runtime overheads.

We note that the detailed design of the components shown in Fig. 3 is system specific. However, we believe that the basic design paradigm is applicable to a wide range of CPSes. In the rest of this paper, we will apply it to the domains of AGC and PPC, which are fundamental control systems used in real-world power systems, and design accordingly the domain-specific components. Note that we will primarily focus on the AGC case study in §V, VI, and VII first. Then, we similarly apply the approach to the PPC case study and present a summary of the results in §VIII due to space constraints.

V. Stability-Safety Assessment for AGC

Since AGC involves long-range communications and its malfunction can cause grid-wide failures and infrastructure damage, it can be an attractive target for attackers. In §V-A, we present necessary background of the AGC for our discussions. §V-B presents extensive simulations to understand the AGC's stability and safety under the delay attack. §V-C applies the proposed tandem stability-safety assessment to the AGC.

A. Background of AGC

AGC maintains the grid frequency at a nominal value (e.g., $60\,\mathrm{Hz}$) by adjusting the setpoints of generators. It also maintains the net power interchanges among neighboring areas at scheduled values [9]. Here, an area is a part of the grid and it is usually operated by a utility. Two areas are connected by *tie-lines*. Fig. 4 illustrates a three-area 37-bus system[1], where the dotted lines represent the tie-lines. As illustrated in Fig. 5, the AGC, located in the grid control center, receives over a communication network measurements of the deviations of the grid frequency (from the standard frequency) and the ith area's power export from their respective setpoints (which are denoted by $\Delta\omega_i$ and ΔP_{Ei}), and it computes the *area control error* (ACE) as $\mathrm{ACE}_i = \alpha_i \cdot \Delta P_{Ei} + \beta_i \cdot \Delta\omega_i$, where α_i and β_i are two constants. The control center sends ACE_i to the area's power plants over the communication network. Each plant applies PI controller with a gain of k to generate a reference signal for its generator. Specifically, the reference signal is $-k\int \mathrm{ACE}_i(t)\mathrm{d}t$. The above process is repeated every AGC cycle, which is often two to four seconds. We note that this reference signal is the power setpoint of the PPC system studied in §VIII.

The sensor measurements and ACE are transmitted in long-range communication networks that are susceptible to cybersecurity threats. In this paper, we focus on the delay attack against transmissions of ACE signals. However, our approach can be readily applied to delay attacks on sensor measurements, or both ACE signals and sensor measurements.

B. AGC's Stability and Safety under Delay Attack

This section presents two extensive simulation studies to investigate how the following factors may affect the AGC's stability and safety: (i) the grid's total load, (ii) the distribution of the load among the load buses, (iii) the change of load,

[1]We use the 37-bus system as a case study throughout this paper. It is a test system [29]. Its scale corresponds to a small-/mid-scale grid in real life. According to our rough count based on a grid topology database (http://bit.ly/2vRH5Nd), a major fraction of 130 national grids consist of fewer than 37 buses.

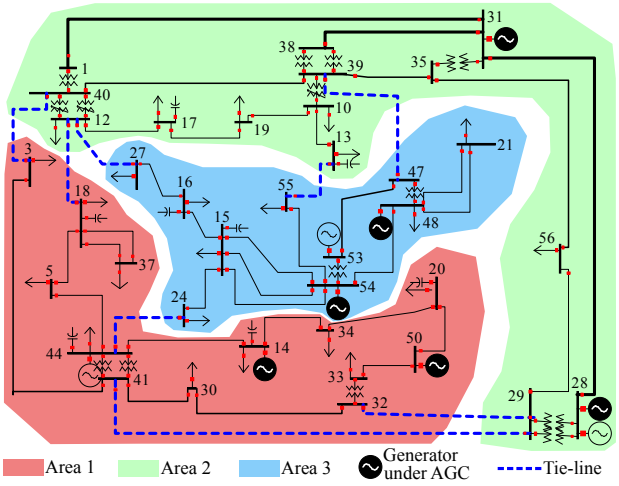


Fig. 4. A three-area 37-bus system. Each area is a part of the grid and operated by a utility. Two areas are connected by tie-lines, i.e., the dashed lines in the figure.

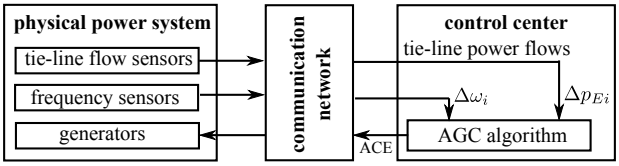


Fig. 5. Overview of AGC.

and (iv) the communication delay. We note that the load distribution determines the power system's state, which is often defined as the union of all the buses' voltage phasors. Thus, the total load can be considered a statistics of the system's initial state. The load change is the primary exogenous disturbance to the AGC. The simulations are conducted using PowerWorld, an industry-strength high-fidelity power system simulator, based on the system model in Fig. 4. The main simulation settings are: the length of a time slot is 1 s; the length of an AGC cycle is 4 s; each simulation lasts for 300 s; the delay attack on the ACE signal is launched at $t = 120$ s; the load change occurs at $t = 140$ s.

1) AGC's stability: The stability is assessed by checking the system's convergence. We have the following observations.

AGC's stability depends on the total load: Fig. 6 shows the AGC's stability under different total loads and delays, where a blue/red point means that the system is stable/unstable, respectively. A total of 7,900 combinations of the total load and delay are tested. We can see that the total load affects the maximum delay that the system can tolerate to keep stable. For instance, when the total load is 600 MW, the maximum tolerable delay is 6 s. When the total load is 1000 MW, the maximum tolerable delay is 2 s only. Fig. 6 also shows a clear cut boundary between the stable and unstable regions.

AGC's stability is independent of the detailed load distribution: We fix the total load at 795 MW and distribute it among the load buses randomly. Simulations using 1,000 random load distributions show that the maximum tolerable

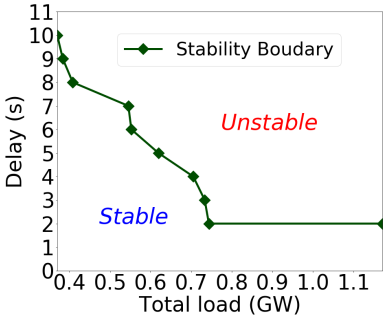


Fig. 6. AGC's stability under different total loads and delays.

TABLE I
MAX TOLERABLE DELAY.

Load change	Total load 715	795	874
–10%	5	3	3
–5%	5	3	3
0	5	3	3
5%	5	3	3
10%	5	3	3

[a]The delays are in seconds.
[b]Total Loads are in MW.

TABLE II
TIME TO CROSS THE SAFE RANGE VS. DELAY AND LOAD CHANGE.

Load change (MW)	Delay (s) 0	1	2	3
-80	105.45	105.45	105.7	105.8
-40	∞	∞	∞	276.1
0	∞	∞	∞	944.3
40	148.6	148.6	148.6	148.6
80	146.1	146.1	146.1	146.1

*The time values are in seconds; ∞ means the system is safe.

delay is always 2 s. Under other settings of the total load, the maximum tolerable delay is also a constant over the different load distributions. This gives strong empirical evidence that the AGC's stability is independent of the load distribution. The observation is consistent with the standard practice of analytical modeling of AGC, which considers the total load only but not the load distribution [9].

AGC's stability is independent of load change: Table I shows the maximum tolerable delay under different settings of the total load and the load change as percentage of the total load. The load change consists of step changes at all the load buses at $t = 140$ s. The step change is realistic given increasing adoption of demand response and distributed renewable energy sources that can trigger sudden changes in load. From the table, for each tested total load setting, the AGC's stability is unaffected by the change. This result is consistent with our discussions in §III-C2. Moreover, with less total load, the system can tolerate longer delays, which is consistent with the results in Fig. 6.

2) AGC's safety: We impose the following two safety requirements. First, the grid frequency deviation must be within $[-0.5\,\text{Hz}, 0.5\,\text{Hz}]$. In real systems, if the deviation exceeds this safe range, disruptive remedial actions such as load shedding will be automatically initiated to protect the grid from infrastructural damage [9]. Second, the power flows must be within capacities of the transmission lines. Otherwise, the lines will trip due to overheating. In our simulations, we adopt the default line capacities of the 37-bus system.

AGC's safety depends on load change: The total load is 800 MW. Table II shows the time from the launch of the delay attack to the breach of the safety requirement under different delays and load changes. The symbol ∞ means that the safety limits are never crossed, i.e., the system is safe. From the table,

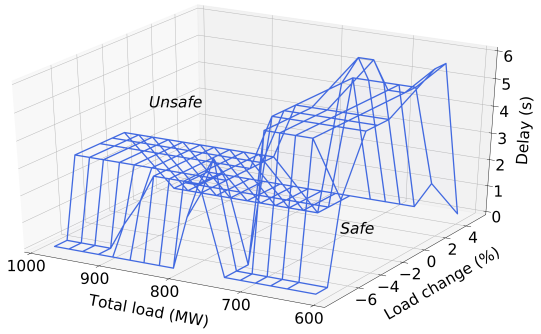


Fig. 7. The minimum delay leading to unsafety vs. total load and load change.

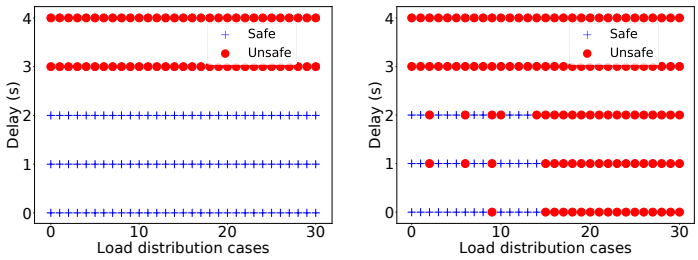


Fig. 8. AGC's safety under different load distributions. (a) Infinite line capacity; (b) Finite line capacity.

the AGC's safety is affected by the load change, which is consistent with our discussion in §III-C2. For instance, when the load change is 5% of the total load (i.e., 40 MW), the system will be unsafe, regardless of the delay. When the load change is small, the system will be safe if the delay is also small. Thus, the load change and delay jointly affect the safety.

AGC's safety depends on total load: Fig. 7 shows the minimum delays that lead to unsafety under different total loads and load changes. Each grid point represents such a minimum delay obtained by running a set of simulations under different delays. Note that, to simplify the illustration, we relax the transmission line capacities to infinite, such that the load distribution does not affect the safety. The next set of experiments will show the impact of the load distribution on the safety under finite line capacities. In Fig. 7, the surface formed by the grid points that represent the obtained minimum delays leading to unsafety divides the space into safe and unsafe regions, which are below and above the surface, respectively. The result shows that the total load, the load change, and the delay jointly affect the AGC's safety.

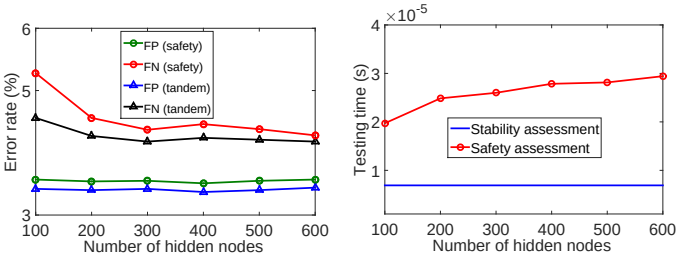
AGC's safety depends on load distribution: We fix the total load at 800 MW and distribute it among the load buses randomly. Fig. 8(a) and Fig. 8(b) show the classification of the AGC's safety given different delays in 30 cases of the load distributions, when the line capacities are set to be infinite and finite, respectively. Although the line capacities are finite in practice, we present the infinite case to help understand the affecting factors of the AGC's safety. Under infinite line capacities, the AGC's safety depends on the frequency deviation only. The deviation depends on the total load, rather than the load distribution. Thus, in Fig. 8(a), the safety is independent of the load distribution. In contrast, since power flows depend on the load distribution, under finite line capacities, the load distribution will affect the AGC's safety. In Fig. 8(b), for a given delay, the system may be safe or unsafe depending on the load distribution.

3) Summary: The above experiments show that the AGC's stability depends on the total load and the delay, while its safety additionally depends on the load change and the load distribution. This observation is mostly consistent with that for the barebone control system in §III-C, except that the AGC's stability depends on the total load, a statistics of the system state. This can be explained from the fact that AGC is a nonlinear system, although its control-theoretic analysis is often based on a linearization at the system's current condition as characterized by the total load [9]. Thus, the AGC's stability condition is also affected by the total load. However, this minor deviation will not impede the application of the tandem stability-safety assessment, since the scalar total load will not lead to a state explosion problem.

C. Stability-Safety Assessment for AGC under Delay Attack

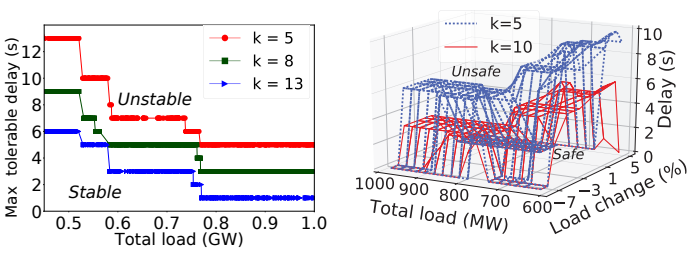
This section applies the proposed tandem stability-safety assessment to AGC. From Fig. 6, since the AGC's stability has a clear cut boundary in the two-dimensional space formed by the total load and the delay, it can be classified quickly at run time based on the boundary *a priori* obtained through extensive offline transient simulations. We call this classification approach *boundary-based stability classification*. Specifically, if the system's current operating point (i.e., total load and delay) is below the boundary, such as that shown in Fig. 6, the system is stable; otherwise, it is unstable. This classification avoids running a time-consuming online transient simulation based on the system's current operating point. In particular, due to the limited dimension of the stability space (i.e., two), we can achieve any granularity in enumerating operating points within any specified range. As a result, the boundary-based approach achieves perfect classification accuracy asymptotically as the enumerating granularity goes to zero.

In contrast, AGC's safety additionally depends on the load distribution vector, which has exponential complexity with respect to the number of load buses that is often tens to hundreds. To avoid the exponential complexity, we use a Monte Carlo method to randomly sample the operating points in a discretized state space and generate extensive offline simulation results with determined safety labels to train an ELM [11] to characterize the AGC's safety. The ELM is a single hidden layer feedforward neural network with a training algorithm much faster than conventional gradient-based learning algorithms. At run time, the trained ELM classifies the AGC's safety based on the current operating point (i.e., total load, load change, load distribution, and delay). In



(a) (b)

Fig. 9. FP, FN rates, and testing time versus the number of hidden nodes in ELM. (a) FP and FN rates. (b) Testing time.



(a) Stability boundaries (b) Safety boundaries

Fig. 10. System stability and safety boundaries under different k settings.

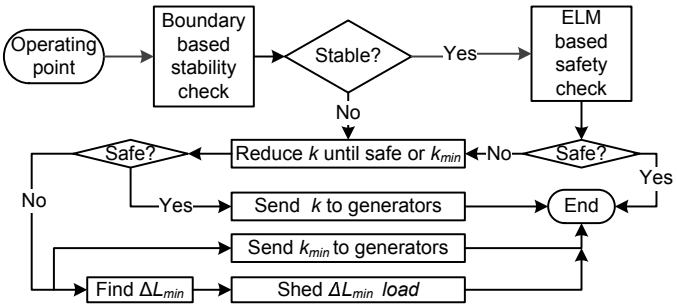


Fig. 11. Two-tiered delay attack impact mitigation.

§VII, we will compare the performance of the ELM with a baseline approach that also uses the training data to classify safety.

We present the following numeric results to show the effectiveness of the ELM-based safety assessment. The training and testing data sets consist of 11,000 and 7,000 operating points and their safety labels, respectively. We use the false positive (FP) and false negative (FN) rates as the accuracy metrics, which are the percentages of safe (resp., unsafe) cases that are wrongly classified to be unsafe (resp., safe). The green and red curves in Fig. 9(a) show the ELM's FP and FN rates versus the number of hidden nodes in the ELM. The two rates are generally below 5%. In §VII-C, we will discuss how to deal with the FPs and FNs. When the number of hidden nodes is 300, both the two rates reach their knee points. Thus, 300 is a satisfactory setting, since using more hidden nodes does not improve the accuracy much, but increases the testing time as shown by the red curve in Fig. 9(b). Under the setting of 300, the testing time is around 0.03 ms only on an Intel i7 2.2GHz CPU. This time is short compared with the time horizon of a power grid's fault clearing (e.g., 200 ms for lightning strike overcurrent clearing). The testing time can be further reduced significantly by using hardware acceleration.

Lastly, we show the benefits of the tandem stability-safety assessment. First, as the boundary-based stability classification gives asymptotically perfect accuracy, it helps reduce FNs of the ELM-based safety classification. The blue and black curves in Fig. 9(a) show the FP and FN rates of the tandem stability-safety assessment. FN rate is reduced by up to 1%. Second, the blue curve in Fig. 9(b) shows the testing time of the boundary-based stability classification, which is 11 microseconds only, 3 times shorter than that of the ELM's testing time with 300 hidden nodes. Thus, under the tandem approach, any instability will be detected by the fast stability classification, which improves the timeliness of the needed mitigation (cf. §VI). In §VII-C, we will evaluate the impact of an FP and describe an approach to further reduce the FN rate.

VI. Mitigating Impact of Attack against AGC

This section presents an approach to mitigating the delay attack impact on AGC. As the total load is an important determining factor for both stability and safety, a feasible approach is to shed load to restore safety. However, clearly, load shedding will affect customers adversely, sometimes severely. Hence, it should be avoided if possible. This section proposes a two-tiered approach that firstly tunes the AGC gain as a first-line defense, and resorts to shedding load only when the gain tuning is insufficient. This section studies the impact of the gain on the AGC's stability and safety first in §VI-A. Then, it presents the two-tiered approach in §VI-B.

A. Impact of AGC Gain on Stability and Safety

As discussed in §V-A, each power plant applies a PI controller with a gain of k to the received ACE to produce a reference signal for the plant's generator. We conduct simulations based on the 37-bus system model to investigate the impact of k on the AGC's stability and safety. The curves and surfaces in Figs. 10(a) and (b) show the stability and safety boundaries, respectively, under different settings of k. By reducing k, we can expand the stable and safe regions. However, from control theory, a smaller k will result in slower convergence when there is a load change. Hence, we have a trade-off between (i) AGC's tolerance to the delay in terms of stability and safety, and (ii) AGC's convergence speed in response to a load change. As AGC generally also needs to meet some required convergence speed, there exists in practice a minimum allowable setting for k [9], which is denoted as $k_{\min}$. Multiple ELMs are trained to characterize the safety boundaries under different settings of k. This *ELM bank* will be used in §VI-B to find a k to restore safety where needed.

B. Two-Tiered Delay Attack Impact Mitigation

Fig. 11 illustrates the integrated stability-safety assessment and attack impact mitigation. When a system is classified unstable or unsafe, the two-tiered mitigation is activated. No mitigation is needed only when the system is classified safe. The two-tiered mitigation works as follows. First, within the

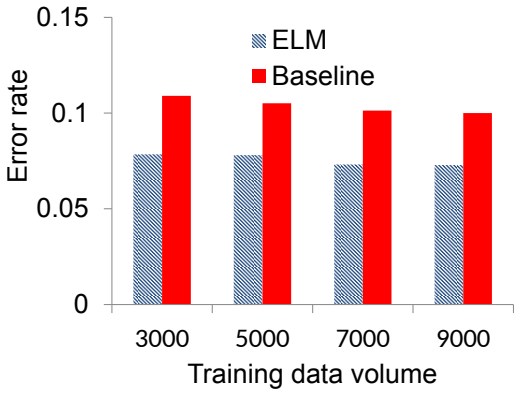


Fig. 12. Comparison of ELM-based and data-driven baseline approaches.

range from $k_{\min}$ to the current setting of k, we search for the maximum setting of k that can restore safety using the ELM bank discussed in §VI-A. If such a k setting is found, it is piggybacked onto the next ACE signal that will be sent to generators. Otherwise, load shedding should be applied. We use the ELM bank to find the minimum amount of load that needs to be shed to restore safety under the setting $k_{\min}$. This minimum amount is denoted by $\Delta L_{\min}$. The grid operator sheds $\Delta L_{\min}$ load and piggybacks the $k_{\min}$ to the next ACE signal that will be sent to generators. The shedding amount can be shared among load buses equally or using existing scheduling algorithms addressing other grid operation optimization objectives and constraints [30]. Once a generator receives the new AGC gain, it updates its setting accordingly.

VII. Performance Evaluation

This section evaluates several key aspects of our attack impact assessment and mitigation designed for the AGC of the 37-bus system shown in Fig. 4.

A. Effectiveness of ELM-Based Safety Classification

We compare the proposed ELM-based approach with a data-driven baseline approach. Specifically, the baseline finds a system operating point within the ELM's training data that has the smallest Euclidean distance to the system's current operating point, and yields the found operating point's safety label. Fig. 12 shows the classification error rates of our ELM-based and the baseline approaches under different settings of training data volume. Consistent with intuition, the error rate decreases with the volume of training data. The ELM-based approach gives lower error rates. Moreover, the running time for ELM-based approach is up to 6,000 times shorter than that of the baseline approach.

B. Effectiveness of Attack Mitigation

We conduct two simulations to show the effectiveness of our two-tiered attack mitigation. The system's total load is 1000 MW. The initial setting for k is 10. The safety requirement for the grid frequency deviation is $[-0.5\,\text{Hz}, 0.5\,\text{Hz}]$.

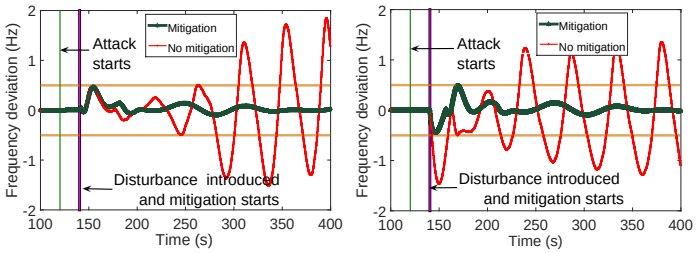
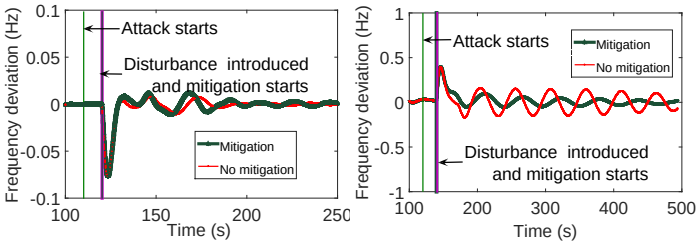


Fig. 13. Attack impact mitigation examples. (a) Tuning k only; (b) Tuning k and shedding load.

The attacker delays the ACE signal by 4 s from $t = 120$ s. The attack impact assessment classifies the system safe until a step load change is introduced at $t = 140$ s. In Fig. 13(a), the load change is 5% of the total load. At this point, the system is classified unsafe. The red curve in Fig. 13(a) shows the system's trajectory if no mitigation is applied. It confirms the assessment result. The mitigation approach starts searching for a k setting to regain safety. By decreasing k from 10 to $k_{\min} = 5$, the system is classified safe under the attack. The thick green curve in Fig. 13(a) shows the system's trajectory after the new setting $k = 5$ is applied. We can see that the system becomes safe after the mitigation. In Fig. 13(b), the load change is 8% of the total load. Because of the increased load change, tuning k to $k_{\min} = 5$ is insufficient and shedding 10% of load is needed to restore safety. The thick green curve in Fig. 13(b) shows the system's trajectory after load shedding and reconfiguring k. The system is safe after the mitigation. The effects of different mitigation approaches on the customers are different. In Fig. 13(b), as tuning k to $k_{\min}$ still cannot mitigate the attack impact, we have to shed some of the customer loads, which results in lower utility to the owners. In Fig. 13(a), as the mitigation is achieved by adjusting the AGC parameters only, no customers will be affected.

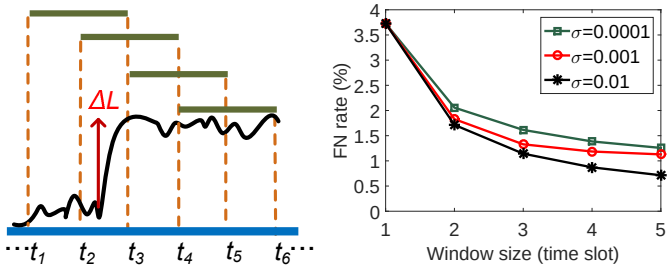
C. False Positives and Negatives in Safety Classification

While the ML deals with the state explosion problem, it results in FPs and FNs. An FP will trigger the attack mitigation. Fig. 14(a) shows the system's trajectory after the mitigation wrongly triggered by a safety classification FP caused by a load change that is 0.5% of the total load, where the ACE signal is delayed by 2 s from $t = 120$ s. As the mitigation applies a small adjustment only (i.e., decrease k from 10 to 8), the frequency deviation has a slightly longer settling time. Moreover, Fig. 14(b) shows another scenario of the system's trajectory after the mitigation wrongly triggered by a safety classification FP caused by a load change that is 0.5% of the total load, where the ACE signal is delayed by 5 s from $t = 120$ s. The mitigation sheds 8% of the total load after decreasing k from 10 to 5, the frequency deviation can even have a smaller setting time. This is the because the mitigation speeds up the system to diminish the small fluctuations due to the delay. Therefore, as FPs mostly occur for marginally safe operating conditions, the caused mitigation is generally of



(a) (b)

Fig. 14. Mitigation wrongly is triggered. (a) By a safety classification false positive and k is tuned; (b) By a safety classification false positive and load shedding is conducted.



(a) (b)

Fig. 15. (a) sliding window approach. The time window is set as two time slots and the step load change will be assessed twice at $t = t_3$ and $t = t_4$. (b) FN rate vs. window size. The window size is increased from 1 to 5 and three different random load fluctuations standard deviations are illustrated.

small strength that leads to slight settling time increase or even can help the system to decrease the settling time, mitigating the concern of FPs.

In contrast, the system may become unsafe due to FNs. We discuss a sliding window approach as illustrated in Fig. 15(a) to reduce the FNs. In this approach, the load change is defined as the difference between the current load and the load that is a time window before. As a result, a step load change will be assessed for multiple times. For instance, in Fig. 15(a), the time window is two time slots and the step load change will be assessed twice at $t = t_3$ and $t = t_4$. Due to the random temporal fluctuations of the load, the probability that an unsafety can be detected in at least one of the multiple assessments will increase, thus reducing the FN rate. By increasing the window size, a load change will be assessed for more times. Fig. 15(b) shows the FN rate versus the window size under different random load fluctuations' standard deviations (σ). The FN rate decreases with the window size. Thus, this approach can effectively reduce the FN rate. The concern of increased FP rate due to this approach is minor since the FPs cause little impact on the system as illustrated earlier.

D. Impact of Load Change Trajectory

In the previous sections, the disturbances to the system are modeled as step load changes. In practice, the load change may take time. For instance, the customers' solar power generation may change due to the movement of clouds and the ensuing load changes may take tens of seconds. This section investigates how the trajectory of the load change may affect the system's stability and safety. Fig. 16 shows

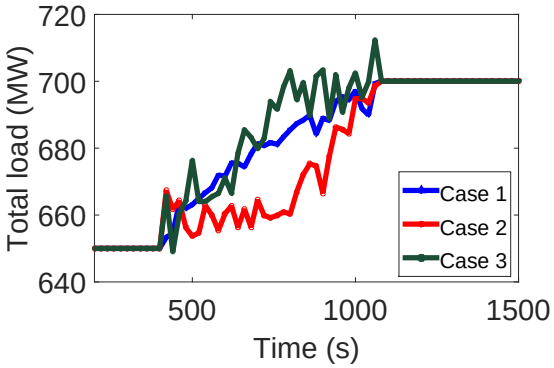
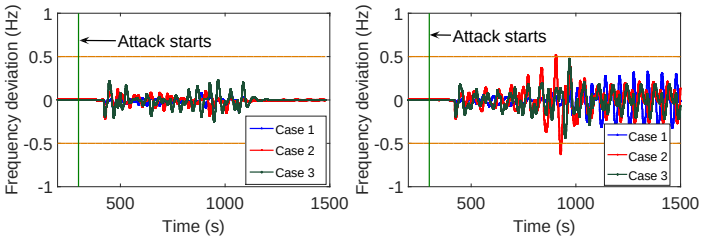


Fig. 16. Trajectories of load increasing. The load starts to change from $t = 450s$ and lasts 600 s. The total load increases by 50 MW in all three randomly generated trajectories.



(a) (b)

Fig. 17. System output for different load increase trajectories in Fig. 16 under different delays. (a) Delay by 3 s (b) Delay by 6 s.

three trajectories for a load change from 650 MW to 700 MW. The trajectories are generated randomly. Fig. 17 shows the system frequency deviation for the different trajectories when the system experiences delays of 3 s and 6 s. From the figure, the system's stability is not affected by the load trajectories. Specifically, in Fig. 17(a), the system frequency converges to the nominal value. In Fig. 17(b), although the system experiences oscillations around the nominal value, the system will eventually converge, although this is not shown in Fig. 17(b) to focus instead on the details around $t = 900$. In contrast, the load trajectories can affect the system's safety. For example, at around $t = 900$, the system frequency of Case 2 crosses the safety range due to some large fluctuations of load in the trajectory. Hence, considering the change trajectory (e.g., by using recurrent neural network [31] to capture the effects of time-series data) may improve the safety assessment. We leave this task to future work.

VIII. Stability-Safety Assessment for PPC

In this section, we apply the stability-safety assessment approach to a power plant control (PPC) system. We use a PPC model in ThermoPower [32], an open-source library based on Modelica. Note that Modelica is an object-oriented complex physical system modeling language [33]. The signal flow graph of the system is shown in Fig. 18. The controlled power plant admits two inputs, the power control signal and the void fraction control signal. The void fraction is also known as

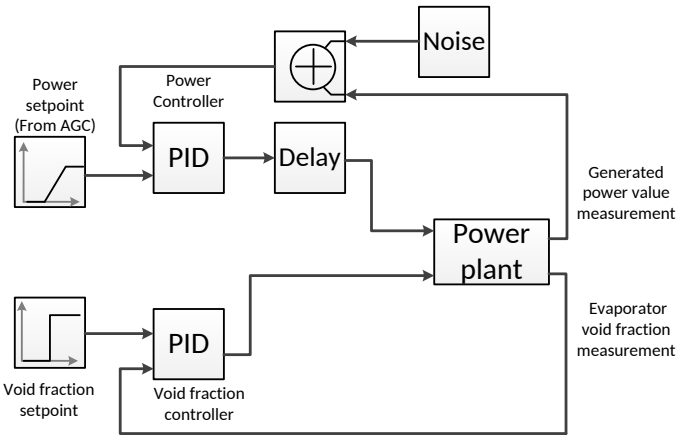


Fig. 18. A PPC system.

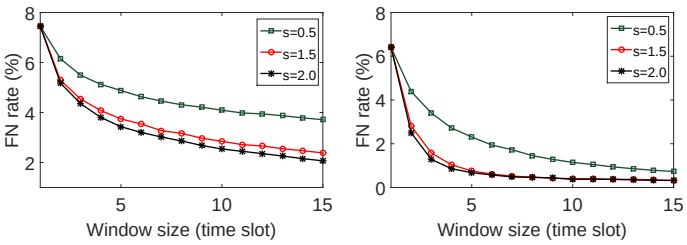


Fig. 19. FN rates of stability (a) and safety (b) assessment vs. window size. The s is the noise generator sampling period. A smaller s value means a larger level of disturbance to the system.

porosity, which is an important parameter characterizing two-phase fluid flow, especially gas-liquid flow. The two control signals are determined respectively by two PID controllers. The power controller's feedback signal is corrupted by additive zero-mean Gaussian noises acting as disturbances to the system. The adversary delays the power controller's output signal. We use offline simulations together with ML to model the system's stability and safety. Specifically, we use OpenModelica [34], a Modelica-based simulator, to run massive offline simulations under a wide range of settings to generate training data. Then, we train the ELMs to model the stability and safety. Inputs to the ELMs include the sampling period and variance of the Gaussian noise as well as the delay. The output is the stability or safety assessment result. We apply the sliding window approach illustrated in Fig. 15 to improve the FN rate. The training and testing data sets consist of 10,000 and 3,500 operating points, respectively.

Figs. 19(a) and 19(b) show the FN rates of the stability and safety assessments, respectively, under various settings of the sliding window size. The different curves are the results under different settings of the Gaussian noise generator's sampling period in seconds, i.e., the s values in the legends. We note that different from the simple control loop in §III-C the AGC, which are discrete-time systems, the PPC is a hybrid system with discrete-time sensing but continuous-time control and actuation. Thus, the Gaussian noise generation, which belongs to the sensing part, is in discrete time. As a result, the frequency at which we update the noise affects the level of the disturbance to the system. Specifically, a smaller noise sampling period causes a higher disturbance. From Figs. 19(a) and 19(b), similar to the results for the AGC, the FN rate increases with the window size and decreases with the disturbance level. For attack mitigation, we can build ELMs for a range of PID configurations (i.e., the P, I, and D coefficients) and then apply the mitigation approach presented in §VI to tune the PID configuration. Due to space constraints, the results are omitted here.

IX. Conclusion

This paper presented a real-time delay attack impact assessment approach that applies a stability classifier and an ML-based safety classifier sequentially. The ML addresses the state explosion problem in the safety classification due to the dependence of the system's safety on the multi-dimensional system state. The tandem stability-safety design improves the accuracy of the unsafety detection and speeds up the overall assessment. We applied our approach to power grid AGC, and developed a two-tiered attack impact mitigation that tunes the control gain as a first-line defense and resorts to shedding load only if the gain tuning is insufficient to regain safety. Simulations based on a 37-bus system model verified and illustrated the effectiveness of our assessment and mitigation approaches. We also applied our approach to assess the stability and safety of a PPC system and presented the evaluation results based on Modelica simulations.

References

[1] S. Viswanathan, R. Tan, and D. Yau, "Exploiting power grid for accurate and secure clock synchronization in industrial IoT," in *IEEE Real-Time Systems Symposium (RTSS)*, 2016.

[2] Y. Zhang and V. Paxson, "Detecting stepping stones," in *USENIX Security Symposium*, 2000.

[3] "Hackers infiltrated power grids in U.S., Spain," https://goo.gl/DUWT1o.

[4] Y. Zhang and V. Paxson, "Stuxnet worm impact on industrial cyberphysical system security," in *IECON*, 2011.

[5] "DDoS attack that disrupted internet was largest of its kind in history," https://goo.gl/Iu0wjc.

[6] B. Chen, S. Mashayekh, K. Butler-Purry, and D. Kundur, "Impact of cyber attacks on transient stability of smart grids with voltage support devices," in *IEEE PES General Meeting*, 2013.

[7] X. Cao, P. Cheng, J. Chen, S. Ge, Y. Cheng, and Y. Sun, "Cognitive radio based state estimation in cyber-physical systems," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 3, pp. 489–502, 2014.

[8] A. Farraj, E. Hammad, and D. Kundur, "A cyber-physical control framework for transient stability in smart grids," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 847–855, 2013.

[9] P. Kundur, *Power System Stability and Control*. McGraw-Hill, 1994.

[10] *PowerWorld (version 17)*, www.powerworld.com.

[11] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1, 2006.

[12] A. M. Lyapunov, *The General Problem of the Stability of Motion*. London, U.K.: Taylor and Francis, 1992.

[13] P. Magnusson, "The transient energy method of calculating stability," *Trans. Amer. Inst. Electr. Eng.*, vol. 66, no. 1, pp. 747–755, 1947.

[14] H. D. Chiang, "Study of the existence of energy functions for power systems with losses," *IEEE Trans. Circuits Syst.*, vol. 36, no. 11, 1989.

[15] T. A. Mikolinnas and B. F. Wallenberg, "An advanced contingency selection algorithm," *IEEE Trans. PAS.*, vol. 100, no. 2, 1981.
[16] V. Brandwajn, Y. Liu, and M. Lauby, "Pre-screening of single contingencies causing network topology changes," *IEEE Trans. Power Syst.*, vol. 6, no. 1, pp. 30–36, 1991.
[17] R. Fischl, "Application of neural networks to power system security: Technology and trends," in *IEEE World Congr. Comput. Intell.*, 1994.
[18] B. Wang, B. Fang, Y. Wang, H. Liu, and Y. Liu, "Power system transient stability assessment based on big data and the core vector machine," *IEEE Trans. Power Syst.*, vol. 7, no. 5, pp. 2561–2570, 2016.
[19] M. He, J. Zhang, and V. Vittal, "Robust online dynamic security assessment using adaptive ensemble decision-tree learning," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4089–4098, 2013.
[20] T. S. Sidhu and C. Lan, "Contingency screening for steady-state security analysis by using FFT and artificial neural networks," *IEEE Trans. Power Syst.*, vol. 15, no. 1, pp. 421–426, 2000.
[21] Y. Xu, Z. Y. Dong, J. H. Zhao, P. Zhang, and K. P. Wong, "A reliable intelligent system for realtime dynamic security assessment of power systems," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1253–1263, 2012.
[22] P. Esfahani, M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "A robust policy for automatic generation control cyber attack in two area power network," in *IEEE Conference on Decision and Control (CDC)*, 2010.
[23] ——, "Cyber attack in a two-area power system: Impact identification using reachability," in *American Control Conference (ACC)*, 2010.
[24] R. Tan, H. Nguyen, E. Foo, X. Dong, D. Yau, Z. Kalbarczyk, R. Iyer, and H. Gooi, "Optimal false data injection attack against automatic generation control in power grids," in *ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*, 2016.
[25] K. Rahimi, A. Parchure, V. Centeno, and R. Broadwater, "Effect of communication time-delay attacks on the performance of automatic generation control," in *North American Power Symposium (NAPS)*, 2015.
[26] M. Roozbehani, M. Dahleh, and S. Mitter, "Volatility of power grids under real-time pricing," *IEEE Trans. Power Syst.*, vol. 27, no. 4, 2012.
[27] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. V. Cutsem, and V. Vittal, "Definition and classification of power system stability," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, 2004.
[28] M. Malek-Zavarei and M. Jamshidi, *Time-Delay Systems: Analysis, Optimization and Applications*. Elsevier Science, 1987.
[29] J. D. Glover, M. S. Sarma, and T. J. Overbye, *Power System Analysis and Design*, 5th ed. Cengage Learning, 2011.
[30] X. Lou, D. Yau, H. Nguyen, and B. Chen, "Profit-optimal and stability-aware load curtailment in smart grids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1411–1420, 2013.
[31] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[32] *ThermoPower*, https://casella.github.io/ThermoPower/.
[33] *Modelica*, https://www.modelica.org/.
[34] *OpenModelica*, https://www.openmodelica.org/.