Authentication of cyber-physical systems under learning-based attacks ⋆

Mohammad Javad Khojasteh * **Anatoly Khina** **
Massimo Franceschetti * **Tara Javidi** *

* *Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093, USA (e-mail: {mkhojast, mfranceschetti, tjavidi}@eng.ucsd.edu).*
** *School of Electrical Engineering, Tel Aviv University, Tel Aviv, Israel 6997801 (e-mail: anatolyk@eng.tau.ac.il)*

Abstract: We study the problem of learning-based attacks in a simple abstraction of cyber-physical systems—the case of a scalar, discrete-time, linear, time-invariant plant that may be subject to an attack that overrides the sensor readings and the controller actions. The attacker attempts to learn the dynamics of the plant and subsequently override the controller's actuation signal, to destroy the plant without being detected. The attacker can feed fictitious sensor readings to the controller using its estimate of the plant dynamics and mimicking the legitimate plant operation. The controller, on the other hand, is constantly on the lookout for an attack, and immediately shuts the plant off if an attack is detected. We study the performance of a specific authentication test and, by utilizing tools from information theory and statistics, we bound the asymptotic detection and deception probabilities for *any measurable* control policy when the attacker uses *an arbitrary* learning algorithm to estimate the dynamic of the plant. Finally, we show how the controller can impede the learning process of the attacker by superimposing a carefully crafted *privacy-enhancing signal* upon its control policy.

Keywords: Secure control, system identification, cyber-physical systems security, man-in-the-middle attack, physical authentication of control systems.

1. INTRODUCTION

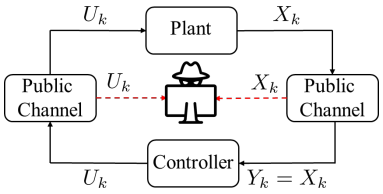
Recent technological advances in wireless communications and computation, and their integration into networked control and cyber-physical systems (CPS), open the door to a myriad of exciting opportunities in cloud robotics (Kehoe et al., 2015).

However, the distributed nature of CPS is often a source of vulnerability. Security breaches in CPS can have catastrophic consequences ranging from hampering the economy by obtaining financial gain, through hijacking autonomous vehicles and drones, and all the way to terrorism by manipulating life-critical infrastructures (Urbina et al., 2016). Real-world instances of security breaches in CPS, that were discovered and made available to the public, include the revenge sewage attack in Maroochy Shire, Australia; the Ukraine power grid cyber-attack; the German steel mill cyber-attack; the Davis-Besse nuclear power plant attack in Ohio, USA; and the Iranian uranium-enrichment facility attack via the Stuxnet malware (Sandberg et al., 2015). Consequently, studying and preventing such security breaches via control-theoretic methods have received a great deal of attention in recent years (Bai et al., 2017; Dolk et al., 2017; Shoukry et al., 2018; Chen et al., 2018; Shi et al., 2018; Dibaji et al., 2018; Tunga et al., 2018).

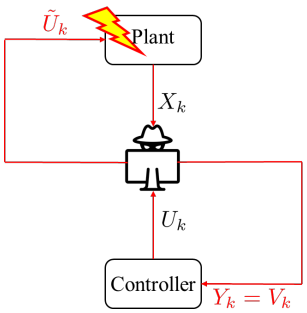
An important and widely used class of attacks in CPS is based on the "man-in-the-middle" (MITM) attack technique (Smith, 2011): an attacker takes over the control and sensor signals of the physical plant. The attacker overrides the control signals with malicious inputs to push the plant toward a catastrophic trajectory. Consequently, many CPS constantly monitor the plant outputs to detect possible attacks. The attacker, on the other hand, aims to override the sensor readings in a manner that would be indistinguishable from the legitimate ones.

The MITM attack has been extensively studied in two special cases (Mo et al., 2015; Zhu and Martínez, 2014; Miao et al., 2013; Satchidanandan and Kumar, 2017; Smith, 2011). The first case is the *replay attack*, in which the attacker observes and records the legitimate system behavior for a given time window and then replays this recording periodically at the controller's input (Mo et al., 2015; Zhu and Martínez, 2014; Miao et al., 2013). The second case is the *statistical-duplicate attack*, which assumes that the attacker has acquired complete knowledge of the dynamics and parameters of the system, and can construct arbitrarily long fictitious sensor readings that are statistically identical to the actual signals (Satchidanandan and Kumar, 2017; Smith, 2011). The replay attack assumes no knowledge of the system parameters—and as a consequence, it is relatively easy to detect it. An effective way to counter the replay attack consists of superimposing a random watermark signal, unknown to the attacker, on top of the control signal (Fang et al., 2017; Hespanhol et al., 2018). The statistical-duplicate attack assumes full knowledge of the system dynamics—and as a consequence, it requires a more sophisticated detection procedure, as well as additional assumptions on the attacker or controller behavior to ensure it can be detected. To combat

⋆ This research was partially supported by NSF award CNS-1446891. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 708932.



(a) Learning: During this phase, the attacker eavesdrops and learns the system, without altering the input signal to the controller ($Y_k = X_k$).



(b) Hijacking: During this phase, the attacker hijacks the system and intervenes as a MITM in two places: acting as a fake plant for the controller ($Y_k = V_k$) by impersonating the legitimate sensor, and as a malicious controller ($\tilde{U}_k$) for the plant aiming to destroy the plant.

Fig. 1. System model during learning-based attack phases.

the attacker's full knowledge, the controller may adopt *moving target* (Weerakkody and Sinopoli, 2015) or *baiting* (Flamholz et al., 2019; Hoehn and Zhang, 2016) techniques. Alternatively, the controller may introduce private randomness in the control input using *watermarking* (Satchidanandan and Kumar, 2017). In this scenario, a vital assumption is made: although the attacker observes the true sensor readings, it is barred from observing the control actions, as otherwise it would be omniscient and undetectable.

Our contribution is threefold. First, we observe that in many practical situations, the attacker does not have full knowledge of the system and cannot simulate a statistically indistinguishable copy of the system. On the other hand, the attacker can carry out more sophisticated attacks than simply replaying the previous sensor readings, by attempting to "learn" the system dynamics from the observations. For this reason, we study *learning-based attacks* and show that they can outperform replay attacks by analyzing the performance using a specific learning algorithm. Second, we derive asymptotic bounds on the detection and deception probabilities for *any* (measurable) control policy when the attacker uses *any arbitrary* learning algorithm to estimate the dynamics of the plant. Third, for any learning algorithm utilized by the attacker to estimate the dynamics of the plant, we show that adding a proper *privacy-enhancing signal* to any measurable control policy provides enhanced guarantees on the detection probability.

Throughout the paper, we assume that the attacker has *full access to both sensor and control signals*. The controller, on the other hand, has perfect knowledge of the system dynamics and tries to discover the attack from the injected observations. The assumed information-pattern imbalance between the controller and the attacker is justified since the controller is tuned in much longer than the attacker and thus has knowledge of the system dynamics to a far greater precision than the attacker. While our analysis is restricted to linear scalar systems, it is natural to extend this framework and develop parallel results for multivariate and nonlinear systems. Finally, studying the relation between our privacy-enhancing signal with the noise utilized to achieve differential privacy (Cortés et al., 2016) is also an interesting future research venue.

A complete list of notations and proofs of all the results appear in (Khojasteh et al., 2018), due to lack of space.

2. PROBLEM SETUP

We consider the networked control system depicted in Fig. 1, where the plant dynamics are described by a scalar, discrete-time, linear time-invariant (LTI) system

$$X_{k+1} = aX_k + U_k + W_k, \tag{1}$$

where X_k, a, U_k, W_k are real numbers representing the plant state, open-loop gain of the plant, control input, and plant disturbance, respectively, at time $k \in \mathbb{N}$. The controller, at time k, observes Y_k and generates a control signal U_k as a function of Y_1^k. We assume that the initial condition X_0 has a known (to all parties) distribution and is independent of the disturbance sequence $\{W_k\}$. For analytical purposes, we assume $\{W_k\}$ is an i.i.d. Gaussian process $\mathcal{N}(0, \sigma^2)$ known to all parties. We assume that $U_0 = W_0 = 0$. Moreover, to simplify the notation, let $Z_k \triangleq (X_k, U_k)$ denote the state-and-control input at time k and its trajectory up to time k—by

$$Z_1^k \triangleq (X_1^k, U_1^k).$$

The controller is equipped with a detector that tests for anomalies in the observed history Y_1^k. When the controller detects an attack, it shuts the system down and prevents the attacker from causing further "damage" to the plant. The controller/detector is aware of the plant dynamics (1) and knows the open-loop gain a of the plant. On the other hand, the attacker knows the plant dynamics (1) as well as the plant state X_k, and control input U_k (or equivalently, Z_k) at time k (see Fig. 1). However, it does not know the open-loop gain a of the plant.

In what follows, it will be convenient to treat the open-loop gain of the plant as a random variable A that is *fixed in time*, whose PDF f_A is known to the attacker, and whose realization a is known to the controller. We assume all random variables to exist on a common probability space with probability measure $\mathbb{P}$, and U_k to be a *measurable function* of Y_1^k for all time $k \in \mathbb{N}$. We also denote the probability measure conditioned on $A = a$ by $\mathbb{P}_a$. Namely, for any measurable event C, we define

$$\mathbb{P}_a(C) = \mathbb{P}(C|A = a).$$

A is further assumed to be independent of X_0 and $\{W_k | k \in \mathbb{N}\}$.

2.1 Learning-based attacks

We define *Learning-based attacks* that consist of two disjoint, consecutive, passive and active phases, as follows.

Phase 1: Learning. During this phase, the attacker passively observes the control input and the plant state to learn the open-loop gain of the plant. As illustrated in Fig. 1a, for all $k \in [0, L]$, the attacker observes the control input U_k and the plant state X_k, and tries to learn the open-loop gain a, where L is the duration of the learning phase. We denote by $\hat{A}$ the attacker's estimate of the open-loop gain a. •

Phase 2: Hijacking. In this phase, the attacker aims to destroy the plant using $\tilde{U}_k$ while remaining undetected. As illustrated in Fig. 1b, from time $L+1$ and onwards the attacker hijacks the

system and feeds a malicious control signal to the plant $\tilde{U}_k$ and a fictitious sensor reading $Y_k = V_k$ to the controller. •

We assume that the attacker can use *any arbitrary* learning algorithm to estimate the open-loop gain a during the learning phase, and upon estimation is completed, we assume that during the hijacking phase the fictitious sensor reading is constructed in the following way

$$V_{k+1} = \hat{A}V_k + U_k + \tilde{W}_k\,, \qquad k = L, \dots, T-1, \quad (2)$$

where $\tilde{W}_k$ for $k = L, \dots, T-1$ are i.i.d. Gaussian $\mathcal{N}(0, \sigma^2)$; U_k is the control signal generated by the controller, which is fed with the fictitious virtual signal V_k by the attacker; $V_L = X_L$; and $\hat{A}$ is the estimate of the open-loop gain of the plant at the conclusion of Phase 1.

2.2 Detection

The controller/detector, being aware of the dynamic (1) and the open-loop gain a, attempts to detect possible attacks by testing for statistical deviations from the typical behavior of the system (1). More precisely, under legitimate system operation (corresponding to the *null hypothesis*), the controller observation Y_k behaves according to

$$Y_{k+1} - aY_k - U_k(Y_1^k) \sim \text{ i.i.d. } \mathcal{N}(0, \sigma^2). \quad (3)$$

In case of an attack, during Phase 2 ($k > L$), (3) can be rewritten as

$$\begin{aligned} &V_{k+1} - aV_k - U_k(Y_1^k) \\ &\quad = V_{k+1} - aV_k + \hat{A}V_k - \hat{A}V_k - U_k(Y_1^k) \quad (4a) \\ &\quad = \tilde{W}_k + \left(\hat{A} - a\right) V_k, \quad (4b) \end{aligned}$$

where (4b) follows from (2). Hence, the estimation error $(\hat{A}-a)$ dictates the ease with which an attack can be detected.

Since the Gaussian PDF with zero mean is fully characterized by its variance, we shall test for anomalies in the latter, i.e., test whether the empirical variance of (3) is equal to the second moment of the plant disturbance $\mathbb{E}\left[W^2\right]$. To that end, we shall use a test that sets a confidence interval of length $2\delta > 0$ around the expected variance, i.e., it checks whether

$$\begin{aligned} &\frac{1}{T}\sum_{k=1}^{T}\left[Y_{k+1} - aY_k - U_k(Y_1^k)\right]^2 \\ &\qquad \in (\mathrm{Var}\left[W\right] - \delta, \mathrm{Var}\left[W\right] + \delta), \end{aligned} \quad (5)$$

where T is called the *test time*. That is, as is implied by (4), the attacker manages to deceive the controller and remain undetected if

$$\begin{aligned} &\tfrac{1}{T}\left(\textstyle\sum_{k=1}^{L} W_k^2 + \sum_{k=L+1}^{T}(\tilde{W}_k + (\hat{A}-a)V_k)^2\right) \\ &\qquad \in (\mathrm{Var}\left[W\right] - \delta, \mathrm{Var}\left[W\right] + \delta). \end{aligned}$$

2.3 Performance Measures

Definition 1. The hijack indicator at test time T is defined as

$$\Theta_T \triangleq \begin{cases} 0, & \forall j \le T:\ Y_j = X_j\,; \\ 1, & \text{otherwise.} \end{cases}$$

At the test time T, the controller uses Y_1^T to construct an estimate $\hat{\Theta}_T$ of Θ_T. More precisely, $\hat{\Theta}_T = 0$ if (5) occurs, otherwise $\hat{\Theta}_T = 1$. •

Definition 2. The probability of deception is the probability of the attacker deceiving the controller and remain undetected at the time instance T

$$P_{\mathrm{Dec}}^{a,T} \triangleq \mathbb{P}_a\left(\hat{\Theta}_T = 0 \middle| \Theta_T = 1\right). \quad (6)$$

In addition, the detection probability at test time T is defined as

$$P_{\mathrm{Det}}^{a,T} \triangleq 1 - P_{\mathrm{Dec}}^{a,T}.$$

Likewise, the probability of false alarm is the probability of detecting the attacker when it is not present, namely

$$P_{\mathrm{FA}}^{a,T} \triangleq \mathbb{P}_a\left(\hat{\Theta}_T = 1 \middle| \Theta_T = 0\right). \quad \bullet$$

In this case, using Chebyshev's inequality, (5), since the system disturbances are i.i.d. Gaussian $\mathcal{N}(0, \sigma^2)$, we have

$$P_{\mathrm{FA}}^{T} \le \frac{\mathrm{Var}[W^2]}{\delta^2 T} = \frac{3\sigma^4}{\delta^2 T}.$$

We further define the deception, detection, and false alarm probabilities w.r.t. the probability measure $\mathbb{P}$, without conditioning on A, and denote them by P_{Dec}^{T}, P_{Det}^{T}, and P_{FA}^{T}, respectively. For instance, P_{Det}^{T} is defined, w.r.t. a PDF f_A of A, as

$$P_{\mathrm{Det}}^{T} \triangleq \mathbb{P}\left(\hat{\Theta}_T = 1 \middle| \Theta_T = 1\right) = \textstyle\int_{-\infty}^{\infty} P_{\mathrm{Det}}^{a,T}\, f_A(a)da \quad (7)$$

3. STATEMENT OF THE RESULTS

In this section, we describe the main results of this work. We want to provide lower and upper bounds on the deception probability (6) of the learning-based attack (2) where $\hat{A}$ in (2) is constructed using *any arbitrary* learning algorithm. In addition, our results are valid for *any measurable* control policy U_k. We find a lower bound on the deception probability by characterizing what attacker can at least achieve using a least-squares (LS) algorithm, and we derive an information theoretic upper bound using Fano's inequality (Polyanskiy and Wu, 2016). While our analysis is restricted to the asymptotic case, $T \to \infty$, it is straightforward to extend this treatment to the non-asymptotic case.

For analytical purposes, we assume that the power of the fictitious sensor reading is equal to $\beta^{-1} < \infty$, namely

$$\lim_{T\to\infty} \frac{1}{T}\sum_{k=L+1}^{T} V_k^2 = 1/\beta \qquad \text{a.s. w.r.t. } \mathbb{P}_a. \quad (8)$$

Remark 3. Assuming the control policy is memoryless, namely U_k is only dependent on Y_k, the process V_k is Markov for $k \ge L+1$. By further assuming that $L = o(T)$ and using the generalization of the law of large numbers for Markov processes (Durrett, 2010), we deduce

$$\lim_{T\to\infty} \frac{1}{T}\sum_{k=L+1}^{T} V_k^2 \ge \mathrm{Var}\left[W\right] \qquad \text{a.s. w.r.t. } \mathbb{P}_a.$$

Consequently, in this case we have $\beta \le 1/\mathrm{Var}\left[W\right]$. In addition, when the control policy is linear and stabilizes (2), that is $U_k = -\Omega Y_k$ and $|\hat{A} - \Omega| < 1$, it is easy to verify that (8) holds true for $\beta = (1 - (\hat{A} - \Omega)^2)/\mathrm{Var}\left[W\right]$. •

3.1 Lower Bound on the Deception Probability

To provide a lower bound on the deception probability $P_{\mathrm{Dec}}^{a,T}$, we consider a specific estimate of $\hat{A}$ at the conclusion of the first phase by the attacker. To this end, we use LS estimation

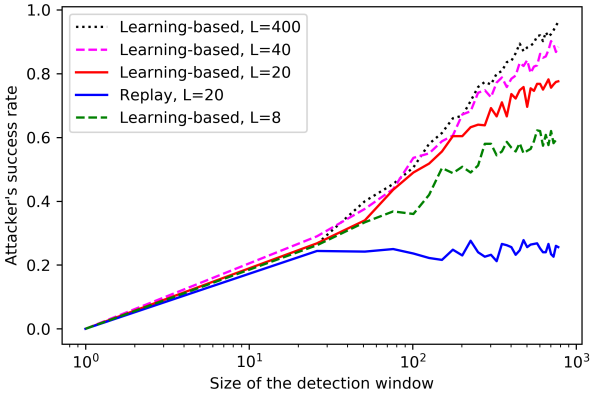


Fig. 2. The attacker's success rate $P_{\text{Dec}}^{a,T}$ versus the size of the detection window T.

due to its efficiency and amenability to recursive update over observed incremental data (Rantzer, 2018; Tu and Recht, 2018; Sarkar and Rakhlin, 2019). The LS algorithm approximates the overdetermined system of equations

$$\begin{pmatrix} X_2 \\ X_3 \\ \vdots \\ X_L \end{pmatrix} = A \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_{L-1} \end{pmatrix} + \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_{L-1} \end{pmatrix},$$

by minimizing the Euclidean distance $\hat{A} = \operatorname{argmin}_A \|X_{k+1} - AX_k - U_k\|$ to estimate (or "identify") the plant, the solution to which is

$$\hat{A} = \frac{\sum_{k=1}^{L-1}(X_{k+1} - U_k)X_k}{\sum_{k=1}^{L-1} X_k^2} \qquad \text{a.s. w.r.t. } \mathbb{P}_a. \tag{9}$$

Remark 4. Since we assumed $W_k \sim \mathcal{N}(0, \sigma^2)$ for all $k \in \mathbb{N}$, $\mathbb{P}_a(X_k = 0) = 0$. Thus, (9) is well-defined. •

Using LS estimation (9), our linear learning-based attack (2) achieves *at least* the asymptotic deception probability stated in the following theorem, for *any measurable* control policy.

Theorem 5. Consider any linear learning-based attack (2) with fictitious sensor reading power that satisfies (8) and an *arbitrary* measurable control policy $\{U_k\}$. Then, the asymptotic deception probability, when using the variance test (5), is bounded from below as

$$\lim_{T\to\infty} P_{\text{Dec}}^{a,T} = \mathbb{P}_a\left(|\hat{A} - a| < \sqrt{\delta\beta}\right) \tag{10a}$$

$$\geq \mathbb{P}_a\left(\frac{\left|\sum_{k=1}^{L-1} W_k X_k\right|}{\sum_{k=1}^{L-1} X_k^2} < \sqrt{\delta\beta}\right) \tag{10b}$$

$$\geq 1 - \frac{2}{(1+\delta\beta)^{L/2}}. \tag{10c}$$

Example 6. In this example, we compare the empirical performance of the variance-test with our developed bound in Thm. 5. At every time T, the controller tests the empirical variance for abnormalities over a detection window $[1, T]$, using a confidence interval $2\delta > 0$ around the expected variance (5). Here, $a = 1$, $\delta = 0.1$, $U_k = -0.88aY_k$ for all $1 \leq k \leq 800$, and $\{W_k\}$ are i.i.d. Gaussian $\mathcal{N}(0, 1)$, and 500 Monte Carlo simulations were performed.

The learning-based attacker (2) uses the LS algorithm (9) to estimate a, and as illustrated in Fig. 2, the attacker's success rate increases as the duration of learning phase L increases. This is in agreement with (10c) since the attacker can improve its estimate of a and the estimation error $|\hat{A} - a|$ reduces as L increases. As discussed in Sec. 2.3, the false alarm rate decays to zero as the size of the detection window T tends to infinity. Hence, for a sufficiently large detection window size, the attacker's success rate could potentially tend to one. Indeed, such behavior is observed in Fig. 2 for a learning-based attacker (2) with $L = 400$.

Also, Fig. 2 illustrates that our learning-based attack outperforms the replay attack. A replay attack with a recording length of $L = 20$ and a learning-based attack with a learning phase of length $L = 20$ are compared, and the success rate of the replay attack saturates at a lower value. Moreover, a learning-based attack with a learning phase of length $L = 8$ has a higher success rate than a replay attack with a larger recording length of $L = 20$. •

3.2 Upper Bound on the Deception Probability

We now derive an upper bound on the deception probability (6) of any learning-based attack (2) where $\hat{A}$ in (2) is constructed using *any arbitrary* learning algorithm, for *any measurable* control policy, when A is distributed over a symmetric interval $[-R, R]$. Similar results can be obtained for other interval choices. Since the uniform distribution has the highest entropy among all distributions with finite support (Polyanskiy and Wu, 2016), we further assume A is distributed uniformly over the interval $[-R, R]$. We assume the attacker knows the distribution of A (including the value of R), whereas the controller knows the true value of A (as before).

Theorem 7. Let A be distributed uniformly over $[-R, R]$ for some $R > 0$, and consider *any measurable* control policy $\{U_k\}$ and *any learning-based attack* (2) with fictitious sensor reading power (8) that satisfies $\sqrt{\delta\beta} \leq R$. Then, the asymptotic deception probability, when using the variance test (5), is bounded from above as

$$\lim_{T\to\infty} P_{\text{Dec}}^{T} = \mathbb{P}(|A - \hat{A}| < \sqrt{\delta\beta}) \tag{11a}$$

$$\leq \Lambda \triangleq \frac{I(A; Z_1^L) + 1}{\log(R/\sqrt{\delta\beta})}. \tag{11b}$$

In addition, if for all $k \in \{1, \ldots, L\}$, $A \to (X_k, Z_1^{k-1}) \to U_k$ is a Markov chain, then for any sequence of probability measures $\{\mathbb{Q}_{X_k|Z_1^{k-1}}\}$, such that for all $k \in \{1, \ldots, L\}$ $\mathbb{P}_{X_k|Z_1^{k-1}} \ll \mathbb{Q}_{X_k|Z_1^{k-1}}$, we have

$$\Lambda \leq \frac{\sum_{k=1}^{L} D\left(\mathbb{P}_{X_k|Z_1^{k-1},A} \middle\| \mathbb{Q}_{X_k|Z_1^{k-1}} \middle| \mathbb{P}_{Z_1^{k-1},A}\right) + 1}{\log\left(R/\sqrt{\delta\beta}\right)}. \tag{12}$$

Remark 8. By looking at the numerator in (11b), it follows that the bound on the deception probability becomes looser as the amount of information revealed about the open-loop gain A by the observation Z_1^L increases. On the other hand, by looking at the denominator, the bound becomes tighter as R increases. This is consistent with the observation of Zames (Raginsky, 2010) that system identification becomes harder as the uncertainty about the open-loop gain of the plant increases. In our case, a larger uncertainty interval R corresponds to a poorer estimation of A by the attacker, which leads, in turn, to a decrease in the achievable deception probability. The denominator can also, be interpreted as the intrinsic uncertainty of A when it is observed at resolution $\sqrt{\delta\beta}$, as it corresponds to the entropy of the random variable A when it is quantized at such resolution.•

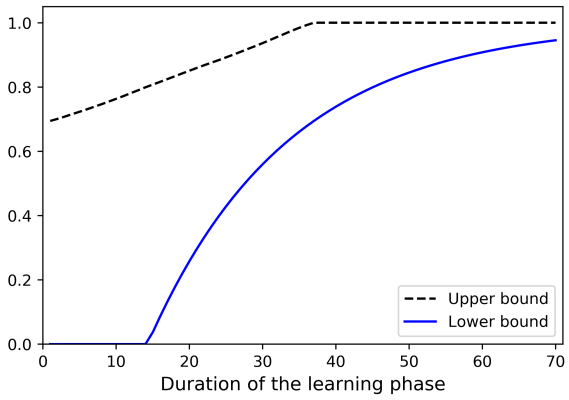


Fig. 3. Comparison of the lower and upper bounds on the deception probability, of Thm. 5 and Corol. 9, respectively.

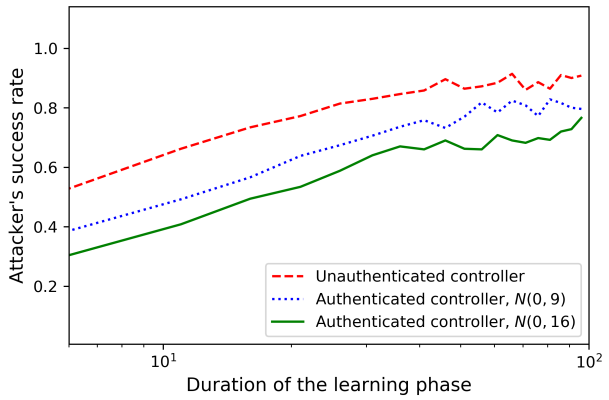


Fig. 4. The attacker's success rate $P_{\text{Dec}}^{a,T}$ versus the duration of the exploration phase L.

In conclusion, Thm. 7 provides two upper bounds on the deception probability. The first bound (11b) clearly shows that increasing the privacy of the open-loop gain A—manifested in the mutual information between A and the state-and-control trajectory Z_1^L during the exploration phase—reduces the deception probability. The second bound (12) allows freedom in choosing the auxiliary probability measure $\mathbb{Q}_{X_k|Z_1^{k-1}}$, making it a rather useful bound. For instance, by choosing $\mathbb{Q}_{X_k|Z_1^{k-1}} \sim \mathcal{N}(0,\sigma^2)$, for all $k \in \mathbb{N}$, we can rewrite the upper bound (12) in term of $\mathbb{E}_{\mathbb{P}}\left[(AX_{k-1}+U_{k-1})^2\right]$ as follows.

Corollary 9. Under the assumptions of Thm. 7, if for all $k \in \{1,\dots,L\}$, $A \to (X_k, Z_1^{k-1}) \to U_k$ is a Markov chain, then asymptotic deception probability is bounded from above by

$$\lim_{T\to\infty} P_{\text{Dec}}^T \le G(Z_1^L), \tag{13a}$$

$$G(Z_1^L) \triangleq \frac{\frac{\log e}{2\sigma^2}\sum_{k=1}^{L}\mathbb{E}_{\mathbb{P}}\left[(AX_{k-1}+U_{k-1})^2\right]+1}{\log\left(R/\sqrt{\delta\beta}\right)}. \tag{13b}$$

Example 10. Thm. 5 provides a lower bound on the deception probability given $A = a$. Hence, by applying the law of total probability w.r.t. the PDF f_A of A as in (7), we can apply the result of Thm. 5 to provide a lower bound also on the average deception probability for a random open-loop gain A. In this context, Fig. 3 compares the lower and upper bounds on the deception probability provided by Thm. 5, $\max\{0, 1-(2/(1+\delta\beta)^{L/2})\}$, and Corol. 9, $\min\{1, G(Z_1^L)\}$, respectively, where A is distributed uniformly over $[-0.9, 0.9]$. (13a) is valid when the control input is not a function of random variable A; hence, we assumed $U_k = -0.045Y_k$ for all time $k \in \mathbb{N}$. Here $\delta = 0.1$, $\{W_k\}$ are i.i.d. Gaussian with zero mean and variance of 0.16, and for simplicity, we let $\beta = 1.1$. Although, in general, the attacker's estimation of the random open-loop gain A and consequently the power of fictitious senor reading (8) vary based on the learning algorithm and the realization of A, the comparison of the lower and upper bounds in Fig. 3 is restricted to a fixed β. 2000 Monte Carlo simulations were performed.

3.3 Privacy-enhancing signal

For a given duration of learning phase L, to increase the security of the system, at any time k the controller can add a privacy-enhancing signal Γ_k to an unauthenticated control policy $\{\bar{U}_k | k \in \mathbb{N}\}$:

$$U_k = \bar{U}_k + \Gamma_k\,, \qquad k \in \mathbb{N}. \tag{14}$$

We refer to such a control policy U_k as the *authenticated* control policy $\bar{U}_k$. We denote the states of the system that would be generated if only the unauthenticated control signal $\bar{U}_1^k$ were applied by $\bar{X}_1^k$, and the resulting trajectory—by $\bar{Z}_1^k \triangleq (\bar{X}_1^k, \bar{U}_1^k)$.

The following numerical example illustrates the effect of the privacy-enhancing signal on the deception probability.

Example 11. Here, the attacker uses the LS algorithm (9), the detector uses the variance test (5), $a = 1$, $T = 600$, $\delta = 0.1$, and $\{W_k\}$ are i.i.d. Gaussian $\mathcal{N}(0,1)$. Fig. 4 compares the attacker's success rate, the empirical $P_{\text{Dec}}^{a,T}$, as a function of the duration L of the learning phase for three different control policies: I) unauthenticated control signal $\bar{U}_1^k = -aY_k$ for all k, II) authenticated control signal (14), where Γ_k are i.i.d. Gaussian $\mathcal{N}(0,9)$, III) authenticated control signal (14), where Γ_k are i.i.d. Gaussian $\mathcal{N}(0,16)$. As illustrated in Fig. 4, for the authenticated and unauthenticated control signals, the attacker's success rate increases as the duration of the learning phase increases. This is in agreement with (10c) since the attacker can improve its estimate of a as L increases. Also, for a fix L the attacker performance deteriorates as the power of privacy-enhancing signal Γ_k increases. Namely, Γ_k hampers the learning process of the attacker and the estimation error $|\hat{A} - a|$ increases as the power of privacy-enhancing signal increases. 500 Monte Carlo simulations were performed. •

Remark 12. A "good" privacy-enhancing signal entails little increase in the control cost (Bertsekas, 1995) compared to its unauthenticated version while providing enhanced detection probability (6) and/or false alarm probability. Finding the optimal privacy-enhancing signal is an interesting research venue.•

One may envisage that superimposing any noisy signal Γ_k on top of the control policy $\{\bar{U}_k | k \in \mathbb{N}\}$ would necessarily enhance the detectability of *any* learning-based attack (2) since the observations of the attacker are in this case noisier. However, it turns out that injecting a strong noise for some learning algorithm may speed up the learning process as it improves the power of the signal magnified by the open-loop gains with respect to the observed noise. Any signal Γ_k that satisfies the condition proposed in the following corollary will provide enhanced guarantees on the detection probability when the attacker uses *any arbitrary* learning algorithm to estimate the uniformly distributed A over the symmetric interval $[-R, R]$.

Corollary 13. For any control policy $\{\bar{U}_k | k \in \mathbb{N}\}$ with trajectory $\bar{Z}_1^k = (\bar{X}_1^k, \bar{U}_1^k)$ and its corresponding authenticated

control policy U_1^k (14) with trajectory $Z_1^k = (X_1^k, U_1^k)$, under the assumptions of Corollary 9, if for all $k \in \{2, \dots, L\}$

$$\mathbb{E}_{\mathbb{P}}\left[\Psi_{k-1}^2 + 2\Psi_{k-1}(A\bar{X}_{k-1} + \bar{U}_{k-1})\right] < 0, \qquad (15)$$

where $\Psi_{k-1} \triangleq \sum_{j=1}^{k-1} A^{k-1-j}\Gamma_j$, for any $L \geq 2$, the following majorization of G (13b) holds:

$$G(Z_1^L) < G(\bar{Z}_1^L). \qquad (16)$$

Example 14. In this example, we describe a class of privacy-enhancing signal that yield better guarantees on the deception probability. For all $k \in \{2, \dots, L\}$, clearly $\Psi_{k-1} = -(AX_{k-1} + U_{k-1})/\eta$ satisfies the condition in (15) for any $\eta \in \{2, \dots, L\}$. Thus, by choosing the privacy-enhancing signals $\Gamma_1 = -(AX_1 + U_1)/\eta$, and $\Gamma_k = -(AX_k + U_k)/\eta - \sum_{j=1}^{k-2} A^{k-1-j}\Gamma_j$ for all $k \in \{3, \dots, L\}$, (16) holds. •

4. FUTURE WORK

Future work will explore the extension of the established results to the vector (possibly partially observable) case, designing optimal privacy-enhancing signals, and investigating the more realistic scenario where neither the attacker nor the controller are aware of the open-loop gain of the plant.

REFERENCES

Bai, C.Z., Pasqualetti, F., and Gupta, V. (2017). Data-injection attacks in stochastic control systems: Detectability and performance tradeoffs. *Automatica*, 82, 251–260.

Bertsekas, D.P. (1995). *Dynamic programming and optimal control*, volume 1. Athena scientific Belmont, MA.

Chen, Y., Kar, S., and Moura, J.M. (2018). Cyber-physical attacks with control objectives. *IEEE Tran. Auto. Cont.*, 63(5), 1418–1425.

Cortés, J., Dullerud, G.E., Han, S., Le Ny, J., Mitra, S., and Pappas, G.J. (2016). Differential privacy in control and network systems. In *2016 IEEE Conf. Decision and Cont. (CDC)*, 4252–4272.

Dibaji, S., Pirani, M., Annaswamy, A., Johansson, K.H., and Chakrabortty, A. (2018). Secure control of wide-area power systems: Confidentiality and integrity threats. In *2018 IEEE Conf. Decision and Cont. (CDC)*, 7269–7274.

Dolk, V., Tesi, P., De Persis, C., and Heemels, W. (2017). Event-triggered control systems under denial-of-service attacks. *IEEE Tran. Cont. Net. Sys.*, 4(1), 93–105.

Duchi, J.C. and Wainwright, M.J. (2013). Distance-based and continuum fano inequalities with applications to statistical estimation. *arXiv preprint arXiv:1311.2669*.

Durrett, R. (2010). *Probability: theory and examples*. Cambridge university press.

Fang, C., Qi, Y., Cheng, P., and Zheng, W.X. (2017). Cost-effective watermark based detector for replay attacks on cyber-physical systems. In *2017 Asian Cont. Conf. (ASCC)*, 940–945. IEEE.

Flamholz, D.B., Annaswamy, A.M., and Lavretsky, E. (2019). Baiting for defense against stealthy attacks on cyber-physical systems. In *AIAA Scitech 2019 Forum*, 2338.

Hespanhol, P., Porter, M., Vasudevan, R., and Aswani, A. (2018). Statistical watermarking for networked control systems. In *2018 American Cont. Conf. (ACC)*, 5467–5472. IEEE.

Hoehn, A. and Zhang, P. (2016). Detection of covert attacks and zero dynamics attacks in cyber-physical systems. In *2016 American Cont. Conf. (ACC)*, 302–307. IEEE.

Kehoe, B., Patil, S., Abbeel, P., and Goldberg, K. (2015). A survey of research on cloud robotics and automation. *IEEE Tran. on auto. science and eng.*, 12(2), 398–409.

Khojasteh, M.J., Khina, A., Franceschetti, M., and Javidi, T. (2018). Authentication of cyber-physical systems under learning-based attacks. *arXiv preprint arXiv:1809.06023*.

Lai, T.L. and Wei, C.Z. (1982). Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *The Annals of Statistics*, 154–166.

Miao, F., Pajic, M., and Pappas, G.J. (2013). Stochastic game approach for replay attack detection. In *2013 IEEE Conf. Decision and Cont.*, 1854–1859.

Mo, Y., Weerakkody, S., and Sinopoli, B. (2015). Physical authentication of control systems: designing watermarked control inputs to detect counterfeit sensor outputs. *IEEE Control Systems*, 35(1), 93–109.

Polyanskiy, Y. and Wu, Y. (2016). Lecture notes on information theory.

Raginsky, M. (2010). Divergence-based characterization of fundamental limitations of adaptive dynamical systems. In *2010 Allerton Conference on Communication, Control, and Computing*, 107–114. IEEE.

Rantzer, A. (2018). Concentration bounds for single parameter adaptive control. In *2018 American Cont. Conf. (ACC)*, 1862–1866. IEEE.

Sandberg, H., Amin, S., and Johansson, K.H. (2015). Cyber-physical security in networked control systems: An introduction to the issue. *IEEE Control Systems*, 35(1), 20–23.

Sarkar, T. and Rakhlin, A. (2019). Near optimal finite time identification of arbitrary linear dynamical systems. In *Int. Conf. on Machine Learning (ICML)*, 5610–5618.

Satchidanandan, B. and Kumar, P.R. (2017). Dynamic watermarking: Active defense of networked cyber–physical systems. *Proceedings of the IEEE*, 105(2), 219–240.

Shi, D., Guo, Z., Johansson, K.H., and Shi, L. (2018). Causality countermeasures for anomaly detection in cyber-physical systems. *IEEE Tran. Auto. Cont.*, 63(2), 386–401.

Shoukry, Y., Chong, M., Wakaiki, M., Nuzzo, P., Sangiovanni-Vincentelli, A., Seshia, S.A., Hespanha, J.P., and Tabuada, P. (2018). SMT-based observer design for cyber-physical systems under sensor attacks. *ACM Trans. on CPS*, 2(1), 5.

Smith, R.S. (2011). A decoupled feedback structure for covertly appropriating networked control systems. *IFAC Proceedings Volumes*, 44(1), 90–95.

Tu, S. and Recht, B. (2018). Least-squares temporal difference learning for the linear quadratic regulator. In *Int. Conf. on Machine Learning (ICML)*, 5005–5014.

Tunga, R., Murguia, C., and Ruths, J. (2018). Tuning windowed chi-squared detectors for sensor attacks. In *2018 Annual American Cont. Conf.*, 1752–1757. IEEE.

Urbina, D.I., Giraldo, J.A., Cardenas, A.A., Tippenhauer, N.O., Valente, J., Faisal, M., Ruths, J., Candell, R., and Sandberg, H. (2016). Limiting the impact of stealthy attacks on industrial control systems. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 1092–1105.

Weerakkody, S. and Sinopoli, B. (2015). Detecting integrity attacks on control systems using a moving target approach. In *2015 IEEE Conf. Decision and Cont. (CDC)*, 5820–5826.

Zhu, M. and Martínez, S. (2014). On the performance analysis of resilient networked control systems under replay attacks. *IEEE Tran. Auto. Cont.*, 59(3), 804–808.