Sporadic Data Integrity for Secure State Estimation

Ilija Jovanov Miroslav Pajic

***Abstract*— We consider the problem of network-based attacks, such as *Man-in-the-Middle* attacks, on standard state estimators. To ensure graceful control degradation in the presence of attacks, existing results impose very strict integrity requirements on the number of noncompromised sensors. We study the effects of sporadic data integrity enforcement, such as message authentication, on control performance under stealthy attacks. We show that even with sporadic data integrity guarantees, the attacker cannot introduce an unbounded state estimation error while remaining stealthy. We present a design-time framework to derive safe integrity enforcement policies, and illustrate its use; we show that with even 20% of authenticated messages we can ensure satisfiable state estimation errors under attacks.**

I. Introduction

Increasing complexity and communication capabilities in modern control systems have also opened new alleyways for malicious interference, as was recently illustrated in Stuxnet attack [1], Ukrainian power-grid breach [2], and a few automotive attacks (e.g., [3]). Critically, network connectivity potentially allows for a *remote* attacker to compromise system components and obtain access to the low-level network used to communicate control-related messages.

From the controls perspective, these attacks can be modeled as additional malicious signals injected into the system [4]. Thus, significant research efforts have focused on development of techniques for attack detection and attack-resilient control (e.g., [5]–[11]). One avenue explored the use of unknown input observers (e.g., [8], [12]) and resilient state estimators [5]–[7]. Another has considered the use of standard residual probability-based detectors, such as χ^2 detectors, to detect attacks on sensor measurements [9]–[11]. Each of these methods also establishes prerequisites to enable attack detection. However, these conditions introduce very conservative constraints on the system, by limiting the number of sensors whose measurements can be compromised while providing suitable resiliency guarantees. For instance, for attack-resilient state estimation at least half of the sensors cannot be tampered with [5]. The reason is that the common assumption used to develop these techniques is that once a sensor or its channel to the estimator is compromised, all measurements received from the sensor could be corrupted.

In case of network-based attacks, these assumptions can be satisfied with continuous use of standard cryptographic tools, such as adding message authentication codes (MACs) to the communicated measurements in order to guarantee data integrity and authentication. This, however, imposes significant bandwidth requirements on the network; adding MAC bits can significantly increase the size of communication packets, making scheduling of several control loops over a shared network unfeasible. As an illustration, consider scheduling communication packets for two control loops over a shared network (Fig. 1). Without data authentication, communication packets can be sent within sampling periods (Fig. 1(a)). With the increase in packet sizes due to authentication of *all* sensor transmissions, there does not exist a feasible communication schedule (Fig. 1(b)). Yet, when it is not necessary to authenticate every sensor message, messages with MACs can be sent in every other period, resulting in feasible schedules for both control loops (Fig. 1(c)).

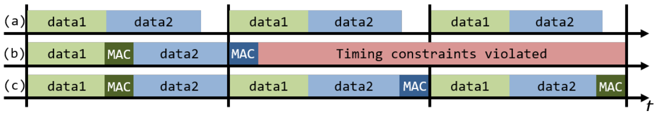
In this paper we study the effects of *sporadic* integrity guarantees for communicated sensor measurements on state estimation error in the presence of *Man-in-the-Middle* (MitM) attacks. We extend the system models from [9], [11] by considering systems in which, by the use of authentication mechanisms, it is possible to occasionally ensure that the obtained sensor measurements are valid. We show that even sporadic enforcement of sensed data integrity can significantly limit a stealthy attacker's capabilities, even if the attacker knows the times when data integrity will be enforced; specifically, we show that the attacker will be unable to introduce unbounded state estimation errors while remaining stealthy. Furthermore, we introduce a design framework to evaluate the impact of stealthy attacks on the state estimation error under a specific integrity enforcement policy, as well as design safe integrity enforcement policies.

This paper is organized as follows. In Sec. II, we introduce the considered problem and system model. Sec. III describes a method to capture attacker's impact on the system, while Sec. IV introduces sporadic integrity enforcement policies and shows that they significantly limit effects of the attacks. In Sec. V, we present a framework for design and analysis of integrity enforcement policies. Finally, Sec. VI provides an illustrative case-study, before concluding remarks in Sec VII.

Notation and Terminology: The transpose of matrix $\mathbf{A}$ is specified as $\mathbf{A}^T$, while the i^{th} element of a vector $\mathbf{x}_k$ is denoted by $\mathbf{x}_{k,i}$. Also, $\|\mathbf{A}\|_i$ denotes the i-norm of a matrix $\mathbf{A}$ and, for a positive definite matrix $\mathbf{P}$, $\|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}} = \|\mathbf{P}^{-1/2}\Delta\mathbf{z}_k\|_2$. $\mathcal{N}(\mathbf{A})$ is the null space of $\mathbf{A}$, and nullity$(\mathbf{A})$ denotes the dimension of $\mathcal{N}(\mathbf{A})$. $\mathbb{R}, \mathbb{N}$ and $\mathbb{N}_0$ denote the sets of reals, natural numbers and nonnegative integers, respectively. For a set $\mathcal{S}$, we use $|\mathcal{S}|$ to denote the cardinality (i.e., size) of the set. In addition, for a set $\mathcal{K} \subset \mathcal{S}$, with $\mathcal{K}^{\complement}$ we denote the complement set of $\mathcal{K}$ with respect to $\mathcal{S}$ – i.e.,

This work was supported in part by the NSF CNS-1652544 and CNS-1505701 grants, and the Intel-NSF Partnership for Cyber-Physical Systems Security and Privacy. This work is also based on research sponsored by the ONR under agreements number N00014-17-1-2012 and N00014-17-1-2504.

I. Jovanov and M. Pajic are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA 27708. Email: {ilija.jovanov, miroslav.pajic}@duke.edu.



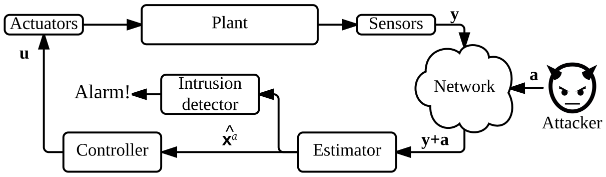


Fig. 2. System architecture – by launching *Man-in-the-Middle (MitM)* attacks, the attacker can inject adversarial signals into plant measurements delivered to the estimator.

$\mathcal{K}^{\complement} = \mathcal{S} \setminus \mathcal{K}$. Finally, *the support* of vector $\mathbf{v} \in \mathbb{R}^p$ is the set $\text{supp}(\mathbf{v}) = \{i \mid \mathbf{v}_i \neq 0\} \subseteq \{1, 2, ..., p\}$.

II. Motivation and Problem Description

Before introducing the problem formulation, we describe the considered system and attacker model, shown in Fig. 2.

A. System Model without Attacks

We consider an observable linear-time invariant (LTI) system whose evolution without attacks can be represented as

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \end{aligned} \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ and $\mathbf{u}_k \in \mathbb{R}^m$ denote the plant's state and input vectors, at time k, while the plant's output vector $\mathbf{y}_k \in \mathbb{R}^p$ contains measurements provided by p sensors from the set $\mathcal{S} = \{s_1, s_2, ..., s_p\}$. Also, $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^p$ denote the process and measurement noise; we assume that $\mathbf{x}_0$, $\mathbf{w}_k$, and $\mathbf{v}_k$ are independent Gaussian random variables.

A Kalman filter is used for state estimation. Since Kalman gain usually converges in only a few steps, to simplify the notation we assume that the system in steady state before the attack. Hence, the Kalman filter estimate $\hat{\mathbf{x}}_k$ is updated as

$$\hat{\mathbf{x}}_{k+1} = \mathbf{A}\hat{\mathbf{x}}_k + \mathbf{B}\mathbf{u}_k + \mathbf{K}(\mathbf{y}_{k+1} - \mathbf{C}(\mathbf{A}\hat{\mathbf{x}}_k + \mathbf{B}\mathbf{u}_k)), \tag{2}$$

$$\mathbf{K} = \mathbf{\Sigma}\mathbf{C}^T(\mathbf{C}\mathbf{\Sigma}\mathbf{C}^T + \mathbf{R})^{-1}, \tag{3}$$

where $\mathbf{\Sigma}$ is the estimation error covariance matrix and $\mathbf{R}$ is the sensor noise covariance matrix. Also, the residue $\mathbf{z}_k \in \mathbb{R}^p$ at time k and its covariance matrix $\mathbf{P}$ are defined as

$$\begin{aligned} \mathbf{z}_k &= \mathbf{y}_k - \mathbf{C}(\mathbf{A}\hat{\mathbf{x}}_{k-1} + \mathbf{B}\mathbf{u}_{k-1}), \\ \mathbf{P} &= E\{\mathbf{z}_k\mathbf{z}_k^T\} = \mathbf{C}\mathbf{\Sigma}\mathbf{C}^T + \mathbf{R}. \end{aligned} \tag{4}$$

Finally, the state estimation error is defined as $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$.

The system employs a χ^2 intrusion detector, defined as $g_k = {\mathbf{z}_k}^T\mathbf{P}^{-1}\mathbf{z}_k$. The alarm triggers when the value of detection function g_k satisfies that $g_k > threshold$. Therefore, the probability of the alarm at time k can be captured as

$$\beta_k = P(g_k > threshold). \tag{5}$$

B. Attack Model

We assume that the attacker is capable of launching MitM attacks on communication channels between a subset of the plant's sensors $\mathcal{K} \subseteq \mathcal{S}$ and the estimator; however, we do not assume that the set $\mathcal{K}$ is known to the system or system designers. Thus, to capture the attacker's impact on the system, the system model from (1) becomes

$$\begin{aligned} \mathbf{x}^a_{k+1} &= \mathbf{A}\mathbf{x}^a_k + \mathbf{B}\mathbf{u}^a_k + \mathbf{w}_k \\ \mathbf{y}^a_k &= \mathbf{C}\mathbf{x}^a_k + \mathbf{v}_k + \mathbf{a}_k. \end{aligned} \tag{6}$$

Here, $\mathbf{x}^a_k$ and $\mathbf{y}^a_k$ denote the state and plant outputs in the presence of attacks. In addition, $\mathbf{a}_k \in \mathbb{R}^p$ denotes the signals injected by the attacker at time k; $\mathbf{a}_k$ is a sparse vector with support in set $\mathcal{K}$ – i.e., $\mathbf{a}_{k,i} = 0$ for all $i \in \mathcal{K}^{\complement}$ and $k \geq 0$.[1]

We consider the following ***threat model*** where:

(1) The attacker has full knowledge of the system – plant dynamics, employed Kalman filter and detector, as well as the used secure-communication mechanisms. Specifically, we consider systems that utilize message authentication to ensure data integrity, and thus assume that the attacker is aware at which time points integrity will be enforced; to avoid detection, the attacker will not launch attacks in these steps and will take them into account in attack planning.

(2) The attacker has the required computation power to calculate suitable attack signals, while planning ahead as needed. In addition, he has the ability to inject *any* signal using communication packets mimicking sensors from the set $\mathcal{K}$, except at times when data integrity is enforced; for example, when MACs are utilized to ensure integrity and authenticity of sensor messages, our assumption is that the attacker does not have access to the shared secret key used to generate the MACs.

The goal of the attacker is to *maximize the error of state estimation* while ensuring that *the attack remains stealthy* – i.e., the attacker wishes to maximize $\mathbf{e}^a_k$, while ensuring that increase in the probability of the alarm β^a_k is not significant. Here, to formally capture this objective along with the stealthiness constraint, we denote the state estimation, residue, and estimation error of the compromised system by $\hat{\mathbf{x}}^a_k$, $\mathbf{z}^a_k$, and $\mathbf{e}^a_k$, respectively. We also define as

$$\Delta\mathbf{e}_k = \mathbf{e}^a_k - \mathbf{e}_k, \qquad \Delta\mathbf{z}_k = \mathbf{z}^a_k - \mathbf{z}_k,$$

the change in the estimation error and residue, respectively, caused by the attacks. From (1) and (6), they evolve as

$$\Delta\mathbf{e}_{k+1} = (\mathbf{A} - \mathbf{K}\mathbf{C}\mathbf{A})\Delta\mathbf{e}_k - \mathbf{K}\mathbf{a}_{k+1}, \tag{7}$$

$$\Delta\mathbf{z}_k = \mathbf{C}\mathbf{A}\Delta\mathbf{e}_{k-1} + \mathbf{a}_k, \tag{8}$$

with $\Delta\mathbf{e}_0 = 0$. Note that the above dynamical system is noiseless (and deterministic), with input $\mathbf{a}_k$ controlled by the attacker. Therefore, since $E[\mathbf{e}_k] = \mathbf{0}$ for the non-compromised system in steady state, it follows that

$$\Delta\mathbf{e}_k = E[\Delta\mathbf{e}_k] = E[\mathbf{e}^a_k]. \tag{9}$$

[1] Although a sensor itself may not be directly compromised with MitM attacks, but rather communication between the sensor and estimator, we will also refer to these sensors are *compromised sensors*. In addition, in this work we sometimes abuse the notation by using $\mathcal{K}$ to denote both the set of compromised sensors and the set of indices of the compromised sensors.

Because $\Delta\mathbf{e}_k$ provides expectation of the state estimation error under the attack, this signal can be used to evaluate the impact that the attacker has on the system.[2] Thus, we specify the objective of the attacker as to *maximize the state estimation error* $\|\Delta\mathbf{e}_k\|$. This is additionally justified by the fact that since $\mathbf{a}_k$ is controlled by the attacker, it follows that

$$Cov(\mathbf{e}_k^a) = Cov(\mathbf{e}_k) = \mathbf{\Sigma}.$$

To capture the attacker's stealthiness requirements, we use the probability of alarm in the presence of an attack

$$\beta_k^a = P(g_k^a > threshold), \text{ where } \; g_k^a = \mathbf{z}_k^{a\,T}\mathbf{P}^{-1}\mathbf{z}_k^a. \quad (10)$$

Therefore, to ensure that attacks remain stealthy, the attacker's constraint in each step k is to maintain

$$\beta_k^a \leq \beta_k + \epsilon, \quad (11)$$

for a small predefined value of $\epsilon > 0$.

C. Problem Formulation

As shown in the next section, for a large class of systems, a stealthy attacker can introduce an unbounded state estimation error with MitM attacks on communication channels between some of the plant sensors and the estimator. On the other hand, existing communication protocols commonly incorporate security mechanisms (e.g., MAC) that can ensure integrity of delivered sensor measurements. Specifically, this means that the system could enforce $\mathbf{a}_k = \mathbf{0}$ if integrity for all transmitted sensor measurements is enforced at some time-step k. Yet, the integrity enforcement comes at significant communication and computation cost, effectively preventing their continuous use in resource-constrained control systems. Hence, we focus on evaluating the impact of stealthy attacks in systems with intermittent (i.e., sporadic) use of data integrity enforcement mechanisms.[3]

III. Impact of Stealthy Attacks on State Estimation Error

To capture impact of the stealthy attacks on the system's performance we start with the following definition.

Definition 1: The set of all stealthy attacks up to time k is

$$\mathcal{A}_k = \{\mathbf{a}_{1..k} | \beta_{k'}^a \leq \beta_{k'} + \epsilon, \; \forall k', 1 \leq k' \leq k\}, \quad (12)$$

where $\mathbf{a}_{1..k} = [\mathbf{a}_1^T \dots \mathbf{a}_k^T]^T$. □

Definition 2: [13] The k-reachable region $\mathcal{R}_k$ of the state estimation error under the attack (i.e., $\Delta\mathbf{e}_k$) is the set

$$\mathcal{R}_k = \left\{ \; \Delta\mathbf{e}_k \in \mathbb{R}^n \; \middle| \; \begin{array}{c} \Delta\mathbf{e}_k, \Delta\mathbf{z}_k \text{ satisfy (7) and (8),} \\ \Delta\mathbf{e}_{k-1} \in \mathcal{R}_{k-1}, \beta_k^a \leq \beta_k + \epsilon \end{array} \right\}, \quad (13)$$

where $\mathcal{R}_0 = \mathbf{0} \in \mathbb{R}^n$. Also, the global reachable region $\mathcal{R}$ of the state estimation error $\mathbf{e}_k^a$ is set $\mathcal{R} = \bigcup_{k=0}^{\infty} \mathcal{R}_k$. □

Theorem 1: For any $\epsilon > 0$ there exists a unique $\alpha > 0$ such that $\beta_k^a \leq \beta_k + \epsilon$ if and only if $\|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}} \leq \alpha$. □

Proof: Without attacks, in steady-state $\mathbf{g}_k$ has χ^2 distribution with p degrees of freedom, since the residue $\mathbf{z}_k$ is zero-mean ($E[\mathbf{z}_k] = \mathbf{0}$) with covariance matrix $\mathbf{P} = \mathbf{C\Sigma C}^T + \mathbf{R}$. Furthermore, from (7) and (8), $\Delta\mathbf{z}_k = \mathbf{z}_k^a - \mathbf{z}_k$, is output of a deterministic system controlled by $\mathbf{a}_{1..k}$, and thus $\mathbf{z}_k^a$ is a non-zero mean with covariance matrix $\mathbf{P}$ – i.e., the attacker is only influencing the $\Delta\mathbf{z}_k = E[\mathbf{z}_k^a - \mathbf{z}_k] = E[\mathbf{z}_k^a]$. Therefore, $\mathbf{g}_k^a = \mathbf{z}_k^{a\,T}\mathbf{P}^{-1}\mathbf{z}_k^a$ will have a non-central χ^2 distribution with p degrees of freedom; the non-centrality parameter of this distribution will be $\lambda = \|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}}^2$.

Let h be the detector threshold in (5), and (10); the alarm probabilities $\beta_k = 1 - P(\mathbf{g}_k \leq h)$ and $\beta_k^a = 1 - P(\mathbf{g}_k^a \leq h)$ can be computed from the distributions for $\mathbf{g}_k$ and $\mathbf{g}_k^a$ as

$$\beta_k = 1 - F_{\chi^2}(h, p), \qquad \beta_k^a = 1 - F_{nc\chi^2}(h; p, \lambda),$$

where $F_{\chi^2}(h,p)$ and $F_{nc\chi^2}(h;p,\lambda)$ are cumulative distribution functions of χ^2 and noncentralized χ^2 respectively, at h, with p degrees of freedom and noncentrality parameter λ. Since p and h are fixed by the system design, it follows that β_k will be a constant, and β_k^a will be a function of λ.

Consider $\epsilon = \beta_k^a - \beta_k$. Since for noncentralized χ^2, the cumulative distribution function can be expressed as $F_{nc\chi^2}(h;p,\lambda) = 1 - Q_{\frac{p}{2}}(\sqrt{\lambda}, \sqrt{h})$, where $Q_{\frac{p}{2}}(\sqrt{\lambda}, \sqrt{h})$ is a Marcum Q-function [14], it follows

$$\epsilon = Q_{\frac{p}{2}}(\sqrt{\lambda}, \sqrt{h}) - \beta_k. \quad (14)$$

Since $Q_{\frac{p}{2}}(\sqrt{\lambda}, \sqrt{h})$ is smooth and monotonously increasing with respect to $\sqrt{\lambda}$ [15], for any ϵ there will exist exactly one $\sqrt{\lambda} = \|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}} = \alpha$ that (14) holds. Also, for any ϵ' lower than ϵ, the corresponding $\sqrt{\lambda'} = \|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}}$ from (14) must be lower than α, and vice versa, concluding the proof. ■

Since the bound α for $\|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}}$ in Theorem 1 depends on ϵ, h and the fact that the χ^2 detector with p degrees of freedom is used, we can denote such value as $\alpha = \alpha_{\chi^2}(\epsilon, p, h)$.

Remark 1: Related results from [10], [13] focus only on sufficient conditions for stealthy attacks; showing that if $\|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}} \leq \alpha$, it follows that $\beta_k^a \leq \beta + \epsilon$, – i.e., satisfying the stealthiness requirement. However, the proven full equivalence between conditions $\|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}} \leq \alpha$ and $\beta_k^a \leq \beta + \epsilon$ will allow us to use both conditions interchangeably when reasoning about the boundness of the reachability region. □

The previous results introduce an equivalent 'robustness-based' representation for the set of stealthy attacks.

Theorem 2: $\mathcal{R}_k$ is bounded if and only if the set

$$\hat{\mathcal{R}}_k = \left\{ \; \Delta\mathbf{e}_k \in \mathbb{R}^n \; \middle| \; \begin{array}{c} \Delta\mathbf{e}_k, \Delta\mathbf{z}_k \text{ satisfy (7) and (8),} \\ \Delta\mathbf{e}_{k-1} \in \hat{\mathcal{R}}_{k-1}, \|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}} \leq \alpha \end{array} \right\} \quad (15)$$

is bounded, where $\hat{\mathcal{R}}_0 = \mathbf{0} \in \mathbb{R}^n$ and $\alpha > 0$. □

Proof: Follows directly from Def. 1 and Thm. 1. ■

A. Perfectly Attackable Systems

We start by considering dynamical systems for which there exists a stealthy attack sequence that results in an unbounded expected state estimation error – i.e., for such systems, given enough time, the attacker can make arbitrary changes in the system states without risking detection.

Definition 3: A system is *perfectly attackable* (PA) if the reachable set $\mathcal{R}$ of the system is unbounded. □

[2] For this reason, and to simplify our presentation, in the rest of the paper we will refer to $\Delta\mathbf{e}_k$ as *the (expected) state estimation error* instead of *the change of the state estimation error caused by attacks*.

[3] Formal definition of such policies is introduced in Section IV.

As shown in [9], [11], for LTI systems additional data integrity guarantees, the set $\hat{\mathcal{R}}$ can be bounded or unbounded depending on dynamics and the set of compromised senso Theorem 2, this property is preserved for th well. For this reason, we will be using the $\hat{\mathcal{R}}$ to analyze the boundedness of $\mathcal{R}$, and to notation due to linearity of the constraint we that $\alpha = 1$. Furthermore, observe that $\frac{1}{|\lambda_{max}}$ $\|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}} \leq \frac{1}{|\lambda_{min}|}\|\Delta\mathbf{z}_k\|_2$, where λ_{max}, λ largest and smallest, respectively, eigenvalues of region $\hat{\mathcal{R}}$ will be bounded for the $\mathbf{P}^{-1}$-norm con only if it is bounded with a 2-norm stealthiness constraint – i.e., for sake of determining boundedness of $\hat{\mathcal{R}}$ we consider the stealthiness attack constraint as

$$\|\Delta\mathbf{z}_k\|_2 \leq 1, \qquad k \in \mathbb{N}_0. \tag{16}$$

Now, the next result follows from [9], [11].

Theorem 3: A system from (6) is perfectly attackable if and only if the matrix $\mathbf{A}$ is unstable, and at least one eigenvector $\mathbf{v}$ corresponding to an unstable mode satisfies that supp$(\mathbf{Cv}) \subseteq \mathcal{K}$ and $\mathbf{v}$ is a reachable state of the dynamical system (7). □

IV. Stealthy Attacks in Systems with Sporadic Integrity Enforcement

In this section, we analyze the effects of global sporadic data integrity enforcement policy. To formalize this notion, we start with the following definition.

Definition 4: Global sporadic data integrity enforcement policy (μ, f, L), where $\mu = \{t_k\}_{k=0}^{\infty}$, with $t_{k-1} < t_k$ for all $k > 0$ and $L = \sup_{k>0} t_k - t_{k-1}$, ensures that

$$\mathbf{a}_{t_k} = \mathbf{a}_{t_k+1} = ... = \mathbf{a}_{t_k+f-1} = \mathbf{0}, \forall k \geq 0.$$ □

Intuitively, a global sporadic data integrity policy ensures that the injected attack $\mathbf{a}_k$ will be equal to zero in at least f consecutive points, where the starts of these 'blocks' are at most L time-steps apart. The definition also captures *periodic* integrity enforcements when $L = t_k - t_{k-1}$ for all $k > 0$. Finally, the definition also captures policies with continuous integrity enforcements, by specifying $L \leq f$.

Note that in most systems networks are commonly shared among several control loops; thus, sporadic increases of network communication for these processes can be carefully interleaved without violating the network's bandwidth constraints. This would not be the case if all these processes impose integrity enforcements at every step, which is a common practice based on existing secure control methods.

The following theorem specifies that when a global sporadic integrity enforcement policy is used, a stealthy attacker can not insert an unbounded expected state estimation error.

Theorem 4: Consider an LTI system from (1) with a global data integrity policy (μ, f, L), where

$$f = \min\left(\text{nullity}(\mathbf{C}) + 1, q_{un}\right), \tag{17}$$

and q_{un} denotes the number of distinct unstable eigenvalues of $\mathbf{A}$. Then the system is not perfectly attackable. □

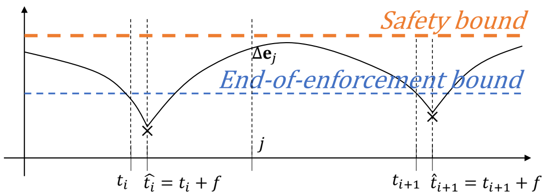


Fig. 3. System evolution between two consecutive endpoints of integrity enforcement intervals.

The above theorem makes no assumptions about the set of compromised sensors $\mathcal{K}$; since in general, system designers do not have this type of guarantees in design-time, no restrictions are imposed on the set, neither regarding the number of elements or whether some sensors belong to it.

To prove Theorem 4, we exploit the following result (we use the notation from Theorem 4).

Theorem 5: Consider $k \in \mathbf{N}$, such that $k+1 \in \mu$ (i.e., at time $k+1$ an integrity enforcement block in the policy μ starts). If $\Delta\mathbf{e}_k$ is a reachable state of the system $(\mathbf{A} - \mathbf{KCA}, \mathbf{K})$ under stealthy attacks and if the vectors $\mathbf{CA}\Delta\mathbf{e}_k, \mathbf{CA}\Delta\mathbf{e}_{k+1}, \ldots, \mathbf{CA}\Delta\mathbf{e}_{k+f-1}$ are bounded, then the vector $\Delta\mathbf{e}_{k+f}$ is bounded for any stealthy attack. □

Proof: Due to space constraints we omit the proof, which is available in [16]. ■

Using the previous theorem, we now prove Theorem 4.

Proof: [Proof of Theorem 4] Consider any time-point $t_k + f$ such that $t_k \in \mu$ – i.e., t_k is the start of an integrity enforcement block. Thus, $\mathbf{a}_{t_k} = ... = \mathbf{a}_{t_k+f-1} = \mathbf{0}$. From (8), we have $\Delta\mathbf{z}_{t_k+j} = \mathbf{CA}\Delta\mathbf{e}_{t_k+j-1}$, and thus from (16)

$$\|\mathbf{CA}\Delta\mathbf{e}_{t_k+j-1}\|_2 \leq 1, \qquad j = 0, ..., f-1.$$

Now, from Theorem 5 it follows that the expected state estimation error $\Delta\mathbf{e}_{t_k+f-1}$ has to be bounded for any stealthy attack; this holds for all time points at the ends of integrity enforcement intervals. Since in the proof of Theorem 5, we have not used any specifics of the time points, there exists a global bound on state estimation error at the end of all integrity enforcement periods (as illustrated in Fig. 3).

Finally, consider an expected state estimation error vector at any time j. From Definition 4, there exists $t_i \in \mu$ such that $j \in \left[\hat{t}_i, \hat{t}_i + L\right)$, where $\hat{t}_i = t_i + f$ (Fig. 3). From (7), (8)

$$\Delta\mathbf{e}_j = \mathbf{A}^{j-\hat{t}_i}\Delta\mathbf{e}_{\hat{t}_i} - \sum_{l=1}^{j-\hat{t}_i}\mathbf{A}^{j-\hat{t}_i-l}\mathbf{K}\Delta\mathbf{z}_{\hat{t}_i+l}. \tag{18}$$

Thus, the evolution of the expected state estimation error vector between two time points with bounded values can be described as evolution over a finite number of steps of a dynamical system with bounded inputs (as $\|\Delta\mathbf{z}_{\hat{t}_i+l}\|_2 \leq 1$); from the triangle and Cauchy-Schwarz inequalities it follows

$$\|\Delta\mathbf{e}_j\|_2 \leq \|\mathbf{A}\|_2^{j-\hat{t}_i}\|\Delta\mathbf{e}_{\hat{t}_i}\|_2 + \sum_{l=1}^{j-\hat{t}_i}\|\mathbf{A}\|_2^{j-\hat{t}_i-l}\|\mathbf{K}\|_2. \tag{19}$$

Hence, the state estimation error vector $\Delta\mathbf{e}_j$ is bounded for any j, and the system is not perfectly attackable. ■

Theorem 4 assumes that the attacker has the full knowledge of the system's integrity enforcement policy – i.e., at which time-points integrity enforcements will occur. As shown in Section VI, this allows the attacker to plan attacks that maximize the error, while ensuring stealthiness of the attack by reducing estimation errors to the levels that will not trigger detection during integrity enforcement intervals. However, if the attacker is not aware of the time points in which integrity enforcements would occur, the integrity enforcement requirements can be additionally relaxed. This is caused by the fact that the attacker has to ensure that if at any point he is unable to inject malicious data, the residue would still remain below the triggering threshold from (16).

Theorem 6: Any LTI system from (1) with a global data integrity policy $(\mu, 1, L)$ is not perfectly attackable for a stealthy attacker that does not have the knowledge of μ. □

Proof: Let us assume that the expected state estimation error $\Delta\mathbf{e}_k$ can be unbounded, and denote by $\mathbf{n}_k$ the unbounded part of the vector, as in the proof of Theorem 5. Note that the unbounded part of $\mathbf{A}\Delta\mathbf{e}_k$ (denoted as $\mathbf{A}\mathbf{n}_k$ in the proof of Theorem 5), can not always belong to $\mathcal{N}(\mathbf{C})$, if the attacker wants to introduce an unbounded expected state estimation error. If it did, (i.e., if $\mathbf{A}\mathbf{n}_k \in \mathcal{N}(\mathbf{C})$, for all k), from (8) and (16) it would follow that $\mathbf{a}_k$ is always bounded. This would imply that the dynamics from (7) has bounded inputs, which since the matrix $(\mathbf{A} - \mathbf{KCA})$ is stable ($\mathbf{K}$ is the Kalman gain), would imply that $\Delta\mathbf{e}_k$ can not diverge – the reachable set $\mathcal{R}$ can not be unbounded.

Thus, there exists k such that $\mathbf{A}\mathbf{n}_k \notin \mathcal{N}(\mathbf{C})$, which implies that $\mathbf{CA}\Delta\mathbf{e}_k$ is unbounded and $\|\mathbf{CA}\Delta\mathbf{e}_k\|_2 > 1$. Then, if global data integrity is enforced at the time-step $k+1$, from (8) it would follow that $\|\Delta\mathbf{z}_{k+1}\| = \|\mathbf{CA}\Delta\mathbf{e}_k\|_2 > 1$, violating the stealthiness requirement from (16). ■

Theorems 4 and 5 consider a worst-case scenario without any constraints or assumptions about the set of compromised sensors $\mathcal{K}$ (e.g., that less than q sensors are compromised). Yet, some knowledge about the set $\mathcal{K}$ may be available at design-time. For instance, for *MitM* attacks some sensors cannot be in set $\mathcal{K}$, such as on-board sensors that do not use a network to deliver information to the estimator, or sensors with built-in continuous data authentication. In these cases, the number of integrity enforcements can be reduced.

Corollary 1: Consider a system from (1) with a global data integrity policy (μ, f, L), where $f = \min\left(\text{nullity}(\mathbf{C}) + 1, q^*_{un}\right)$ and q^*_{un} denotes the number of distinct unstable eigenvalues λ_i of $\mathbf{A}$ for which the corresponding eigenvector $\mathbf{v}_i$ satisfies $\text{supp}\left(\mathbf{C}\mathbf{v}_i\right) \in \mathcal{K}$. Then the system is not perfectly attackable. □

Proof: Again, we assume that the expected state estimation error $\Delta\mathbf{e}_k$ can be unbounded, and denote by $\mathbf{n}_k$ the unbounded part, as in the proof of Theorem 5. Let $\alpha_1, \dots, \alpha_n$ be coefficients such that $\Delta\mathbf{e}_k = \alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n$, where $\mathbf{v}_1, \dots, \mathbf{v}_n$ are generalized eigenvectors of $\mathbf{A}$, and $\mathbf{n}_k = \alpha_1\mathbf{v}_1 + \dots + \alpha_r\mathbf{v}_r$. As shown in the proof of Theorem 6, $\mathbf{A}\mathbf{n}_k$ can not always belong to $\mathcal{N}(\mathbf{C})$. Then, the proof directly follows the proofs of Theorems 4 and 5, with the only difference that all $\alpha_i \to \infty$ in $\mathbf{n}_k$ also have to correspond to the unstable eigenvectors $\mathbf{v}_i$ satisfying that $\text{supp}\left(\mathbf{C}\mathbf{v}_i\right) \in \mathcal{K}$; otherwise, consider $\alpha_i \to \infty$, from a decomposition of a 'large' $\Delta\mathbf{e}_k$ such that $\mathbf{A}\mathbf{n}_k \notin \mathcal{N}(\mathbf{C})$, and the corresponding eigenvector $\mathbf{v}_i$ where $\text{supp}\left(\mathbf{C}\mathbf{v}_i\right) \notin \mathcal{K}$. Then, from (8) the components of the residue $\Delta\mathbf{z}_{k+1}$ whose indices are in $\text{supp}\left(\mathbf{C}\mathbf{v}_i\right)$ but not in the set $\mathcal{K}$ cannot be influenced by attack signal $\mathbf{a}_{k+1}$. Thus, their large values due to $\alpha_i \to \infty$ cannot be compensated for by the attack signal, and thus will violate the stealthiness condition (16). ■

V. Safe Integrity Enforcement Policies

With sporadic integrity enforcements, a stealthy attacker cannot introduce an unbounded state estimation error, irrelevant of the set $\mathcal{K}$. However, we still need to provide methods to evaluate if a specific integrity enforcement policy ensures safe estimation errors, either for a specific set $\mathcal{K}$ or the worst case $\mathcal{K} = \mathcal{S}$. While several metrics can be used to capture safety requirements, we require that the norm of induced estimation errors is always below a prespecified threshold ρ.

1) Evaluation of the State Estimation Error Regions: There exist algorithms that approximate estimation error regions (e.g., [10]), but with significant limitations due to the level of overapproximation or the fact that they do not support analysis with integrity enforcement in addition to the stealthiness constraints. Thus, we developed a method to estimate reachable regions in our case.

As $\|\Delta\mathbf{z}_k\|_\infty \le \|\Delta\mathbf{z}_k\|_2 \le \|\Delta\mathbf{z}_k\|_{\mathbf{P}^{-1}}|\lambda_{max}| \le \alpha|\lambda_{max}|$, the k-reachable regions can be overapproximated by capturing the stealthiness constraint as $\|\Delta\mathbf{z}_k\|_\infty \le \alpha|\lambda_{max}|$. Due to linearity of the constraints, we set the constraint to be $\|\Delta\mathbf{z}_k\|_\infty \le 1$, and multiply obtained values by $\alpha|\lambda_{max}|$ after. Thus, the system and attacker have to satisfy

$$\Delta\mathbf{e}_k = -\underbrace{\left[\begin{array}{c|c|c} (\mathbf{A} - \mathbf{KCA})^{k-1}\mathbf{K} & ... & \mathbf{K} \end{array}\right]}_{\mathbf{M}_k}\mathbf{a}_{1..k}$$

$$\Delta\mathbf{z}_k = \underbrace{\left[\begin{array}{c|c} -\mathbf{CAM}_{k-1} & \mathbf{I} \end{array}\right]}_{\mathbf{N}_k}\mathbf{a}_{1..k} \tag{20}$$

$$|\Delta\mathbf{z}_{j,i}| \le 1, \qquad i \in \{1, \dots, p\}, j \in \{1, \dots, k\},$$

where $\mathbf{a}_{1..k} = [(\mathbf{a}_1)^T \text{ ... } (\mathbf{a}_k)^T]^T$. This can be summarized as

$$\underbrace{\left[\begin{array}{c|c} \mathbf{I}_{n+pk} & \begin{array}{c|c} \multicolumn{2}{c}{\mathbf{M}_k P_{\mathcal{Q}}^\dagger} \\ \hline -\mathbf{N}_1 P_{\mathcal{K}}^\dagger & \mathbf{0}_{p\times(k-1)q} \\ \hline \multicolumn{2}{c}{\cdots} \\ \hline \multicolumn{2}{c}{-\mathbf{N}_k P_{\mathcal{Q}}^\dagger} \end{array} \\ \hline \multicolumn{2}{c}{\begin{array}{c|c|c} \mathbf{0}_{2pk\times n} & \begin{array}{c} \mathbf{I}_{kp} \\ \hline -\mathbf{I}_{kp} \end{array} & \mathbf{0}_{2pk\times kq} \end{array}} \end{array}\right]}_{(\mathbf{\Omega}_k)_{(n+3pk)\times(n+pk+kq)}} \underbrace{\left[\begin{array}{c} \Delta\mathbf{e}_k \\ \Delta\mathbf{z}_1 \\ \cdots \\ \Delta\mathbf{z}_k \\ P_{\mathcal{K}}\mathbf{a}_1 \\ \cdots \\ P_{\mathcal{K}}\mathbf{a}_k \end{array}\right]}_{\mathbf{r}_k^{rez}} \ge \underbrace{\left[\begin{array}{c} \mathbf{0}_{n+pk} \\ -\mathbf{1}_{2pk} \end{array}\right]}_{\mathbf{b}_k} \tag{21}$$

Here, $P_{\mathcal{K}} \in \mathbb{R}^{|\mathcal{K}|\times p}$ is the projection matrix that keeps only elements from the set $\mathcal{K}$, and $P_{\mathcal{Q}}$ is block-diagonal with k matrices $P_{\mathcal{K}}$ on the diagonal.

Let us introduce a k-reachable region $\tilde{\mathcal{R}}_k$ as in (15) with one difference - instead of the $\|\Delta\mathbf{z}_k\|_{\mathcal{P}^{-1}} \le \alpha$ requirement, we impose that $\|\Delta\mathbf{z}_k\|_\infty \le \alpha|\lambda_{max}|$. In addition, we introduce $\tilde{\mathcal{R}} = \cup_{k=0}^{\infty}\tilde{\mathcal{R}}_k$. Since, $\|\Delta\mathbf{z}_k\|_{\mathcal{P}^{-1}} \le \alpha \Rightarrow \|\Delta\mathbf{z}_k\|_\infty \le \alpha|\lambda_{max}|$ it follows that $\hat{\mathcal{R}} \subseteq \tilde{\mathcal{R}}$, and we use $\tilde{\mathcal{R}}$ to bound $\hat{\mathcal{R}}$.

From (21), $\tilde{\mathcal{R}}_k$ is a polyhedron in $\mathbb{R}^n$. Note that the maximal value of $\|\Delta\mathbf{e}_k\|_2$ over a polyhedron can be obtained in a vertex of the polyhedron [17]. The vertices of $\tilde{\mathcal{R}}_k$ satisfy

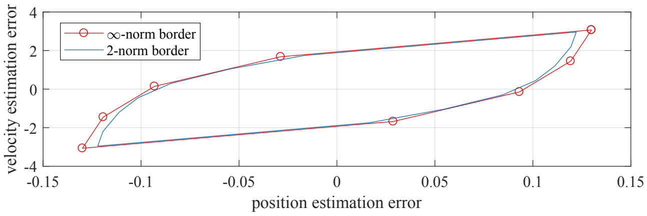


Fig. 4. 4-step reachable regions for 2-norm and ∞-norm.

that kq constraints from (20) are active. This means that all equalities and kq inequalities from (20) are active rows in (21). We define matrix $\mathbf{\Omega}_k^{act}$ that contains all active rows of $\mathbf{\Omega}_k$ from (21). Let $\{(\mathbf{\Omega}_k^{act})^1, ..., (\mathbf{\Omega}_k^{act})^{full}\}$ be the set of all such $\mathbf{\Omega}_k^{act}$ with the full rank. Then, if $(\mathbf{b}_k^{act})^i, 1 \leq i \leq full$ represent the corresponding values from $\mathbf{b}_k$, we define

$$(\mathbf{r}_k^{rez})^i = ((\mathbf{\Omega}_k^{act})^i)^\dagger (\mathbf{b}_k^{act})^i \tag{22}$$

where $((\mathbf{\Omega}_k^{act})^i)^\dagger$ denotes the pseudoinverse matrix of $(\mathbf{\Omega}_k^{act})^i$. Thus, the set of vertices of $\tilde{\mathcal{R}}_k$ can be expressed as $\{(\mathbf{r}_k^{rez})^i : (i \in \{1, 2, ..., full\}) \wedge ((\mathbf{r}_k^{rez})^i \text{ satisfies } (21))\}$. Finally, to determine the vertices of $\tilde{\mathcal{R}}_k$ in case when integrity is enforced at time points $H \subseteq \{1, \ldots, k\}$, we add $\mathbf{a}_j = \mathbf{0}$, $\forall j \in H$ to the system (20), and repeat the process.

The above procedure provides an estimate of the maximal estimation error in each step k. On one hand, the computation time grows exponentially with k and calculations for higher numbers of steps could become unfeasible. On the other hand, since we evaluate reachable state estimation errors to provide guidance on the effects and design of integrity enforcement policies, we did not face limitations caused by the computation times for analyzed systems; even after first integrity enforcements we would observe that new $\tilde{\mathcal{R}}_k \subset \cup_{i=0}^{k-1} \tilde{\mathcal{R}}_i$, which was exploited to reduce problem size.

In addition to the case study presented in Section VI, we evaluated the proposed reachable region estimation method on a simple vehicle model from [10], and analyzed the level of over-approximation due to the use of ∞-norm. For example, as shown in Fig. 4, the 4-step reachable region obtained by our method is a very good approximation of the actual reachability region. Similar results were obtained for other reachable regions. We also compared our method to the algorithm from [10] that recursively over- and under-approximates the estimation error with outer and inner ellipsoids. Although the method from [10] requires lower computation time, it does not directly allow for capturing the effects of integrity enforcement. It may also make arbitrarily large over-approximation for the outer ellipsoids, depending on the shape of the actual region. For example, using 4-step outer ellipsoidal region estimations presented in [10] would approximately double the required frequency of integrity enforcements compared to the one obtained by our method.

2) Determining safe integrity enforcement policies: Parameters of a safe integrity enforcement policy (μ, f, L) are obtained as follows. Value of f is chosen directly from (17). We determine elements of μ in iterative manner. First, we determine $t_1 \in \mu$ as the maximal time such that the first integrity enforcement at t_1 causes the attacker to reduce estimation error before $\|\Delta\mathbf{e}_k\|_2$ reaches ρ.[4] We similarly obtain $t_2 \in \mu$, but starting from $\mathcal{R}_{t_1}$ as the initial region. Note that we do not search through all possible $t_2 - t_1$; we rather evaluate candidates obtained as the minimal time an overapproximationsimilar to (19) needs to reach the safety threshold and return to the initial region. When this method was repeated, we observed that the time between starts of consecutive blocks would quickly settle and the policy would effectively move to periodic enforcement blocks.

Two caveats are in order. First, while this procedure is computationally heavy, it is performed at design-time (as opposed to runtime). Also, while the proposed policy would ensure system safety even in the presence of attack, providing optimal policies (e.g., that minimize average frequency of integrity enforcements) is an avenue for future work.

VI. Case Studies

We illustrate the use of sporadic data integrity enforcements on vehicle trajectory tracking, based on a model from [18]. The two-axis model is decoupled into separate one-axis models. After discretizing the system with sampling period $T_s = 0.01\ s$, we obtain state-space system matrices

$$\mathbf{A}_d = \begin{bmatrix} 1 & 0.01 \\ 0 & 1 \end{bmatrix} \quad \mathbf{B}_d = \begin{bmatrix} 0.0001 \\ 0.01 \end{bmatrix}, \tag{23}$$

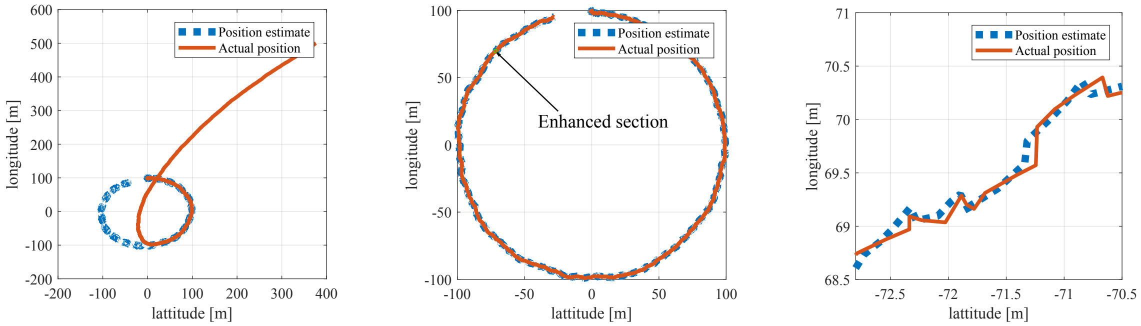
with Kalman gain $\mathbf{K} = \begin{bmatrix} 0.6180 & 0.0011 \\ 0.0011 & 0.6180 \end{bmatrix}$. We assume that the system is equipped with position and speed sensors, and consider the case where the attacker modifies values sent from both sensors. The system is perfectly attackable since the matrix $\mathbf{A}_d$ is unstable and supp$(\mathbf{Cv}) \in \mathcal{K}$.

We set the largest safe estimation error on position and speed as $0.025\ m$ and $0.025\ \frac{m}{s}$ respectively, resulting in $\|\Delta\mathbf{e}_{max}\|_2 = 0.0354$. Furthermore, we use $\alpha \leq 0.013$ as the attacker's constraint from Theorem 1. Since the system has a double-eigenvalue in 1, a stealthy attacker can force the state estimation error to increase linearly, crossing this threshold after three steps if no integrity enforcements are used. We studied several integrity enforcements policies – for all these policies $f = 1$. Specifically, we started with three global integrity enforcement policies that are periodic with periods $T = 3, 5$ and 6.

Using our framework from Section V, we show that the first two policies are safe and the third with period $T = 6$ could violate safety constraints, as illustrated in Fig. 6. Also, we generated a safe integrity enforcement policy μ_4 with $t_1 = 6$, followed by $t_2 = 5 = t_3 = ...$; Fig. 6 presents the state estimation error with this policy for a stealthy attack that also considers integrity enforcement times (referred to as *Epoch Attack*). Even when enforcing integrity on less than 20% of messages, we obtain low estimation errors in the presence of attacks on all sensors.

Finally, we show these effects on a simulation of a vehicle moving in a circular path of 100 m radius, at the speed of 3.14 $\frac{m}{s}$, and under attack from 100 s. Behavior of the

[4] The attacker may breach the threshold but would not be able to remain stealthy. To also ensure that the stealthiness constraint is violated before the bound is breached, a lower threshold should be used to generate the policy.



(a) State estimates for system under without integrity enforcements

(b) State estimates under stealthy Epoch Attack with integrity enforcements.

(c) Enhanced section of the Fig. 5(b).

Fig. 5. State estimation of the vehicle trajectory, the attack starts at 100 s - without integrity enforcements a stealthy attacker can introduce a significant estimation error in a short period of time. Yet, even with the sporadic integrity enforcement in less than 20% of steps, the attack effects are negligible.

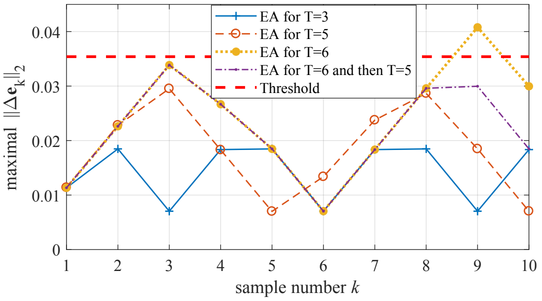


Fig. 6. Maximal estimation error for Epoch Attacks (EA) with four integrity enforcement policies (all with $f = 1$) – three periodic with periods $T = 3, 5, 6$, and one obtained by the framework from Sec. V with $t_1 = 6$ and $t_2 = 5 = t_3 = \dots$. All policies are safe, except the periodic with $T = 6$.

system under a stealthy attack without integrity enforcement is shown in Fig. 5(a). When a periodic integrity enforcement policy with $T = 4$ is used, state estimates are shown in Fig. 5(b), with a close-up portion presented in Fig. 5(c).

VII. Conclusion

We have considered the problem of network-based data-injection attacks on Kalman filter-based estimators, where the attacker can compromise sensor measurements sent from a subset of system sensors to the estimator (e.g., Man-in-the-Middle attacks). For these scenarios existing results impose very strict requirements to ensure limited control degradation. Thus, we have studied the effects of sporadic integrity enforcement policies, such as message authentication, on control performance in the presence of stealthy attacks. We have shown that even a sporadic enforcement of sensor data integrity significantly limits a stealthy attacker's capabilities – the attacker will not be able to introduce unbounded state estimation errors. We have also presented a design-time framework to derive integrity enforcement policies that ensure safe estimation errors. Finally, we have illustrated that even with very low utilization of integrity enforcement mechanisms, we can ensure satisfiable control performance, enabling resilient control implementation even on resource-constrained platforms shared by several control loops.

References

[1] T. M. Chen and S. Abu-Nimeh, "Lessons from stuxnet," *Computer*, vol. 44, no. 4, pp. 91–93, 2011.

[2] K. Zetter, "Inside the cunning, unprecedented hack of ukraine's power grid, wired magazine," March 2016.

[3] S. Checkoway and et. al, "Comprehensive experimental analyses of automotive attack surfaces," in *USENIX Security*, 2011.

[4] A. Teixeira, D. Pérez, H. Sandberg, and K. H. Johansson, "Attack models and scenarios for networked control systems," in *Conf. on High Confid. Net. Sys. (HiCoNS)*, 2012, pp. 55–64.

[5] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Trans. Autom. Control*, vol. 59, pp. 1454–1467, 2014.

[6] M. Pajic, I. Lee, and G. J. Pappas, "Attack-resilient state estimation for noisy dynamical systems," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 82–92, 2017.

[7] M. Pajic, J. Weimer, N. Bezzo, P. Tabuada, O. Sokolsky, I. Lee, and G. Pappas, "Robustness of attack-resilient state estimators," in *ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*, April 2014, pp. 163–174.

[8] S. Sundaram, M. Pajic, C. Hadjicostis, R. Mangharam, and G. Pappas, "The Wireless Control Network: Monitoring for Malicious Behavior," in *49th IEEE Conference on Decision and Control (CDC)*, Dec 2010, pp. 5979–5984.

[9] Y. Mo and B. Sinopoli, "False data injection attacks in control systems," in *First Workshop on Secure Control Systems*, 2010.

[10] ——, "On the performance degradation of cyber-physical systems under stealthy integrity attacks," *IEEE Transactions on Automatic Control*, vol. PP, no. 99, pp. 1–1, 2015.

[11] C. Kwon, W. Liu, and I. Hwang, "Analysis and design of stealthy cyber attacks on unmanned aerial systems," *Journal of Aerospace Information Systems*, vol. 1, no. 8, 2014.

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

[13] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, "False data injection attacks against state estimation in wireless sensor networks," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 5967–5972.

[14] A. Gil, J. Segura, and N. M. Temme, "Computation of the marcum q-function," *CoRR*, vol. abs/1311.0681, 2013.

[15] Y. Sun, Á. Baricz, and S. Zhou, "On the monotonicity, log-concavity, and tight bounds of the generalized marcum and nuttall-functions," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1166–1186, 2010.

[16] I. Jovanov and M. Pajic, "Relaxing integrity requirements for resilient control systems," *CoRR*, vol. abs/1707.02950, 2017. [Online]. Available: https://arxiv.org/abs/1707.02950

[17] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[18] A. J. Kerns, D. P. Shepard, J. A. Bhatti, and T. E. Humphreys, "Unmanned aircraft capture and control via gps spoofing," *Journal of Field Robotics*, vol. 31, no. 4, pp. 617–636, 2014.