Robustness of Attack-Resilient State Estimators

Miroslav Pajic  
*University of Pennsylvania*, pajic@seas.upenn.edu

James Weimer  
*University of Pennsylvania*, weimer@seas.upenn.edu

Nicola Bezzo  
*University of Pennsylvania*, nicbezzo@seas.upenn.edu

Paulo Tabuada  
*UCLA*, tabuada@ee.ucla.edu

Oleg Sokolsky  
*University of Pennsylvania*, sokolsky@cis.upenn.edu

See next page for additional authors

Follow this and additional works at: [http://repository.upenn.edu/cis_papers](http://repository.upenn.edu/cis_papers)

Part of the Computer Engineering Commons, and the Computer Sciences Commons

**Recommended Citation**

[http://dx.doi.org/10.1109/ICCPS.2014.6843720](http://dx.doi.org/10.1109/ICCPS.2014.6843720)


This paper is posted at ScholarlyCommons. [http://repository.upenn.edu/cis_papers/771](http://repository.upenn.edu/cis_papers/771)  
For more information, please contact repository@pobox.upenn.edu.
Robustness of Attack-Resilient State Estimators

Abstract
The interaction between information technology and physical world makes Cyber-Physical Systems (CPS) vulnerable to malicious attacks beyond the standard cyber attacks. This has motivated the need for attack-resilient state estimation. Yet, the existing state-estimators are based on the non-realistic assumption that the exact system model is known. Consequently, in this work we present a method for state estimation in presence of attacks, for systems with noise and modeling errors. When the the estimated states are used by a state-based feedback controller, we show that the attacker cannot destabilize the system by exploiting the difference between the model used for the state estimation and the real physical dynamics of the system. Furthermore, we describe how implementation issues such as jitter, latency and synchronization errors can be mapped into parameters of the state estimation procedure that describe modeling errors, and provide a bound on the state-estimation error caused by modeling errors. This enables mapping control performance requirements into real-time (i.e., timing related) specifications imposed on the underlying platform. Finally, we illustrate and experimentally evaluate this approach on an unmanned ground vehicle case-study.

Keywords

Disciplines
Computer Engineering | Computer Sciences

Comments

Author(s)
Miroslav Pajic, James Weimer, Nicola Bezzo, Paulo Tabuada, Oleg Sokolsky, Insup Lee, and George Pappas

This conference paper is available at ScholarlyCommons: http://repository.upenn.edu/cis_papers/771
Robustness of Attack-resilient State Estimators

Miroslav Pajic
Dept. of Electrical & Systems Eng.
University of Pennsylvania
pajic@seas.upenn.edu

James Weimer
Dept. of Computer & Information Sc.
University of Pennsylvania
{weimerj, nicbezzo}@seas.upenn.edu

Nicola Bezzo
Dept. of Computer & Information Sc.
University of Pennsylvania
pappasg@seas.upenn.edu

ABSTRACT

The interaction between information technology and physical world makes Cyber-Physical Systems (CPS) vulnerable to malicious attacks beyond the standard cyber attacks. This has motivated the need for attack-resilient state estimation. Yet, the existing state-estimators are based on the non-realistic assumption that the exact system model is known. Consequently, in this work we present a method for state estimation in presence of attacks, for systems with noise and modeling errors. When the estimated states are used by a state-based feedback controller, we show that the attacker cannot destabilize the system by exploiting the difference between the model used for the state estimation and the real physical dynamics of the system. Furthermore, we describe how implementation issues such as jitter, latency and synchronization errors can be mapped into parameters of the state estimation procedure that describe modeling errors, and provide a bound on the state-estimation error caused by modeling errors. This enables mapping control performance requirements into real-time (i.e., timing related) specifications imposed on the underlying platform. Finally, we illustrate and experimentally evaluate this approach on an unmanned ground vehicle case-study.

Categories and Subject Descriptors
K.6.5 [Management of Computing and Information Systems]: Security and Protection—Unauthorized access; C.3 [Special-purpose and Application-based Systems]: Process control systems

1. INTRODUCTION

Tight coupling of computation and communication substrates with sensing and actuation components in Cyber-Physical Systems (CPS) has introduced significant changes in the system design process; the heterogeneity of these systems has challenged the standard design methods that completely ignore cross-cutting constraints, as component-level understanding usually does not translate to the system level. A great example for CPS are modern vehicles that present a complex interaction of a large number of embedded Electronic Control Units, interacting with each other over different types of networks. In addition, there is a current shift in vehicle architectures, from isolated control systems to more open automotive architectures that would introduce new services such as remote diagnostics and code updates, and vehicle-to-vehicle communication.

Until recently, security of CPS (and embedded control systems before) has usually been an afterthought. However, the increasing set of functionalities, network interoperability, and system design complexity may introduce security vulnerabilities that are easily exploitable. The interaction between information technology and physical world have made CPS vulnerable to malicious attacks beyond the standard cyber attacks [5]. As shown in [11][6], using simple methods an attacker can disrupt the operation of a car to either disable the vehicle or hijack it. This problem is even more emphasized with
the rise of vehicle autonomy; thus, criticality analysis for various automotive components will have to be completely re-done. Similarly, attacks on CPS could hamper the critical infrastructure with undesired consequences as illustrated in the Maroochy Water breach \cite{18} and StuxNet virus attack on a SCADA system used in industrial processes control \cite{7}.

Relying exclusively on cyber-security techniques for securing CPS is insufficient. This is highlighted in cases when non-invasive sensor attacks occur – i.e., when the physical environment around a sensor is compromised to allow for injection of a malicious signal \cite{17}. For example, non-invasive attacks on Anti-lock Braking Systems are presented in \cite{17}. In addition, attacks on GPS sensors by spoofing a GPS to misguide a yacht off route are demonstrated in \cite{2}, while \cite{22} presents the steps and equipment required for GPS spoofing.

These results have spanned research into control-level techniques that address the problem of state estimation and intrusion detection under attacks on the environment of the controller, such as attacks on sensors, actuators and communication networks (e.g., \cite{19,21,15,9,20,13}). Attack-resilient state estimation has drawn considerable attention since knowing the system’s state even when some components have been compromised would allow for the use of the same controllers as in the case without attacks. For deterministic linear systems this problem has been recently mapped into $l_0$ optimization problem \cite{9,15}. However, the proposed techniques consider systems without noise and for which the exact model of the system’s dynamics is known. As a result, it is not clear whether most of the resilience guarantees hold when these assumptions are violated.

This problem is even more emphasized when issues inherently present in any system implementation are taken into account. For example, real-time issues such as sampling and actuation jitter, and synchronization errors between system components directly affect the controlled plant’s model that should be used for state estimation. Consequently, there is a need to provide a robust method for attack-resilient state estimation in the presence of noise and modeling errors. This would also allow for the extraction of system level requirements imposed by control algorithms on the underlaying OS and utilized networking, and facilitate reasoning about attack-resilience across different implementation layers.

We address this problem in the paper. Building on the work from \cite{9} we present a procedure for attack-resilient state estimation in presence of noise and modeling errors. We show that the attacker cannot destabilize the system by exploiting the difference between the model used for state estimation and the real physical dynamics of the system. Furthermore, we describe how implementation issues such as jitter, latency and synchronization errors can be mapped into parameters of the state estimation procedure (describing modeling errors), and provide a bound on the state-estimation error due to the modeling error. This effectively enables mapping control performance requirements into real-time (i.e., timing related) specifications imposed on the underlying platform. Finally, we illustrate the use of this approach on an autonomous robot case-study.

The rest of the paper is organized as follows. In Section \ref{sec:method}, we present an $l_0$-based state-estimator for deterministic linear systems, before describing common modeling errors caused by system implementation. Section \ref{sec:attack} presents a resilient state estimator that can be used in systems with modeling errors. In Section \ref{sec:experiments}, we show that the maximal state-estimation error is bounded, before we present a procedure that can be used to compute a bound. We evaluate this procedure in Section \ref{sec:evaluation} and illustrate its use on unmanned ground vehicle (UGV) case study (Section \ref{sec:case_studies}). Finally, we provide concluding remarks in Section \ref{sec:conclusion}.

\section{Notation and Terminology} \label{sec:notations}

For a set $S$, $|S|$ denotes the cardinality (i.e., size) of the set, while for two sets $S$ and $R$, we use $S \setminus R$ to denote the set of elements in $S$ that are not in $R$. In addition, for a set $K \subset S$, we specify the complement set of $K$ with respect to $S$ as $\overline{K} = S \setminus K$. We use $R$ to denote the set of reals, and $1^n$ to denote the row vector of size $N$ containing all ones.

We use $A^T$ to indicate the transpose of matrix $A$, while the $i^{th}$ element of a vector $x_k$ is denoted by $x_{k,i}$. For vector $x$ and matrix $A$, we use $|x|$ and $|A|$ to denote the vector and matrix whose elements are absolute values of the initial vector and matrix, respectively. Also, for matrices $P$ and $Q$, $P \preceq Q$ specifies that the matrix $P$ is element-wise smaller than or equal to the matrix $Q$.

For a vector $e \in \mathbb{R}^p$, the support of the vector is the set

\[ \text{supp}(e) = \{ i \mid e_i \neq 0 \} \subseteq \{1, 2, ..., p\}, \]

while the $l_0$ norm of vector $e$ is the size of supp$(e)$ – i.e., $\|e\|_{l_0} = |\text{supp}(e)|$. For a matrix $E \in \mathbb{R}^{p \times N}$, we use $e_1, e_2, ..., e_N$ to denote its columns and $E_1, E_2, ..., E_p$ to denote its rows. We define the row support of matrix $E$ as the set

\[ \text{rowsupp}(E) = \{ i \mid E_i \neq 0 \} \subseteq \{1, 2, ..., p\}. \]

As for vectors, the $l_0$ norm for a matrix $E$ is defined as $||E||_{l_0} = |\text{rowsupp}(E)|$.

\section{Motivation and Problem Description} \label{sec:method}

Consider a Linear-Time Invariant (LTI) system

\begin{align}
    x_{k+1} & = Ax_k + Bu_k + v_k, \\
    y_k & = Cx_k + w_k + e_k,
\end{align}

where $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^m$ is the input vector, $v_k \in \mathbb{R}^q$ is the actuation vector, $y_k \in \mathbb{R}^p$ is the output vector, $w_k \in \mathbb{R}^r$ is the sampling vector, and $e_k \in \mathbb{R}^l$ is the actuation vector. The matrices $A$, $B$, $C$ are the system matrices, while $v_k$, $w_k$, $e_k$ are the input, output, and actuation errors, respectively.

The goal is to estimate the state $x_k$ of the system using the output $y_k$ and the input $u_k$. This is a standard state estimation problem, and the solution is given by the Kalman filter.

The Kalman filter is a recursive algorithm that estimates the state of a system from a series of measurements. It uses a model of the system to predict the state, and then updates the prediction with the actual measurements.

The Kalman filter is a powerful tool for state estimation, but it assumes that the system is deterministic, and that the measurements are noise-free. In practice, these assumptions are often violated, and the Kalman filter may not be able to accurately estimate the state.

To address this problem, we propose a new method that takes into account the modeling errors caused by system implementation. This method is based on a robust optimization framework, and it can be used to estimate the state of a system even when the system is non-deterministic.

The method is summarized as follows:

1. **Modeling Errors**: The system is modeled as a linear time-invariant system, and the state is estimated using the Kalman filter. However, the model is not perfect, and there are modeling errors. The modeling errors are represented by the matrices $v_k$, $w_k$, and $e_k$.

2. **Robust Optimization**: The problem is formulated as a robust optimization problem, where the objective is to minimize the expected value of the state estimation error. The robust optimization problem is solved using a convex optimization algorithm.

3. **State Estimation**: The state is estimated using the solution of the robust optimization problem.

The method is illustrated on a case study of an autonomous robot, and it is shown that the method is able to accurately estimate the state of the system even when the system is non-deterministic.

The proposed method has several advantages over existing methods. First, it is able to handle modeling errors caused by system implementation. Second, it is able to provide a bound on the state estimation error. Third, it is able to be used on a wide range of systems, including non-deterministic systems.

The proposed method has several limitations. First, it assumes that the system is linear and time-invariant. Second, it assumes that the measurements are noise-free. Third, it assumes that the modeling errors are known.

The method can be extended to handle more complex systems, such as non-linear systems and systems with non-stationary measurements.

In summary, the proposed method is a robust and accurate method for state estimation in the presence of modeling errors. It is able to handle non-deterministic systems, and it is able to provide a bound on the state estimation error. The method is illustrated on a case study of an autonomous robot, and it is shown that the method is able to accurately estimate the state of the system even when the system is non-deterministic.
where $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ denote the plant’s state and input vectors, respectively, while $y \in \mathbb{R}^p$ is the plant’s output vector obtained from measurements of $p$ sensors from the set $S = \{ s_1, s_2, ..., s_p \}$. Accordingly, the matrices $A$, $B$ and $C$ have suitable dimensions. Furthermore, $v \in \mathbb{R}^n$ and $w \in \mathbb{R}^p$ denote the process and measurement noise vectors, while $e \in \mathbb{R}^p$ denotes the attack vector. To model attacks on sensors, we assume that sensors with indices in set $K \subseteq \{ 1, 2, ..., p \}$ are under attack. This means that $\theta_k = 0$ for all $i \in K^C$ and $k \geq 0$, where $K^C = S \setminus K$, and thus $\text{supp}(e_k) \subseteq K, \forall k \geq 0$.

**Remark 1.** The setup presented in this paper can be easily extended to include attacks on system actuators. In this case additional vector $e_k^A$ is added to the plant input at each step $k \geq 0$. As shown in [8], the same technique used for resilient-state estimation in the presence of attacks on sensors can be used to obtain the plant’s state when both sensors and actuators are compromised. Consequently, the analysis and results presented in this paper can be easily extended to the case when a subset of the actuators is also under attack.

**2.1 Attack-resilient State Estimation for Noisefree Systems**

For linear systems without noise (i.e., systems from (1) where $w_k = 0$ and $v_k = 0$, for all $k \geq 0$), a $l_0$-norm based method to extract state estimate in presence of attacks is introduced in [9]. To obtain the plant’s state at any time-step $t$ (i.e., $x_t$), the proposed procedure utilizes the previous $N$ sensor measurement vectors $(y_{t-N+1}, ..., y_t)$ and actuator inputs $(u_{t-N+1}, ..., u_{t-1})$ to evaluate the state $x_{t-N+1}$; the state is computed as the minimization argument of the following optimization problem$^2$

$$
\min_{x \in \mathbb{R}^n} ||Y_{t,N} - \Phi_N(x)||_{l_0}.
$$

Here, $Y_{t,N} = [\hat{y}_{t-N+1} | \hat{y}_{t-N+2} | \ldots | \hat{y}_t] \in \mathbb{R}^{p \times N}$ aggregates the last $N$ sensor measurements while taking into account the inputs applied during that interval

$$
\hat{y}_k = y_k, \quad k = t - N + 1
$$

$$
\hat{y}_k = y_k - \sum_{i=0}^{k-t-N+2} CA^iBu_{k-i}, \quad k = t - N + 2, ..., N
$$

In addition, $\Phi_N : \mathbb{R}^n \rightarrow \mathbb{R}^{p \times N}$ is a linear mapping defined as $\Phi_N(x) = [Cx | CAx | \ldots | CAX^{N-1}x]$, which captures the system’s evolution over $N$ steps caused by the initial state $x$.

The rationale behind the problem (2) is that the matrix $E_{t,N} = Y_{t,N} - \Phi_N(x_{t-N+1})$ presents the history of the last $N$ attacks vectors $e_{t-N+1}, ..., e_t$ i.e.,

$$
E_{t,N} = [e_{t-N+1} | e_{t-N+2} | \ldots | e_t] \in \mathbb{R}^{p \times N}.
$$

The critical observation here is that for a noisefree LTI system there is a pattern of zeros (i.e., zero-rows) in the matrix $E_{t,N}$ that corresponds to the non-attacked sensors and which remains constant over time; if $\mathcal{K}$ is the set of compromised sensors then $\text{rowsupp}(E_{t,N}) \subseteq \mathcal{K}$, for all $N$ and $t$ such that $N \geq 0$ and $t \geq N - 1$.

As shown in [8, 9] for noisefree systems the state estimator from (2) is optimal in the sense that if another estimator can recover $x_{t-N+1}$ then the one defined in (2) can as well. In addition, the estimator from (2) can extract the system’s state after $N$ steps when up to $q$ sensors are under attack if and only if for all $x \in \mathbb{R} \setminus \{0\}$,

$$
|\text{supp}(Cx) \cup \text{supp}(CAx) \cup \ldots \cup \text{supp}(CAX^{N-1}x)| > 2q
$$

We use $q_{\text{max}}$ to denote the maximal number of compromised sensors for which the system’s state can be recovered after $N$ steps despite attacks on sensors. Hence, if the number of compromised sensors $q$ satisfies that $q \leq q_{\text{max}}$, for noisefree systems the minimal $l_0$ norm of (2) is equal to $q$. Note that for these systems $q_{\text{max}}$ does not decrease with $N$, and due to Cayley-Hamilton theorem $\Phi_N$ cannot be further increased when more than $n$ previous measurements are used – i.e., $q_{\text{max}}$ obtains the maximal value for $N = n$. Finally, beside the measurement window size $N$, $q_{\text{max}}$ only depends on the system’s dynamics (i.e., matrices $A$ and $C$).

**2.2 Sources of Modeling Errors**

Besides measurement and process noise, vectors $v_k$ and $w_k$ in (1) can be used to capture any deviation in the plant model (1) from the real dynamics of the controlled physical system. Here, we present some of the common modeling errors introduced by non-idealities of control system implementation and limitations of the utilized computation and communication platforms. Specifically, we focus on the modeling errors caused by sampling and computation/actuation jitter, and synchronization errors between system components in scenarios where continuous-time plants are being controlled.

The described attack-resilient state estimator (2) is based on discrete-time model (1) of the system. Consequently, to be able to deal with continuous-time plants it is necessary to discretize the controlled plant, while taking into account real-time issues introduced by communication and computation schedules. To illustrate

$^1$Later we will assume that the noise vectors are constrained in certain ways. Furthermore, we will use $v$ and $w$ to capture different types of modeling errors which may be caused by some implementation (e.g., real-time) issues. We will address this in more details in Section 2.2.

$^2$The state $x_t$ can then be obtained from $x_{t-N+1}$ and the history of actuator inputs $(u_{t-N+1}, ..., u_{t-1})$ by applying the system evolution from (1) for $N - 1$ steps.

$^3$The size of the utilized measurement history $N$ is considered to be an input parameter to the resilient-state estimator. In the general case we should use the notation $q_{\text{max},N}$. 

this, consider a standard continuous-time plant model
\[
\dot{x}(t) = A_c x(t) + B_c u(t) \\
y(t) = C_c x(t),
\]
with state \(x(t) \in \mathbb{R}^n\), output \(y(t) \in \mathbb{R}^p\) and input vector \(u(t) \in \mathbb{R}^m\), where matrices \(A_c, B_c, C_c\) are of the appropriate dimensions.

We first consider setups where all plant’s output are sampled (i.e., measured) at times \(t_k, k \geq 0\), and where all actuators apply newly calculated inputs at times \(t_k + \tau_k, k \geq 0\), as shown in Fig. 1. We denote the \(k^{th}\) sampling period of the plant by \(T_{s,k} = t_{k+1} - t_k\), and note that the the input signal will have the form shown in Fig. 1(b). Using the approach from [10, 23], we describe the system as
\[
\begin{align*}
x_{k+1} &= A_k x_k + B_k u_k + B_k^- u_{k-1}, \\
y_k &= C x_k,
\end{align*}
\]
where \(u(t^+)\) is a piecewise continuous function that only changes values at time instances \(t_k + \tau_k, k \geq 0\). From the above equation, the discretized model of the system can be represented as
\[
\begin{align*}
x_k &= e^{A \tau_k} x_k + \int_0^{\tau_k} e^{A \xi} B d\xi, \\
y_k &= C e^{A \tau_k} x_k,
\end{align*}
\]
where \(x_k = x(t_k), k \geq 0\), and
\[
\begin{align*}
A_k &= e^{A \tau_k}, \\
B_k &= \int_0^{\tau_k} e^{A \xi} B d\xi.
\end{align*}
\]
Note that the matrices \(A_k, B_k, \) and \(B_k^-\) are time-varying (with \(k\)) and depend on the continuous-time plant dynamics, inter-sampling time \(T_{s,k}\), and latency \(\tau_k\). On the other hand, when control (and state estimation) is performed using resource constrained CPUs, the designers usually utilize the ‘ideal’ discrete-time model of the system of the form (1) where for all \(k \geq 0\), \(T_{s,k} = T_s\) and \(\tau_k = 0\)
\[
\begin{align*}
A &= e^{A T_s}, \\
B &= \int_0^{T_s} e^{A \xi} B d\xi.
\end{align*}
\]
Hence, by comparing the discrete-time models (1) and (6), in this case sampling and actuation jitter, and actuation latency (caused by computation and/or communication) introduce the error component \(v_k^{jit}\) \((k \geq 0)\) defined as
\[
v_k^{jit} = (e^{A \tau_k} - e^{A T_s}) x_k + \int_0^{\tau_k} e^{A \xi} B d\xi u_k + B_k^- u_{k-1}.
\]

Finally, from the equation above it follows that a bound on the size of the error \(v_k^{fit}\) can be obtained from the conservative bounds on the sampling jitter (i.e., \(T_{s,k} - T_s\)) and latency (i.e., \(\tau_k\)), for a predefined range of acceptable system states and actuator inputs.

2.2.1 Effects of Synchronization Errors

Consider a setup where the sensors do not have a common clock source, resulting in a bounded error between the individual clocks at sensors [1]. In this case, although scheduled to measure corresponding plant outputs at the same time-instance \(t_k\), each sensor \(s_j\) will actually perform the measurement at time \(t_{kj}\). Hence, for every \(j = 1, ..., p\), \(y_{k,j} = C' x(t_{kj})\) instead of \(C x(t_k)\), where \(C'\) denotes the \(j^{th}\) row of \(C\), and the synchronization error introduces a measurement error defined as
\[
v_k^{syn} = C' (x(t_{kj}) - x(t_k)) = C' \int_0^{\Delta t_{kj}} e^{A \xi} B d\xi u_k - \int_0^{\Delta t_{kj}} e^{A \xi} B d\xi u_{k-1}.
\]
Here, \(\Delta t_{kj} = t_{kj} - t_k\) captures the synchronization error for each sensor \(s_j\). Hence, for a predefined actuation range it is possible to provide a bound on the size of the measurement error vector \(v_k^{syn} \in \mathbb{R}^p\) describing modeling errors due to synchronization errors between sensors.

2.3 Problem Description

The existence of modeling error [1] limits the use of the attack-resilient state estimator from (2). For example, in this case the \(l_0\) norm of a solution of the problem

4To simplify the presentation we only describe the case where there possibly exist synchronization errors between sensors, because the same approach can be extended to scenarios where there exist synchronization errors between plant actuators.

5In the rest of this paper, unless otherwise specified we will include process and measurement noise as part of modeling errors. We will also use the term noise to capture the modeling errors – i.e., discrepancy between the model used to design the state-estimator and real dynamics of the controlled system.
in (2) may be larger than \( q_{\text{max}} \), indicating that more than the allowed number of sensors has been compromised, which violates requirements for correct operation of the state estimator. Therefore, in this paper we focus on the following problems:

- How can we design attack-resilient state estimators in presence of noise and modeling errors?
- Can the attacker exploit the modeling errors in the state-estimator’s design in order to destabilize the system when the state estimates are used for control by a stable state-based feedback controller?
- Is it possible to obtain a bound on the worst-case performance degradation of the introduced resilient-state estimator due to bounded modeling errors?

3. RESILIENT STATE ESTIMATION IN PRESENCE OF MODELING ERRORS

As illustrated in Section 2, the effects of the input vectors \( \mathbf{u}_k \) are taken into account when computing the matrix \( \mathbf{Y}_{l,N} \). Thus, in the rest of this paper (unless otherwise stated) we will assume that in (1) \( \mathbf{u}_k = \mathbf{0} \) for all \( k \geq 0 \). In addition, to further simplify the notation we consider the case for \( t = N - 1 \), meaning that our goal is to obtain \( \mathbf{x}_0 \), and we denote the matrices \( \mathbf{Y}_{t,N}, \mathbf{E}_{t,N} \) and \( \Phi_N(x) \) as \( \mathbf{Y}, \mathbf{E} \) and \( \Phi(x) \), respectively.

We assume that the state of the plant at \( k = 0 \) is \( \mathbf{x}_0 \) and that the system evolves for \( N \) steps as specified in (1) (for \( \mathbf{u}_k = \mathbf{0} \)) for some attack vectors \( \mathbf{e}_0, \ldots, \mathbf{e}_{N-1} \) applied on the sensors from set \( \mathcal{K} = \{ s_{i_1}, \ldots, s_{i_q} \} \subseteq \mathcal{S} \), where \( |\mathcal{K}| \leq q_{\text{max}} \), and the corresponding matrix \( \mathbf{E} = [\mathbf{e}_0 | \mathbf{e}_1 | \ldots | \mathbf{e}_{N-1}] \). Furthermore, we assume that \( |\mathbf{w}_k| \geq \epsilon_{\mathbf{w}_k} \) and \( |\mathbf{v}_k| \geq \epsilon_{\mathbf{v}_k} \) for \( k = 0, 1, \ldots, N - 1 \), and we define \( \mathbf{Y}_{w,v} = [\mathbf{y}_0 | \mathbf{y}_1 | \ldots | \mathbf{y}_{N-1}] \).

Note that the matrix \( \mathbf{Y}_{w,v} \) contains measurements of the system with noise. Finally, we use \( \mathbf{Y} = [\mathbf{y}_0 | \mathbf{y}_1 | \ldots | \mathbf{y}_{N-1}] \) to denote the sensor measurements (plant outputs) that would be obtained in this case if the system was noiseless – i.e., for \( \|\epsilon_{\mathbf{w}_k}\|_2 = \|\epsilon_{\mathbf{v}_k}\|_2 = 0 \) (meaning that \( \mathbf{y}_k = \mathbf{CA}^k\mathbf{x}_0 + \mathbf{e}_k, \ k = 0, 1, \ldots, N - 1 \)).

We consider the following optimization problem:

\[
P_0(\mathbf{Y}) : \quad \min_{\mathbf{E}, \mathbf{x}} \|\mathbf{E}\|_{l_0} \quad \text{s.t.} \quad \mathbf{E} = \mathbf{Y} - \Phi(x)
\]

As described in Section 2, we have that

\[
(x_0, \mathbf{E}) = \arg \max_{\mathbf{Y}} P_0(\mathbf{Y})
\]

where \( q = \|\mathbf{E}\|_{l_0} \leq q_{\text{max}} \). However, the ‘ideal’ (noiseless) measurements from \( \mathbf{Y} \) are not available to the estimator; the estimator can only use the measurements specified by the matrix \( \mathbf{Y}_{w,v} \). In addition, it is worth noting that \((x_0, \mathbf{E})\) may not even be a feasible point for problem \( P_0(\mathbf{Y}_{w,v}) \) that utilizes noisy sensor measurements. Consequently, there is need to adapt problem \( P_0(\mathbf{Y}) \) to non-ideal models that capture noise and modeling errors.

To achieve this we consider the following problem that relaxes the equality constraint from (9) by including a noise allowance

\[
P_{0,\Delta}(\mathbf{Y}) : \quad \min_{\mathbf{E}, \mathbf{x}} \|\mathbf{E}\|_{l_0} \quad \text{s.t.} \quad \|\mathbf{Y} - \Phi(x) - \mathbf{E}\|_2 \leq \Delta
\]

(11)

In the above problem, the matrix \( \Delta \in \mathbb{R}^{p \times N} \) contains non-negative tolerances \( \delta_{j,i} \) for each sensor \( s_i \), \( i = 1, \ldots, p \), in each of the \( N \) steps \( j \) – i.e.,

\[
\Delta = [\delta_0 | \delta_1 | \ldots | \delta_{N-1}].
\]

We use the following notation

\[
(x_0, \Delta, E_\Delta) = \arg \max_{\mathbf{Y}} P_{0,\Delta}(\mathbf{Y}_{w,v})
\]

(12)

\[
q_\Delta = \|E_\Delta\|_{l_0}
\]

Note that \( P_{0,\Delta}(\mathbf{Y}) = P_0(\mathbf{Y}) \), for all \( \mathbf{Y} \in \mathbb{R}^{p \times N} \).

To allow for the use of (11) as an attack-resilient state estimator it is necessary to ensure that \( P_{0,\Delta}(\mathbf{Y}) \) has a feasible point \( (\mathbf{x}, \mathbf{E}) \) such that \( \|\mathbf{E}\|_{l_0} \leq q_{\text{max}} \). This can be guaranteed with an appropriate initialization of the matrix \( \Delta \). From (1) we have that for \( k = 0, 1, \ldots, N - 1 \)

\[
\mathbf{y}_k = \mathbf{CA}^k\mathbf{x}_0 + \mathbf{e}_k + \mathbf{C} \sum_{i=0}^{k-1} \mathbf{A}^{k-1-i}\mathbf{v}_i + \mathbf{w}_k
\]

\[
= \hat{\mathbf{y}}_k + \mathbf{C} \sum_{i=0}^{k-1} \mathbf{A}^{k-1-i}\mathbf{v}_i + \mathbf{w}_k
\]

If we use \( |\mathbf{A}^{k-1-i}| \) to denote the matrix whose elements are absolute values of the corresponding elements of the matrix \( \mathbf{A}^{k-1-i} \), then we can specify the following bound

\[
|\mathbf{y}_k - \hat{\mathbf{y}}_k| \leq |\mathbf{C}| \sum_{i=0}^{k-1} |\mathbf{A}^{k-1-i}| |\mathbf{v}_i| + |\mathbf{w}_k|
\]

\[
\leq |\mathbf{C}| \sum_{i=0}^{N-1} |\mathbf{A}^{k-1-i}| |\epsilon_{\mathbf{v}_i}| + |\mathbf{w}_i| = \tilde{\delta}_k.
\]

Therefore, for \( \delta_k \geq \tilde{\delta}_k \) \( (k = 0, \ldots, N - 1) \) we have that \( (x_0, \mathbf{E}) \) from (10) is a feasible point for the problem \( P_{0,\Delta}(\mathbf{Y}_{w,v}) \), meaning that there exists a solution of the problem – i.e., there exists \( (x_0, \Delta, E_\Delta) \) from (12) such that \( q_\Delta \leq q_{\text{max}} \). This means that the solution of \( P_{0,\Delta}(\mathbf{Y}_{w,v}) \) from (11) can be used as a state-estimator in the sense that if at most \( q_{\text{max}} \) sensors have been compromised it would provide a solution where the size of row-support of \( E_\Delta \) is not larger than \( q_{\text{max}} \).

\(^6\)This condition has to be satisfied for all \( \mathbf{Y} \in \mathbb{R}^{p \times N} \) that could be ‘generated’ by the system when at most \( q_{\text{max}} \) sensors have been attacked.

\(^7\)We assume that \( \sum_{i=0}^{N-1} \alpha_i = 0 \) for any sequence of \( \alpha_i \).
4. ROBUSTNESS OF $P_{0,\Delta}(Y)$ STATE ESTIMATION

In this section, we provide robustness analysis for $P_{0,\Delta}(Y)$ optimization problem when matrix $\Delta$ satisfies conditions from the previous section. We start by showing that the attacker cannot exploit the modeling errors to destabilize the system before we present a method to bound the error caused by noise and modeling errors – i.e., we provide a bound on $||x_0,\Delta - x_0||_2$.

Consider $(x_0,\Delta, E\Delta)$ from $[12]$, and a matrix $\Sigma \in \mathbb{R}^{P \times N}$ such that
\[
Y - \Phi(x_0,\Delta) - E\Delta = \Sigma. \tag{14}
\]
Here, $|\Sigma| \leq \Delta$. In addition, because $(x_0, E)$ is a feasible point for $P_{0,\Delta}(Y)$, it follows that
\[
q = ||E||_{l_0} \geq ||E\Delta||_{l_0} = q_\Delta,
\]
implying that $||E - E\Delta||_{l_0} \leq 2q$. Our goal is to provide a bound on $||\Delta x||_2$
\[
\Delta x = x_0,\Delta - x_0. \tag{15}
\]
If we also define $\Delta E = E\Delta - E$, it holds that
\[
\Delta E = (Y_{w,v} - \Phi(x_0,\Delta) - \Sigma) - (Y - \Phi(x_0)) = (Y_{w,v} - \bar{Y} - \Sigma) - \Phi(\Delta x_0)
\]
Let us denote by $\Delta y_0, ..., \Delta y_{N-1}$ the columns of the matrix $\Omega$ (i.e., $\Omega = [\Delta y_0|...|\Delta y_{N-1}]$). From $[13]$ and $[14]$ it follows that
\[
|\Delta y_k| \leq \bar{d}_k + d_k \leq 2\bar{d}_k
\]
Accordingly, to provide a bound on $||\Delta x||_2$ we consider the following problem
\[
\max_{\Delta x} \; ||\Delta x||_2 \tag{16}
\]
\[
||\Phi(\Delta x) - \Omega||_{l_0} \leq 2q \tag{17}
\]
\[
\Omega \leq 2\Delta \tag{18}
\]
Since $q \leq q_{\text{max}}$, we can increase the feasible space by relaxing constraint (17) to
\[
||\Omega - \Phi(\Delta x)||_{l_0} \leq 2q_{\text{max}} \tag{19}
\]
Therefore, our goal is to bound $\Delta x$ for which there exists $\Omega \in \mathbb{R}^{P \times N}$ that satisfies $[13]$, and for which at least $p - 2q_{\text{max}}$ rows of the matrix $\Phi(\Delta x) - \Omega$ are zero-rows. Let us use $F$ and $K_F \subset S$ to denote the number of rows $\Phi(\Delta x)$ that are zero-rows and the set of corresponding sensors, respectively. This means that at least $F_c = p - 2q_{\text{max}} - F$ rows of $\Phi(\Delta x)$ are equal to the rows of $\Omega$, which are non-zero, and we use $K_{F_1} \subset S$ to denote sensors corresponding to those rows. It is worth noting here that $|K_F \cup K_{F_1}| = p - 2q_{\text{max}}$ and $K_F \cap K_{F_1} = \emptyset$.

We use the following notation – for any set $K \subseteq S$ we define the matrix $O_K$ as
\[
O_K = \begin{bmatrix} P_F \cdot C && \cdots \\ P_F \cdot CA & \cdots & \cdots \\ P_F \cdot CA^{N-1} \end{bmatrix}. \tag{20}
\]
Here, $P_F$ denotes the projection from the set $S$ to the set $K$ by keeping only rows of $C$ with indices that correspond to sensors from $K$. Since $K_F \subset S$ contains indices of zero-rows of $\Phi(\Delta x)$ we have that $O_{K_F} \cdot \Delta x = 0$. In addition, $O_{K_{F_1}} \cdot \Delta x = \Omega_{K_{F_1}}$, where for $\Omega = [\omega_1|\omega_2|...|\omega_N]$ (i.e., $\omega_i, i = 1,...,N$ are columns of $\Omega$ such that $|\omega_i| \leq 2\delta_i$, and we define
\[
\Omega_{K_{F_1}} = \begin{bmatrix} P_{K_{F_1}} \cdot \omega_1 \\ P_{K_{F_1}} \cdot \omega_2 \\ \vdots \\ P_{K_{F_1}} \cdot \omega_N \end{bmatrix} \quad \Delta_{K_{F_1}} = \begin{bmatrix} P_{K_{F_1}} \cdot \delta_1 \\ P_{K_{F_1}} \cdot \delta_2 \\ \vdots \\ P_{K_{F_1}} \cdot \delta_N \end{bmatrix}.
\]
Consequently, for $\Delta x$ to satisfy constraints $[19]$ and $[18]$ there have to exist sets $K_F, K_{F_1} \subset S$ such that
\[
|K_F| = F, \; |K_{F_1}| = p - 2q_{\text{max}} - F, \tag{21}
\]
\[
K_F \cap K_{F_1} = \emptyset \tag{22}
\]
\[
O_{K_F} \cdot \Delta x = 0 \tag{23}
\]
\[
|O_{K_{F_1}} \cdot \Delta x| \leq 2\Delta_{K_{F_1}} \tag{24}
\]
We consider the polyhedron $\mathbb{P}$ defined with constraints $[21] - [24]$. Note that the point $\Delta x = 0$ belongs to the polyhedron. In the rest of this section we will first prove that the polyhedron is bounded, before we present a procedure that can be used to provide a bound on $\max ||\Delta x||_2$. We start by introducing the following lemma.

**Lemma 1.** For any two sets $K_F, K_{F_1} \subset S$ such that
\[
|K_F| = F, \; |K_{F_1}| = p - 2q_{\text{max}} - F \; \text{and} \; K_F \cap K_{F_1} = \emptyset, \tag{25}
\]
\[
\text{rank}(O_{K_F}) = n. \quad \text{rank}(O_{K_{F_1}}) = n.
\]

**Proof.** From $[9], q_{\text{max}} = \lfloor s/2 - 1 \rfloor$ where $s$ is the cardinality of the smallest set $K \subseteq S$ for which the matrix $O_K$ has non-trivial kernel. Note that $|K_F^b| = p - s$, and since $s \geq 2q_{\text{max}} + 1 > 2q_{\text{max}}$, it follows that $|K_F^b| < p - 2q_{\text{max}}$. Now consider any set $K_1$ for which $|K_1^b| \geq p - 2q_{\text{max}}$, meaning that $|K_1| \leq 2q_{\text{max}} < s$. Thus, $O_{K_1^b}$ does not have non-trivial kernel (since $K_1$ is the smallest such matrix), meaning that columns of $O_{K_1^b}$ are linearly independent. Thus, $O_{K_1^b} \in \mathbb{R}^{N|K_1^b| \times n}$, \[\text{rank}(O_{K_1^b}) = n.\]

$^8$Formally, $P_K = \begin{bmatrix} \hat{u}_k \\ \vdots \\ \hat{u}_k \end{bmatrix}$, where $K = \{s_k,...,s_{k_{|K|}}\}$ and $k_1 < k_2 < ... < k_{|K|}$, and $\hat{u}_j$ denotes the row vector (of appropriate size) with a 1 in its $j^{th}$ position.
The above relation holds for any $K_F^1$ with at least $p - 2q_{max}$ sensors, and hence (25) holds since the set $K_F \cup K_{F_1}$ contains $p - 2q_{max}$ sensors. □

**THEOREM 1.** The polyhedron $P$ defined by constraints (21)–(24) is bounded.

**Proof.** We start by assuming the opposite, that $P$ is unbounded; there exist a feasible point $\Delta x \in P$ and a direction $d \in \mathbb{R}^n$ such that $d \neq 0$ and for any $\epsilon > 0$, $\Delta x + \epsilon d \in P$. Therefore, $O_{K_F}(\Delta x + \epsilon d) = 0$, and since $\Delta x \in P$ it follows that $O_{K_F} d = 0$. In addition,

$$|O_{K_{F_1}}(\Delta x + \epsilon d)| \geq 2\Delta_{K_{F_1}}$$

implies that $O_{K_{F_1}} d = 0$ (otherwise, for any non-zero element of the vector $O_{K_{F_1}} d$, when $\epsilon \to \infty$ the absolute value of that element in vector $\epsilon O_{K_{F_1}} d$ will be unbounded and the constraint (20) will be violated).

Therefore, $d$ belongs to the kernel of $O_{K_F \cup K_{F_1}}$ i.e., $O_{K_F \cup K_{F_1}} d = 0$. However, from Lemma 1, $O_{K_F \cup K_{F_1}} d$ has full rank (i.e., rank($O_{K_F \cup K_{F_1}}$) $=$ $n$), meaning that it does not have a non-trivial kernel and thus $d = 0$, which violates our initial assumption and concludes the proof. □

As a direct consequence of the above theorem we have that maximal $\|\Delta x\|_2$ is bounded, and the attacker can not use modeling errors and the corresponding relaxation of the I0 optimization problem to introduce unbounded error in the attack-resilient state estimator. Thus, we can formulate the corollary.

**COROLLARY 1.** Consider a matrix $K$ such that $A + BK$ is stable. If $x_0, \Delta$ from (12) is used as state estimate by the state-feedback controller $K$ (i.e., $u = Kx_0, \Delta$) then the closed-loop system will remain stable when at most $q_{max}$ sensors have been compromised.

### 4.1 Bounding the State-estimation Error

**Theorem 2** also facilitates bounding the error of the resilient state estimator $P_{0, \Delta}(Y_{w, v})$ by starting from the following lemma.

**LEMMA 2.** Consider the optimization problem

$$\max_{x \in Q} f(x)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and $Q$ is a bounded polyhedron. Then, there exists an optimal solution $f(x^*)$ where $x^*$ is a vertex of the polyhedron $Q$.

**Proof.** See Appendix □

From the above lemma, to determine the maximal $\|\Delta x\|_2$ over the polyhedron $P$ it is sufficient to compute $\|\Delta x\|_2$ at each vertex of the polyhedron. On the other hand, the vertices of the polyhedron satisfy that

$$\begin{bmatrix} O_{K_F} & O_{K_{F_1}} \end{bmatrix} \Delta x = \begin{bmatrix} 0 \\ 2\Delta_{K_{F_1}}^{+} \end{bmatrix},$$

(27)

where $\Delta_{K_{F_1}}^{+}$ denotes a vector such that $|\Delta_{K_{F_1}}^{+}| = \Delta_{K_{F_1}}$ (i.e., with elements whose absolute values are equal to the corresponding elements of $\Delta_{K_{F_1}}$). It is worth noting that there are $2^{K_{F_1} \cup N}$ such elements and thus $2^{K_{F_1} \cup N}$ vertices of the polyhedron. Finally, since $O_{K_F \cup K_{F_1}}$ is a full rank matrix ($\text{rank}(O_{K_F \cup K_{F_1}}) = n$), vertex points can be found as

$$\Delta x_{\text{ver}} = (O_{K_F \cup K_{F_1}}^T O_{K_F \cup K_{F_1}})^{-1} O_{K_F \cup K_{F_1}}^T \begin{bmatrix} 0 \\ 2\Delta_{K_{F_1}}^{+} \end{bmatrix}.$$  

(28)

Consequently, for any sets $K_F$ and $K_{F_1}$ that satisfy (21) and (22), by checking all $2^{K_{F_1} \cup N}$ vertices defined by (28) we can determine the maximal $\|\Delta x\|_2$ for the corresponding polyhedron. However, since

$$\|\Delta x_{\text{ver}}(\Delta_{K_{F_1}}^{+})\|_2 = \|\Delta x_{\text{ver}}(\Delta_{K_{F_1}}^{-})\|_2,$$

where $\Delta x_{\text{ver}}(\Delta_{K_{F_1}}^{+})$ denotes the solution of (28) for specific $\Delta_{K_{F_1}}^{+}$, we only need to evaluate norms at $2^{K_{F_1} \cup N}$ points (i.e., vertices). Furthermore, to provide a bound on $\|\Delta x\|_2$ for all $\Delta x$ that satisfy (18) and (19), we have to consider all such sets $K_F$ and $K_{F_1}$. Therefore, it is necessary to evaluate all possible values for $F$. From the definition $F \geq 0$. On the other hand, from (23) $O_{K_F}$ has a nontrivial kernel, meaning that as in the proof of Lemma 1, $F = \text{max}_F \leq p - s \leq p - 2q_{max} - 1$.

This allows for the formulation of Algorithm 1 and proves the following theorem.

**THEOREM 2.** Algorithm 1 provides an upper bound on the state estimation error $\|x_0, \Delta - x_0\|_2$.

Finally, note that the matrix $\Delta$ captures several sources of modeling errors (e.g., noise, jitter, synchronization errors). Since (28) is linear in $\Delta$, the estimation error bound obtained by Algorithm 1 for $\Delta$ will be less than or equal to the sum of estimation error bounds computed separately for each error component. Thus, it is possible to analyze the impact of each source of modeling errors on robustness of the state estimator.

### 4.1.1 Complexity of Algorithm 1

The complexity of Algorithm 1 depends on the number of plant states and sensors ($n$ and $p$, respectively), considered plant window size (i.e., history) $N$, and the maximal number of attacked sensors $q_{max}$ for which it is possible to estimate the state of the noiseless system after $N$ steps. Hence, in the general case the number of times that equation (28) needs to be solved is

$$\sum_{F=0}^{p-s} \binom{p}{F} \binom{p-F}{2p-2q_{max}-F}3^{(p-2q_{max}-F)N-1}.$$

However, for almost all systems, meaning that for almost all pairs of matrices $A \times C \in \mathbb{R}^{n \times n} \times \mathbb{R}^{p \times n}$ (i.e.,
Algorithm 1 Design-time procedure used to provide a bound on $\Delta x$

1: $MAX_\|\Delta x\| = 0$
2: for $F = 0, 1, \ldots, p - s$ do
3: \hspace{1em} for all sets $K_F \subset S$ with $F$ elements do
4: \hspace{2em} \hspace{1em} for all sets $K_{Fi} \subset (S \setminus K_F)$ with $p - 2q_{max} - F$ elements do
5: \hspace{3em} for all $\Delta^+_i \subset K_{Fi}$ do
6: \hspace{4em} Compute $\Delta x_{ver}$ using (28)
7: \hspace{4em} if $\|\Delta x_{ver}\|_2 > MAX_\|\Delta x\|_2$ then
8: \hspace{5em} $MAX_\|\Delta x\| = \|\Delta x_{ver}\|_2$
9: \hspace{4em} \hspace{1em} \hspace{1em} end if
10: \hspace{3em} \hspace{1em} end for
11: \hspace{2em} \hspace{1em} end for
12: \hspace{1em} end for
13: end for

the set of matrices for which the property does not hold has Lebesgue measure zero), the number of correctable errors using the previous $N = n$ measurement vectors is (maximal and) equal to $q_{max} = \lfloor p/2 - 1 \rfloor$. Note that in this case $s = p$, and thus $F$ can only take the value 0 in Algorithm 1. On the other hand,

$$q_{max} = \left\{ \begin{array}{ll} \frac{p - 1}{2} & \text{if } p \text{ is odd} \\ \frac{p}{2} & \text{if } p \text{ is even} \end{array} \right.$$ and thus

$$|K_{Fi}| = \left\{ \begin{array}{ll} 1 & \text{if } p \text{ is odd} \\ 2 & \text{if } p \text{ is even} \end{array} \right.$$ Therefore, for almost all systems with $n$ states and $p$ sensors, if we use $N = n$ past measurement vectors the complexity of Algorithm 1 is $p \cdot 2^{n-1}$ if $p$ is an odd number, or $\frac{p(p-1)}{2} \cdot 2^{n-1}$ if the system has an even number of sensors.

5. EVALUATION

To evaluate conservativeness of the error bound obtained using Algorithm 1, we consider two types of systems – systems with $n = 10$ states and $p = 5$ sensors, and with $n = 20$ states and $p = 11$ sensors. For each system type we randomly generated 100 systems with measurement models satisfying that the rows of the $C$ matrix have unit magnitude and matrices $\Delta$ had elements between 0 and 2. In addition, for each of the 200 systems we evaluated the state-estimation error $\Delta x = \|x_0, \Delta - x_0\|_2$ in 1000 experiments for various attack and noise realizations. Attacks and noise profiles were chosen randomly assuming uniform distribution of the following: (a) The number of attacked sensors between 0 and 2 for systems with 5 sensors, and between 0 and 5 for systems with 11 sensors, (b) Attack vectors on the compromised sensors between −10 and 10, chosen independently for each attacked sensor, and (c) Noise realizations between the noise bounds specified by matrices $\Delta$.

In both simulations and Algorithm 1 executions, we considered the case when window size $N$ is equal to the number of system states (i.e., $N = n$). Comparison between the bounds provided by Algorithm 1 and simulation results are shown in Fig. 2 and Fig. 3. Fig. 2(a) and Fig. 3(a) present histograms of $\|\Delta x\|_2$ errors for all 1000 scenarios for three randomly selected systems. As can be seen, the bound provided by Algorithm 1 is an order of magnitude larger than the average state-estimation error for each system. However, for each system $S$ we are more interested in the ratio between the worst-case observed state estimation error for all 1000 simulations – i.e., $\max_{i=1:1000} \|\Delta x_{ei}\|_2$, and the error bound $MAX_\|\Delta x_e\|_2$ provided by Algorithm 1 for the system. Thus, we consider relative estimation error defined for each system $S$ as

$$Rel_{error_S} = \frac{\max_{i=1:1000} \|\Delta x_{ei}\|_2}{MAX_\|\Delta x_e\|_2}.$$ A histogram of the relative errors for both types of systems are presented in Fig. 2(c) and Fig. 3(b). For the systems with $n = 10$ states the maximal relative error reaches almost 20% of computed bounds, while for larger system (with $n = 20$ states) the maximal relative error is 2% of computed bounds.

However, it is worth noting here that conservativeness of the presented results is (at least partially) caused by the fact that for each system we only considered random initial points, and random uncorrelated attack vectors and noise profiles/modeling errors. Thus, the errors obtained through simulation do not represent the worst-case errors; for each system, to obtain scenarios that result in the worst-case estimation errors it is necessary to derive the corresponding attack vector (and the initial state), which is beyond the scope of this paper.

This is especially illustrated in histograms of relative estimation errors for systems with different size. As in the histograms from Fig. 2(c) and Fig. 3(b), in simulations we observed a decrease in the obtained maximal relative estimation error with an increase in the system size $n$ (and thus increase in the window size $N = n$). One of the reasons is that with the increase of $N$ we increase the number of attack vectors, and due to the random selection of the attack vectors we reduce probabilities to incorporate a worst-case attack. On the other hand, for systems with $n = 1$ and $n = 2$ states we were able to generate initial states and attack vectors for which the bounds from Algorithm 1 are tight – i.e., the error $\|\Delta x\|_2$ is equal to the obtained bounds.

6. CASE STUDY

We illustrate the development framework on a design
(a) Histogram for a system with error bound 41.43
(b) Histogram for a system with error bound 35.74
(c) Histogram of the maximal relative state-estimation error for all 100 system

Figure 2: Simulation results for 1000 runs of 100 randomly selected systems with $n = 10$ states and $p = 5$ sensors.

(a) Histogram for a system with error bound 155.98
(b) Histogram of the maximal relative state-estimation error for all 100 system

Figure 3: Simulation results for 1000 runs of 100 randomly selected systems with $n = 20$ states and $p = 11$ sensors.

of secure cruise control of the LandShark vehicle [1], a fully electric Unmanned Ground Vehicle (UGV) shown in Fig. 4(a). In a tethered mode, the robot can be fully tele-operated from the Operator Control Unit (OCU). However, in our scenario the operator only specifies the desired vehicle speed, while the on-board control has to ensure that all of the safety requirements are satisfied even if some of the sensors are under attack.

Vehicle Modeling
To obtain a dynamical model of the vehicle we used the standard differential drive vehicle model (Fig. 4(b)) [12]. Here, $F_l$ and $F_r$ denote forces on the left and right set of wheels respectfully, and $B_r$ is the mechanical resistance of the wheels to rolling. The vehicle position is specified by its $x$ and $y$ coordinates, $\theta$ denotes the heading angle of the vehicle measured from the $x$ axis, while $v$ is the speed of the vehicle in this direction. The LandShark employs skid steering, meaning that in order to make a turn it is necessary to generate enough torque to overcome the sticking force $S_l$. Therefore, when $\frac{B_r}{2}|F_l - F_r| \geq S_l$ the wheels start to slide sideways (i.e., the vehicle begins to turn). Consequently, if we assume that the wheels do not slip, the dynamical model of the vehicle can be specified as

\[
\dot{v} = \begin{cases} 
\frac{1}{m}(F_l + F_r - (B_s + B_r)v), & \text{if turning} \\
\frac{1}{m}(F_l + F_r - B_r v), & \text{if not turning} 
\end{cases}
\]

\[
\dot{\omega} = \begin{cases} 
\frac{1}{J_t} (B_l (F_l - F_r) - B_l \omega), & \text{if turning} \\
0, & \text{if not turning} 
\end{cases}
\]

\[
\dot{\theta} = \omega
\]

\[
\dot{x} = v \sin(\theta), \\
\dot{y} = v \cos(\theta)
\]

Also, $\omega = 0$ if the vehicle is not turning.

Finally, to estimate the state of the vehicle for cruise control (i.e., its speed and position) we use three sensors — two speed encoders, one on each set of wheels, and a GPS. The GPS provides time-stamped global position and speed, while from the encoders we can obtain the rotation angle (which could be translated into rotational velocity and finally into linear velocity). Note that other sensors can be used to estimate the state of the vehicle; for instance, linear acceleration measurements coming from the IMU, or use optical flow algorithms to com-
pute visual odometry from a camera. However, to illustrate the use (and robustness) of the attack-resilient state estimator we only used the encoders and GPS.

The above model presents a high-level model of the vehicle, describing only the motion equations. However, the forces $F_l$ and $F_r$, which can be considered as inputs to the model, are derived from the vehicle's electromotors and are affected by the motors, gearbox and wheels. Thus, we have also derived a 6-state linear model of this low-level electromechanical system based on the model from [12], which is then used to derive a local state (i.e., velocity) feedback controller that provides the desired $F_l$, $F_r$ levels.

System Architecture

On the LandShark, the CPU that implements the state-estimation and controller procedure is connected to all sensors through independent serial buses, while the motors are connected to the CPU via motor drivers (as presented in Fig. 4(c)). Since the speed of the vehicle is bounded, the attack-resilient state-estimator from (11) can be formulated as a mixed linear integer programming (MILP) problem

$$\min_{\gamma, E, x} \quad 1^T_p \gamma$$

$$-\delta_k \leq y_k - CA^k x - e_k \leq \delta_k, \quad k = 0, \ldots, N - 1$$

$$-\gamma_j \alpha \cdot 1'_N \leq E'_j \leq \gamma_j \alpha \cdot 1'_N, \quad j = 1, \ldots, p$$

where $E'_j$ and $e_k$ denote the $j^{th}$ row and $k^{th}$ column of the matrix $E \in \mathbb{R}^{p \times N}$, respectively. Here, $\gamma = (\gamma_1, \ldots, \gamma_p) \in \{0, 1\}^p$ are binary optimization variables representing, for each sensor $j$, whether the sensor is considered attacked ($\gamma_j = 1$) or safe ($\gamma_j = 0$), and $\alpha$ is a sufficiently large positive constant.

The developed resilient controller is executed on top of Linux OS and the Robot Operating System (ROS) middleware [10]. ROS is a meta-operating system that facilitates development of robotic applications using a publish/subscribe mechanism in which a master superintend every operation. There is a driver associated with each sensor, which takes care of getting time stamped informations from the sensor and publishing this data in the ROS format to the ROS master. The controller written in C++ language subscribes to each sensor measurements (called topics) through the master, and sends inputs to the motor driver in order to maintain the desired cruise control speed.

Experiments

Fig. 5 presents a deployment of the robot during experiments run on a tiled uneven surface and a grass uneven field. From the developed GUI we demonstrate that the robot can reach and maintain the desired reference speed even when one of the sensors is under attack, as shown in Fig. 6. Fig. 6(a) presents speed estimates from the encoders and GPS; each of the sensors has been attacked at some point, with attacks such that their measurements would result in the speed estimate equal to 4 m/s, except in the last period of the simulation when we have switched to an alternating attack on the encoder left. However, as shown in Fig. 6(b) when the attack-resilient controller is active the robot reaches and maintains the desired speed of 1 m/s. On the other hand, if the state estimator is disabled and instead a simple observer is employed (as in the interval between 68 s and 73 s – the shredded area in Fig. 6), even when one of the sensors is under attack the robot cannot reach the desired state (e.g., it can even be forced to stop).[10]

Robustness Analysis

All ROS nodes are executed in the run-to-completion manner. Thus, although the execution period for the controller node is 20 ms, other instantiated nodes might affect its execution (i.e., the controller might run with a

---

9 Since the robot cannot obtain speed larger than 20 mph, all sensor measurements larger than the value are obtained from compromised sensors and thus can be discarded. Hence, we can assume that elements of attack vectors can not be larger than the maximal speed.

10 Videos of the Landshark experiments can be seen at http://www.seas.upenn.edu/~pajic/research/CPS_security.html
variable period). Each sensor has its own clock and all measurements are time-stamped before being transmitted to the controller. However, since relative changes in obtained measurements are used, time synchronization error between sensors does not accumulate. In addition, there is a huge discrepancy between sensors’ sampling jitters. For example, encoders’ sampling jitters are bounded by 100 $\mu$s (as shown in Fig. 7), while GPS has highly variable jitter with maximal values up to 125 ms. Therefore, it is not possible to use the idealized discrete-time model from (8), but rather the full input compensation has to be done as in (6) and (7), before the state-estimator is executed.

Consequently, a bound on GPS error is determined from manufacturer specifications, worst-case sampling jitter and synchronization error, and is experimentally validated to be $\delta_{k,1} \leq 0.4$ m/s. On the other hand, each encoder has 192 cycles per revolution, resulting in a measuring error of 0.5%. Thus, since the maximal achievable vehicle speed is 20 m/s, we have that for both encoders $\delta_{k,2} = \delta_{k,2} \leq 0.1$ m/s. For these values Alg. 1 provides a state-estimation error bound of 0.72 m/s.

Note that the conservativeness of the bound was mostly caused by the large worst-case GPS sampling jitter.

7. CONCLUSION

In this paper, we have considered the problem of attack-resilient state estimation in systems with noise and where the exact model of the system dynamics is not known. We have described a $l_0$-norm based state estimator that can be used for these systems, and showed that the attacker cannot exploit the noise and limitations in model accuracy to destabilize the system. Furthermore, we have provided an algorithm to derive a bound for the state estimation error caused by noise and modeling errors, and presented a procedure to map these bounds into a set of implementation specifications imposed on the underlying platform. Finally, we have illustrated our approach by designing an attack-resilient constant speed cruise controller for unmanned ground vehicle.

We have shown that the presented $l_0$-norm optimization procedure for state estimation can be formulated as a mixed integer linear program. Although there exist efficient MILP solvers, MILPs are effectively NP hard.
Therefore, the natural next step would include transforming the $l_0$ state estimator into a convex program based on $l_1/l_r$ optimization (e.g., $r = 2$) as done in [8]. In this case, providing a bound for the state-estimation error when the $l_1/l_r$ convex relaxation is used will be an avenue for future work.

8. APPENDIX

Proof of Theorem 2. The proof uses a generalization of the proof of Theorem 2 from [14]. We start by assuming that there exist a non-vertex point \( x_0 = \arg \max_{x \in Q} f(x) \), and that for every vertex point \( x \in Q \), \( f(x) < f(x_0) \). Note that \( x_0 \) exists since the polyhedron \( Q \) is bounded. Since \( x_0 \) is not a vertex of \( Q \in \mathbb{R}^n \), less than \( n \) constraints used to specify \( Q \) are active at \( x_0 \) [4]; we denote by \( d (d < n) \) the number of active constraints in \( x_0 \) and denote them as \( a_i^T x_0 = b_i, i \in \{ j_1, ..., j_d \} \).

Thus, the vectors \( a_i, i = j_1, ..., j_d \) belong to a proper subspace of \( \mathbb{R}^n \) and there exists \( d \in \mathbb{R}^n \) such that \( d \neq 0 \) and \( a_i^T d = 0 \), for all \( i = j_1, ..., j_d \).

Consider vectors \( x_0 + \epsilon d \) and \( x_0 - \epsilon d \) for any \( \epsilon > 0 \). From the definition of \( x_0 \), \( f(x_0) \) \( \geq f(x_0 + \epsilon d) \) and \( f(x_0) \geq f(x_0 - \epsilon d) \), meaning that

\[
\frac{1}{2}(f(x_0 + \epsilon d) + f(x_0 - \epsilon d)).
\]

However, since \( f \) is a convex function, we have that for all \( \epsilon > 0 \)

\[
f(x_0) = f(x_0 + \epsilon d) = f(x_0 - \epsilon d).
\] (29)

Because \( Q \) is bounded, there exists some \( \epsilon_1 > 0 \) for which additional constraint used to specify \( Q \) gets activated either at point \( x_0 + \epsilon_1 d \) or \( x_0 - \epsilon_1 d \). Let us denote that point as \( x_1 \). Note that from [29], \( f(x_0) = f(x_1) \) and that at \( x_1 \) at least \( d+1 \) constraints used to specify \( Q \) are active (because all previously active constraints will remain active).

The above procedure can be repeated until we reach a point \( x_r \) \( (r \leq n - d, \text{i.e., in at most } n - d \text{ steps}) \) in which \( n \) constraints are activated and where \( f(x_r) = f(x_0) \). However, since in \( x_r \) exactly \( n \) constraints used to specify \( Q \) are active we have that \( x_r \) is a vertex of the polyhedron \( Q \), and therefore there exist an optimal point of the maximization problem that is a vertex of the polyhedron \( Q \). □

9. REFERENCES